

# Eine nicht-parametrische Glättungsmethode, die Randinformationen berücksichtigt

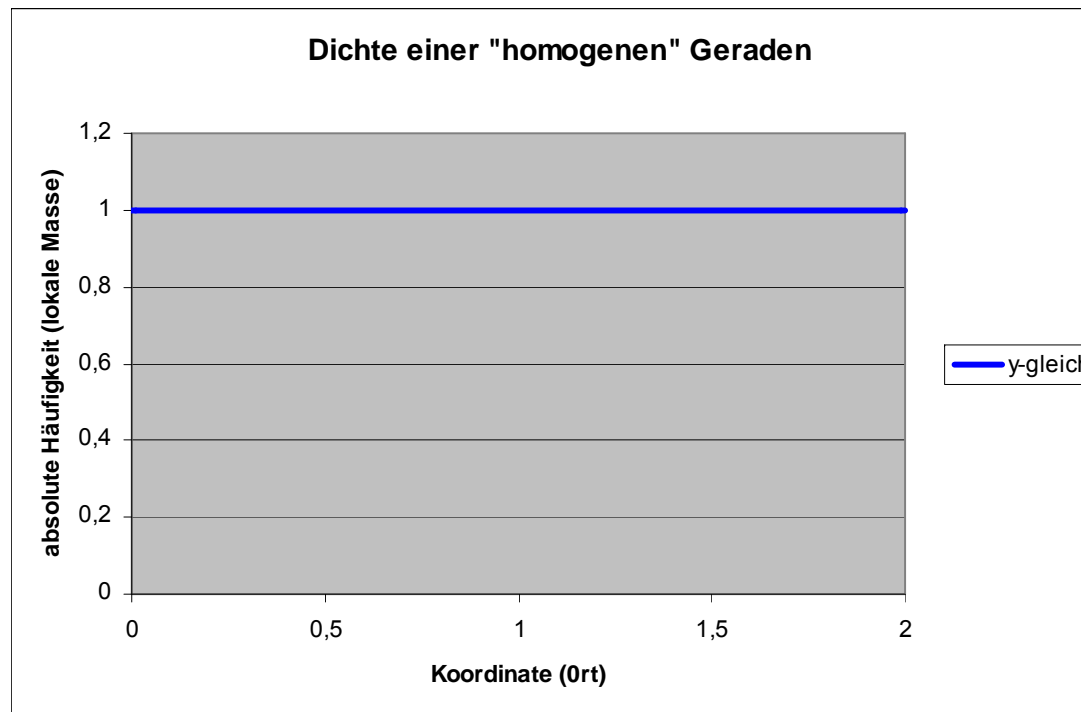
Andreas Fröhlich, Thomas Selhorst, Christoph Staubach  
FLI-Wusterhausen

AG – Räumliche Statistik  
Lübeck, Dezember 2009

# Gliederung

- Dichteschätzungen
  - Probleme am Rand
  - ein minimaler Anspruch
  - ein Lösungsvorschlag
- Theoretische Betrachtungen
  - mit allgemeinen praktischen Bezügen
- Beispiele
  - Dichteberechnungen aus der Praxis und ihre individuellen Interpretationen
- offene Probleme

# Dichte einer homogenen Geraden



Eigenschaften des Objekts: 1. Homogenität,

2. Objekt erfüllt vollständig den Raum

FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Berechnung der Dichte eines homogenen Objektes $O$

- Bezeichne
  - $V(x)$  eine lokale Umgebung um den Punkt  $x$  des Objektes  $O$  und
  - $m(V(x))$  die von der Umgebung  $V(x)$  eingeschlossene Masse sowie
  - $i(V(x))$  den Inhalt der Umgebung  $V(x)$

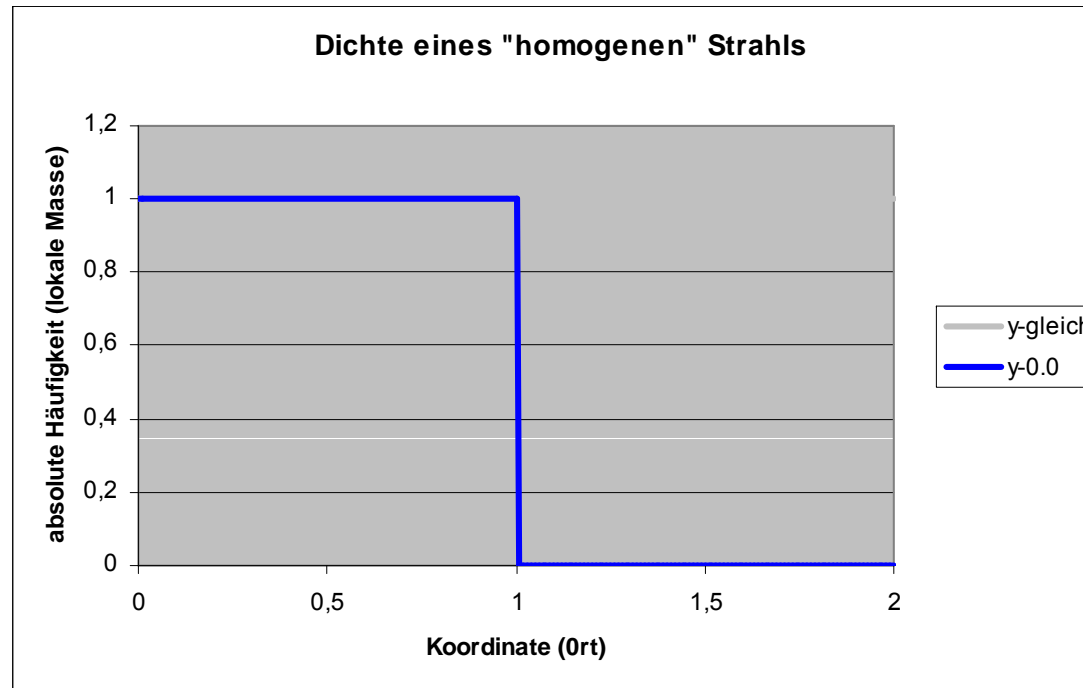
so sei per definitionem

- die Dichte  $d$  des Objekts  $O$  in  $x$
  - $d(x) = m(V(x)) / i(V(x))$
- Anmerkung:
  - Die „Größe“ und „Form“ der lokalen Umgebung  $V(x)$  ist für die Dichtebestimmung in  $x$  ohne Bedeutung!

