Die Erzeugung datenschutzkonformer, hochauflösender Rasterkarten

5. Dresdner Flächennutzungssymposium

Markus Dießelmann, Dr. Gotthard Meinel
Leibniz-Institut für ökologische Raumentwicklung, Dresden

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung

Übersicht

1. Hintergrund
2. Rasterkarten
3. Datenschutzdiskussion
4. Die Erzeugung von Mischrasterkarten
5. Fazit

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

1. Hintergrund

- Von privatwirtschaftlicher Seite werden zunehmend Individualdaten mit Standortbezug erhoben
- Motiviert durch steigende Anforderungen in der Logistik und Navigation, Geomarketing, Standortanalysen etc.
- Im Rahmen des Geomarketing vielfältige Erhebungs- und Anwendungsmöglichkeiten

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

1. Hintergrund

- Klassifizierung von Adressen in soziale Schichten an (bspw. Bürgerliche Mitte oder prekäres Milieu) inkl. Visualisierung
- Mittels Informationen über die Kreditwürdigkeit wird ein (Geo)scoring erstellt, welches über die Zahlungsmoral von Haushalten Auskunft gibt
- Durch vor Ort Begehungen wird der Zustand und der Wert von Häusern erhoben und zur Datenergänzung genutzt
- Befragung von Haushalten nach Konsumverhalten und Lifestyle

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

1. Hintergrund

- Die Daten werden zunehmend mit Geodaten verknüpft und in Kartenform veröffentlicht
- Maßstäbe der administrativen Gebietsstrukturen in vielen Fällen zu klein
- Visualisierung und Bewertung auf Grundlage von Gemeindegrenzen, PLZ-Gebieten oder Baublöcken jedoch mit Interpretationsnachteilen

→Nutzung von Rasterkarten

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

2. Rasterkarten

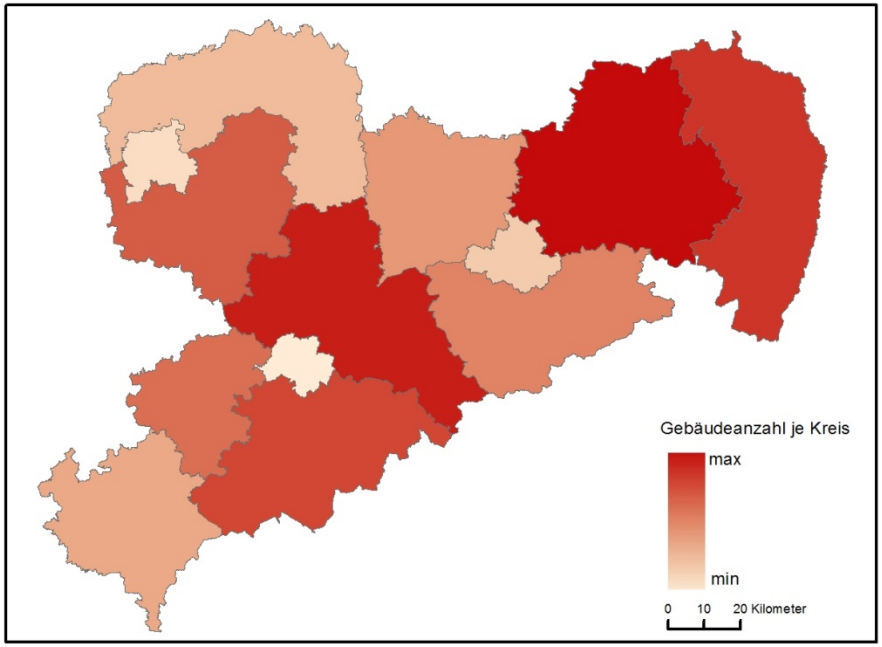
- Schon früh wurde auf die Nachteile von administrativen Grenzen als Bezugsflächen in der thematischen Kartographie hingewiesen (bspw. Arnberger 1977, Witt 1970)

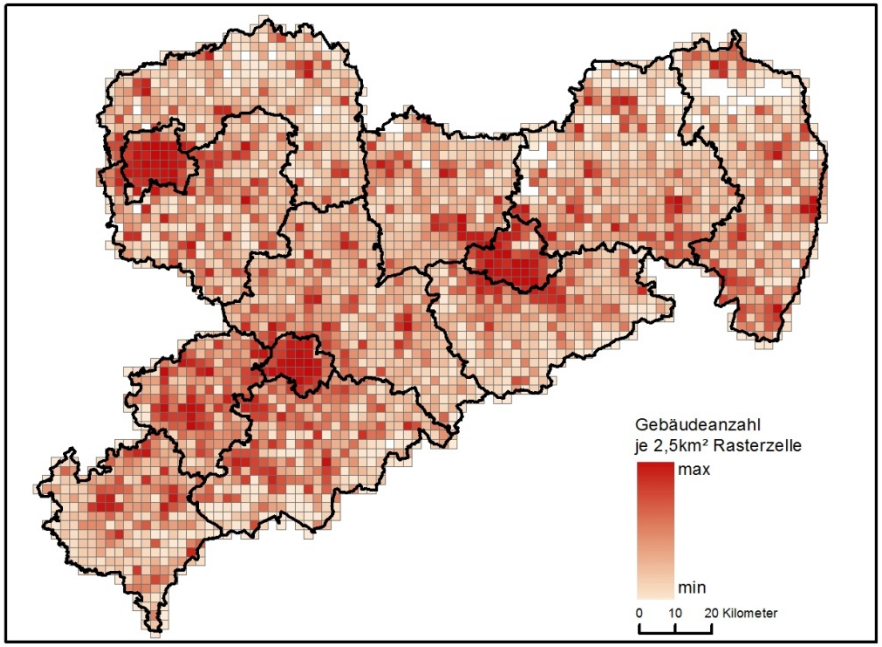
- Bspw. betrifft dies die Ungleichheit der