# Invarianz der Dichte homogener Objekte

- Die Dichte ist invariant bzgl.  $V(x)$ 
  - d.h.  $d(x)$  ist stets gleich groß unabhängig von der gewählten
    - Größe von  $V(x)$ ,
    - Form von  $V(x)$  und
    - Positionierung von  $V(x)$
- **Minimaler Anspruch an Dichtebestimmungen:**
  - **! homogene** Objekte  $O$  besitzen konstante Dichte,  $d(x)=\text{const.}$  für alle  $x$  aus  $O$  !

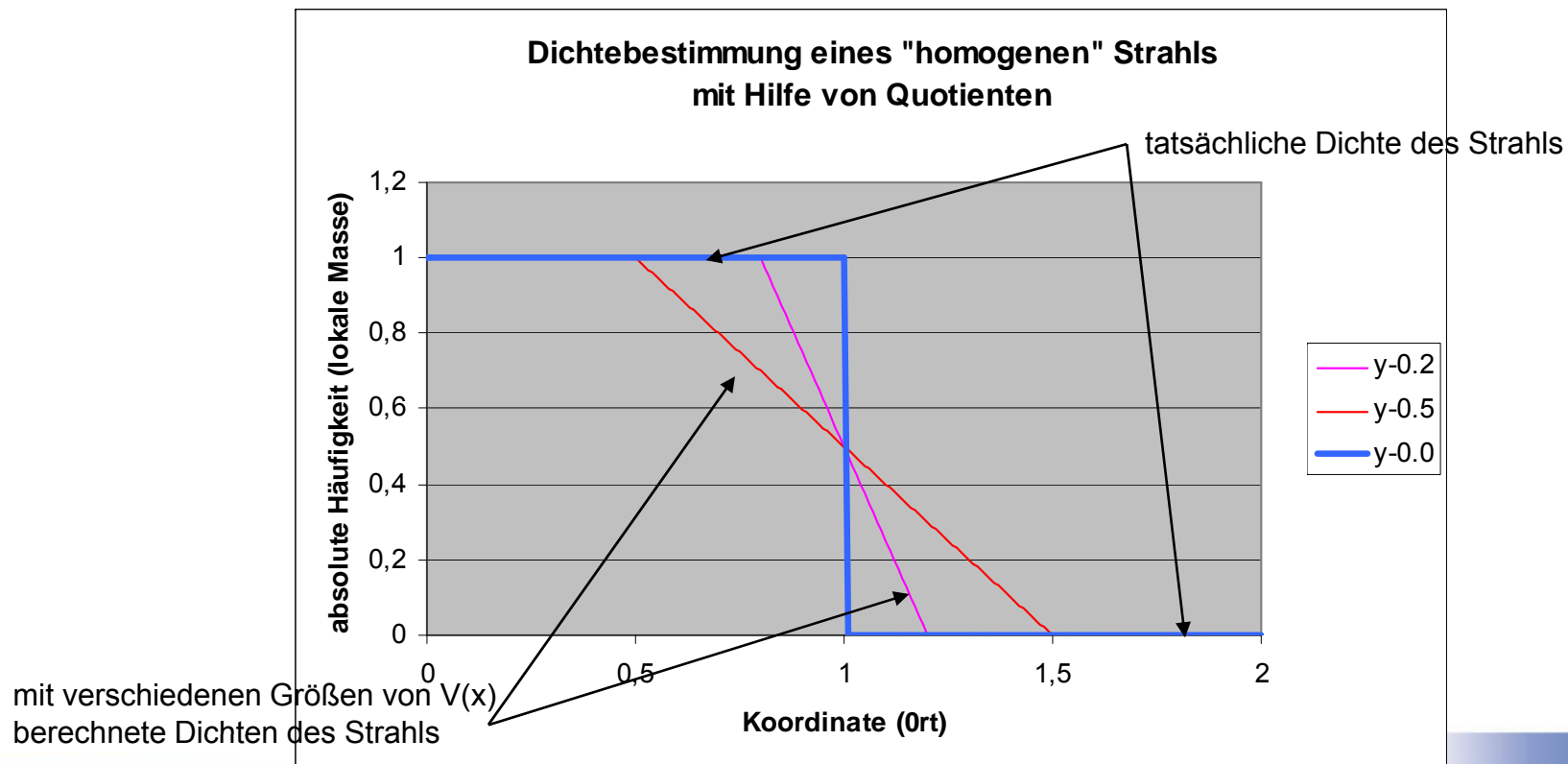
# Dichte eines homogenen Strahls



- Eigenschaften des Objekts:
1. Homogenität des Objekts,
  2. Objekt ist echte Teilmenge des Raumes

# Berechnung der Dichte eines homogenen Strahls

- unter Verwendung von  $V(x)$  von „konstanter“ Form und Positionierung bzgl.  $x$  und verschiedenen Größen



# Berechnung der Dichte eines homogenen Strahls

- Besonderheit (an Stellen „nahe“  $x=1$ ) am Rand von  $O$ 
  - $V(x)$  ist als Teilmenge von  $O$  zu wählen (insbesondere für den homogenen Fall) und/oder
  - Umgebung  $V(x)$  von  $x$  muss vom Inhalt  $i$  (=Größe - Inhalt) hinreichend „klein“ gewählt werden

somit ist

$$d(x) = \lim_{i(V(x)) \rightarrow 0} \frac{m(V(x))}{i(V(x))}$$

eine verallgemeinerte Definition  
der Dichte des Objekts  $O$  in  $x$ ;  $x$  innerer Punkt von  $O$



# Notwendige Bedingung einer korrekten Dichteberechnung im (topologischen) Inneren

- Wird statt des korrekten Limes

$$\lim_{i(V(x)) \rightarrow 0} \frac{m(V(x))}{i(V(x))}$$

ein Quotient zur näherungsweisen Dichteberechnungen verwendet,

dann sind Randinformationen von Objekten  $O$  mit zu berücksichtigen

# Numerisches Problem

- Endlichkeit der Messwerte zum Objekt  $O$  lässt für beliebige  $x \in O$  im Allgemeinen keinen praktikablen Limesübergang zu
- **\* d.h.  $V(x)$  muss eine nicht leere Menge von Messwerten (Beobachtungen) des Objekts  $O$  enthalten, so dass  $m(V(x))$  kontrollierbar bestimmbar bleibt**

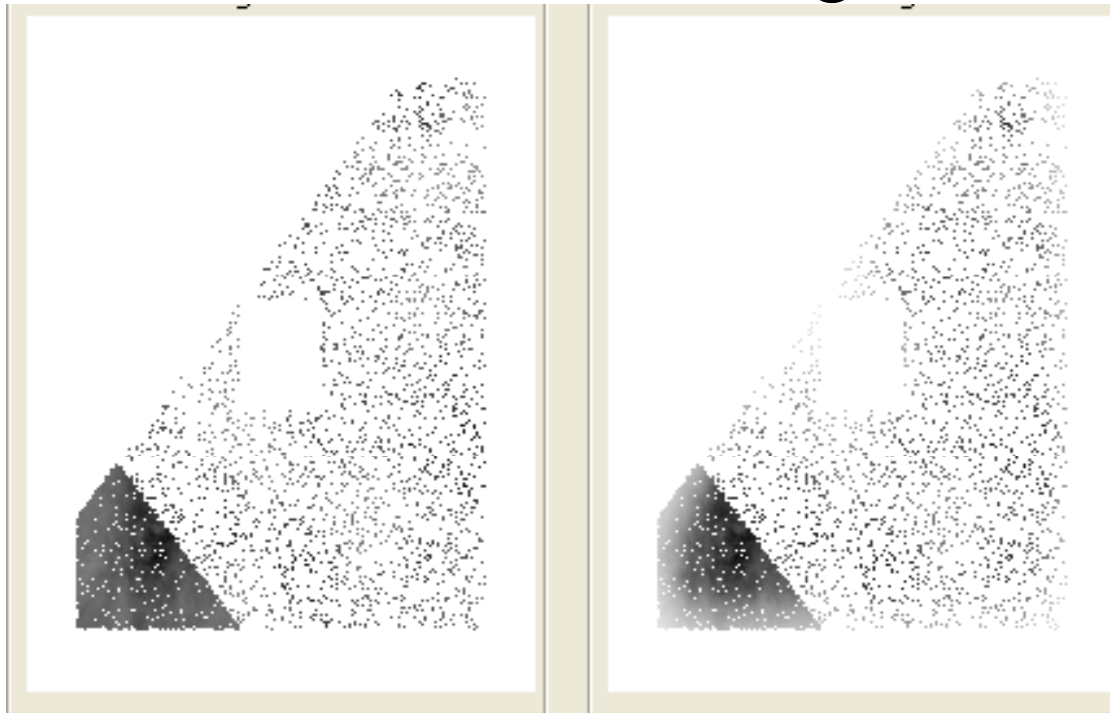
# Gewünschte Eigenschaften von $V(x)$

- Da  $d(x)$  eine lokale Größe ist, sollte  $V(x)$  eine „gut an  $x$  gebundene“ Umgebung darstellen
- Für nicht-homogene Objekte  $O$  sollte  $V(x)$  möglichst klein gewählt werden.  
der Grenzübergang ist wegen Bedingung \*  
(numerisch) praktisch unmöglich

**-> Notwendigkeit der Wahl eines geeigneten  
„Glättungsparameters“, der  $V(x)$  beschreibt**

# Elementare Forderung einer Dichteschätzung

- Reproduzierbarkeit homogener Dichten



modifiziert Methode

Standard

FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Methode ohne zusätzliche Annahmen

- k-nächster Nachbarn (in metrischen Räumen  $R$ )  
( $k$  – entspricht einem notwendigen Glättungsparameter)

$O \ni x \mapsto O_k(x) = V_k(x) \cap O$ , wobei

$V_k(x) = \{y \in R \mid d(y, x) \leq r_{k,x} \wedge r_{k,x} = \min\{r \mid \text{card}(\{m \in M \cap V(x)\}) \geq k\}\}$ , wobei

$M$  die Menge der Messwerte bezeichne .

# Die Dichteschätzung

- Jedem beliebigen Element  $x \in R$  wird wie folgt ein Dichtewert  $d_k(x, O)$  in Abhängigkeit des Glättungsparameters  $k$  zugeordnet:

$$R \ni x \mapsto \frac{k}{\int_{V_k(x)} I_O dy}$$

- Die Inhaltsbestimmung  $i$  von  $V_k(x) \cap O$  ( $= \int_{V_k(x) \cap O} dy$ ) mit  $I_O$  als Indikatorfunktion von  $O$  erfolgt stochastisch durch eine gleichverteilte ZFG auf dem Objekt  $O$

# Wichtige Begründungen und unbeantwortete Fragen

- Für euklidische Räume
  - Warum Umgebungen  $V(x)$  von  $x$  als  $n$ -dimensionale Kugeln
    - Bewegungsinvarianz des Inhaltes - somit Unabhängigkeit von der Lage des Koordinatensystem
- Wie sehen Verallgemeinerungen auf allgemeine metrisierbare Räume aus?

(Theorem über die Metrisierbarkeit top. Räume,  
P.S. Uryson (1925))

Ein top. Raum mit abzählbarer Basis ist genau dann metrisierbar, wenn er normal ist.

# Was zur Berechnung benötigt wird

- Rel. Datenbanksystem zur Speicherung von
  - Messwertmenge,
  - Indikatorfunktion (-menge) des Objekts  $O$ ,
  - Definitionsmenge für Dichtefunktion
  - Resultate der Dichtefunktion
- Programm für die Berechnung (Client)



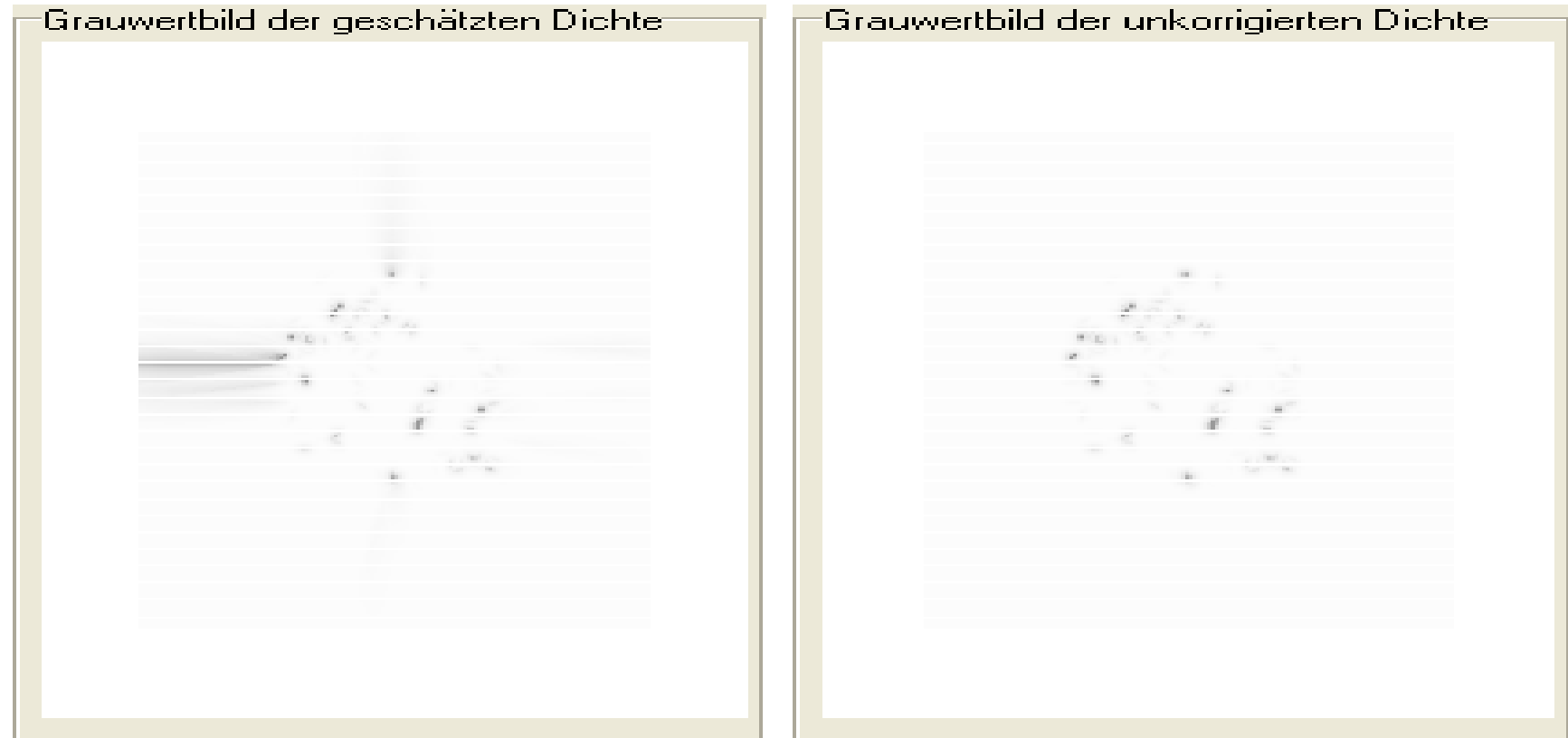
# Vor- und Nachteile der gegenwärtigen Realisierung dieser Methode

- Nachteile:
  - SQL-Datenbank
  - SQL-Datenbankmanagementkenntnisse
  - lange Rechenzeiten
- Vorteile:
  - Möglichkeit der Berechnung von Zerlegungen von Gebieten (symmetrische Differenzen)
  - Parallele Berechnungen
  - Einsicht in alle relevanten Zwischenergebnisse
  - Übernahme der Ergebnisse in GIS-Systeme

# Alle benötigten und verfügbaren Daten auf einen Blick



# Interpretation 1: Identifikation von Randregionen mit erhöhtem Risiko



modifizierte Methode der Schätzung vs. Standardmethode

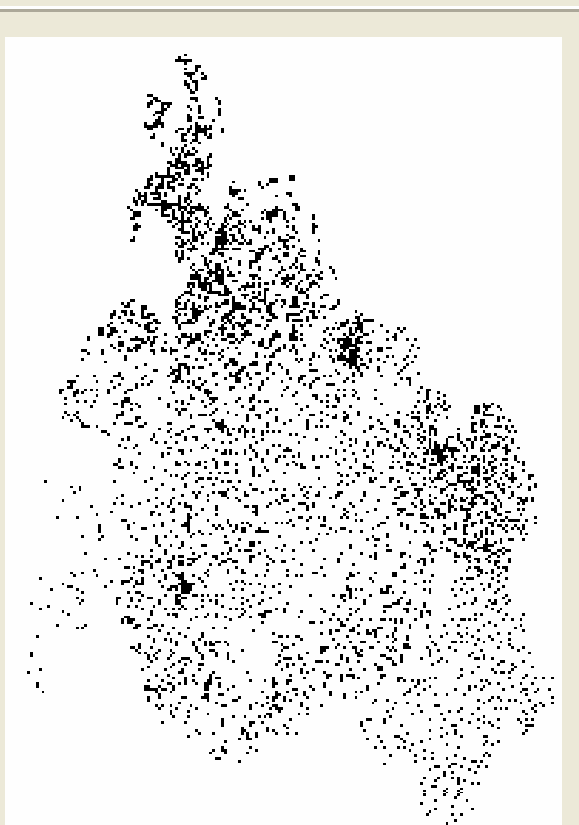
FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

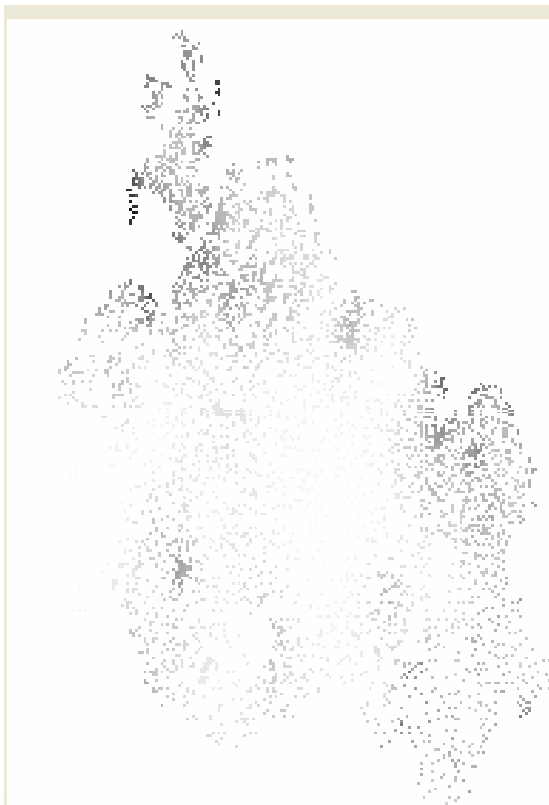
Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Interpretation 2: Dichte von Haltungen bzgl. Landfläche

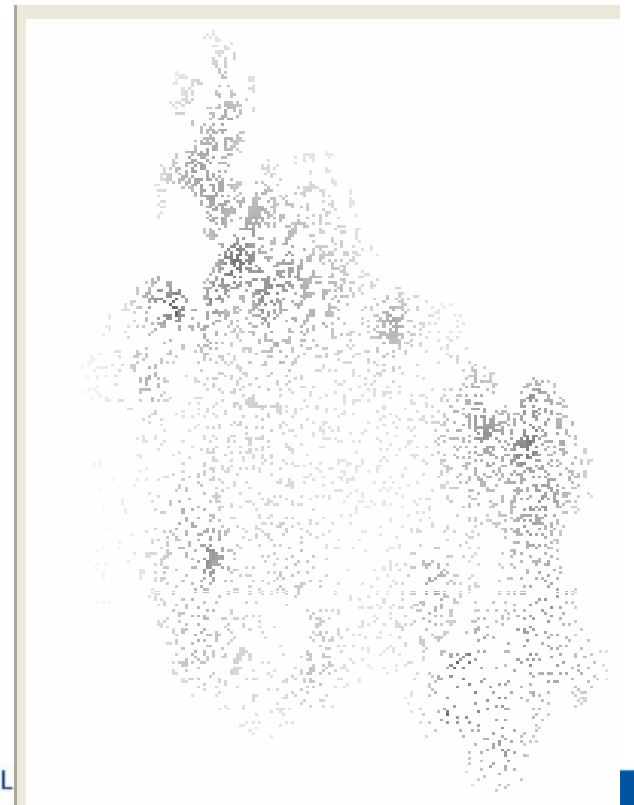
Positionen der  
Haltungen



Dichte der Haltungen  
mit Randinformation



Dichte der Haltungen  
Standard

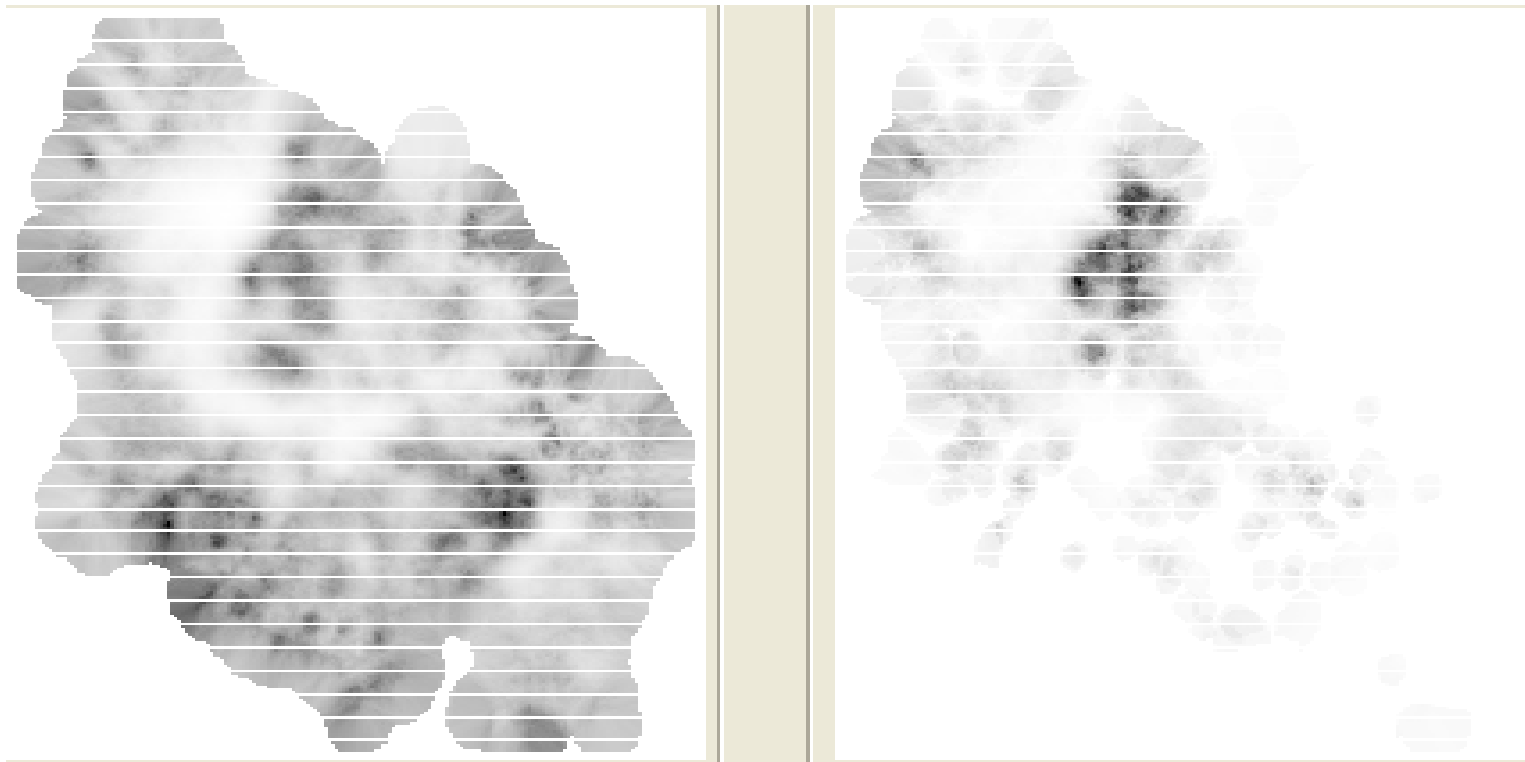


:FFL

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Interpretation 3: Wohin ist eine weitere Ausbreitung wahrscheinlich



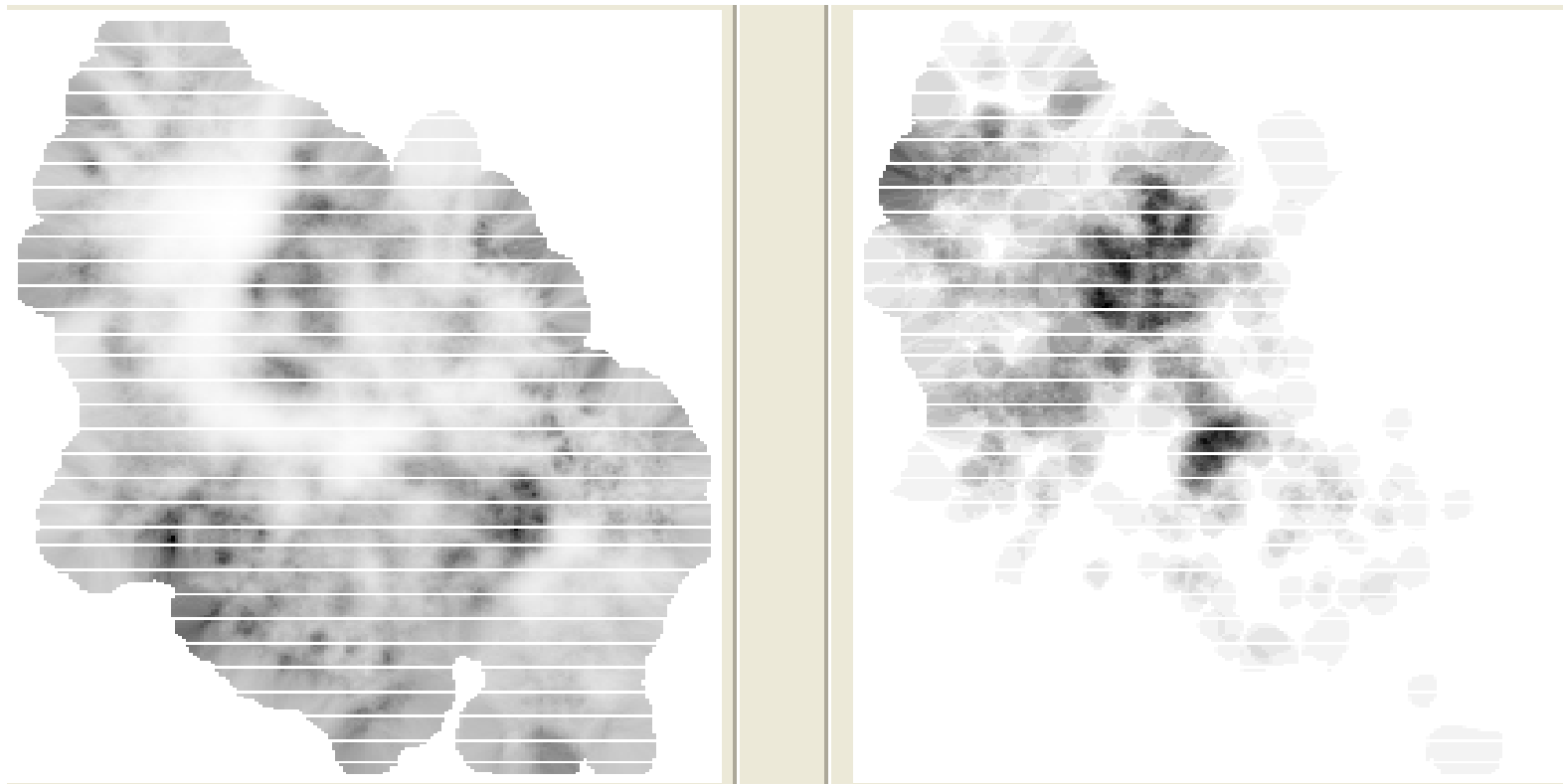
Lokale Beobachtungsintensität vs. Dichte positiver Fälle

FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Interpretation 4: Sperrgebiete und Beobachtungszonen



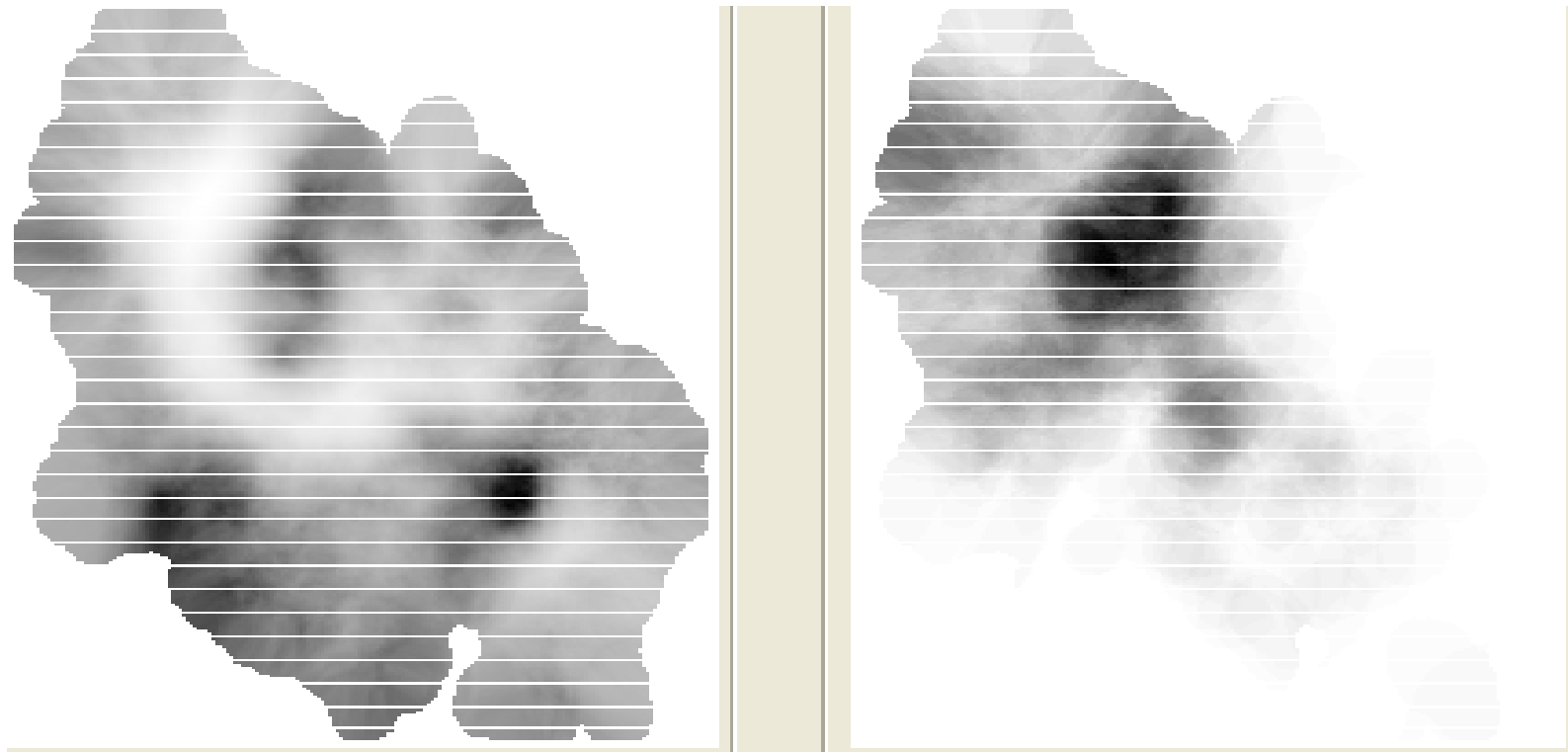
Lokale Beobachtungsintensität vs. lokale Prävalenz

FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Interpretation 4: bei stärkerer Glättung



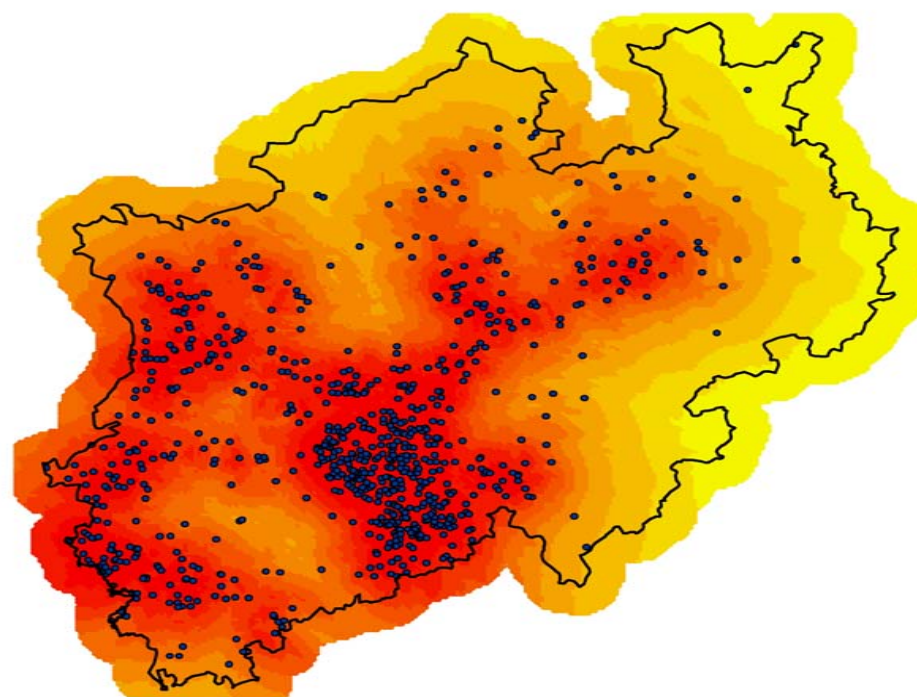
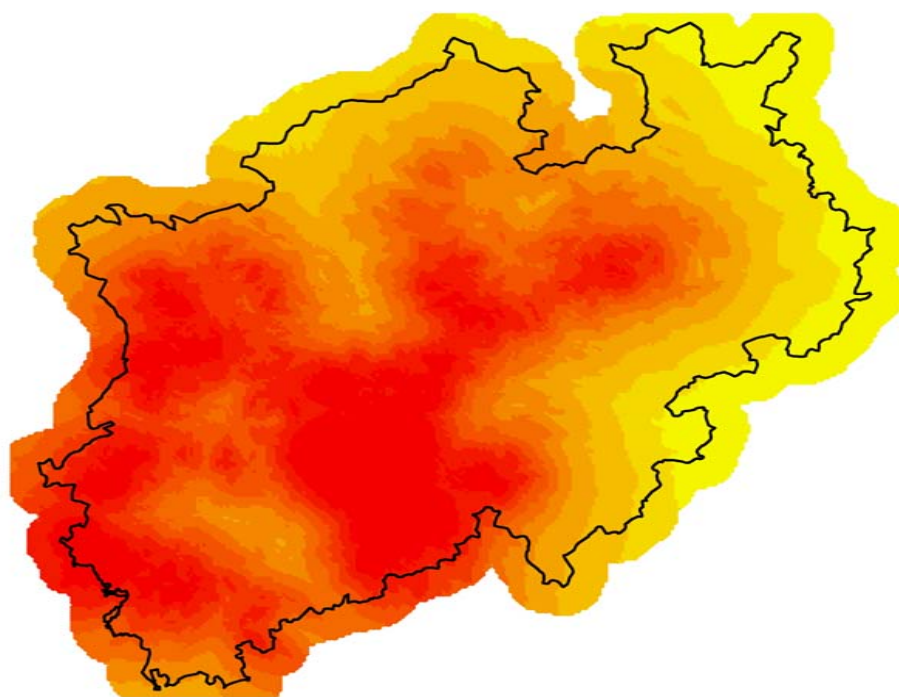
Beobachtungsintensität vs. Prävalenz

FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health

# Interpretation 5: Verschnitt mit GIS-Informationen



Bestandsdichte der Wiederkäuer vs.

BT-Fälle und Bestandsdichte der Wiederkäuer

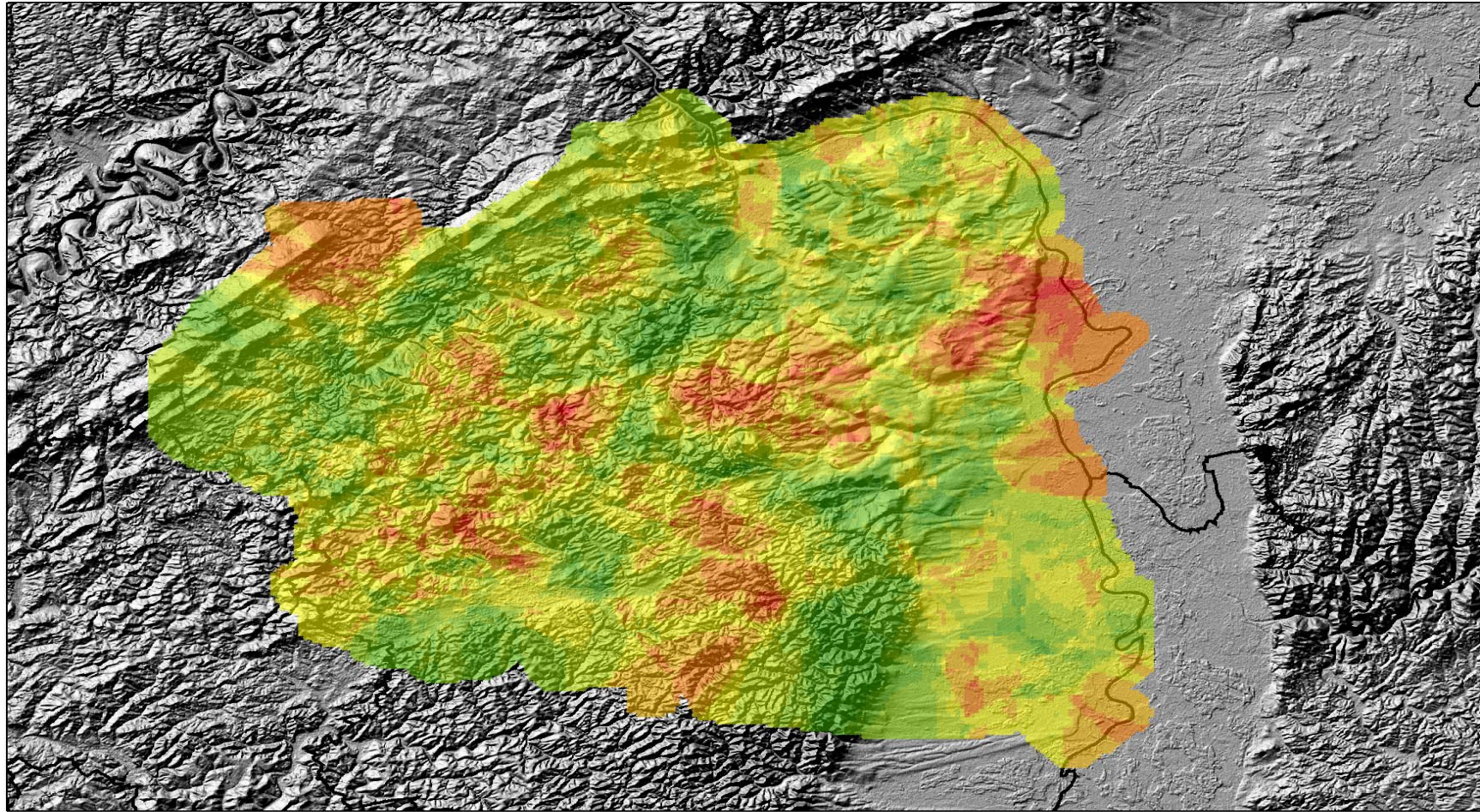
FRIEDRICH-LOEFFLER-INSTITUT

**FLI**

Bundesforschungsinstitut für Tiergesundheit  
Federal Research Institute for Animal Health



# Seroprävalenzen Tollwut bzgl. gegebener Untersuchungsfläche



# Ausblick

- Konstruktion geeigneter Tests zur objektiven Beurteilung von Unterschieden verschiedener Dichten



# Danksagung

- Wir bedanken uns bei
  - den Bundesländern für die Kooperation bei der Bereitstellung der Urdaten
  - Dr. Staubach für die Anfertigung der GIS-Karten zur Blauzungenkrankheit
  - den Frauen Kämer, Kliemt und Teske für die technische Unterstützung bei der Arbeit mit dem GIS-System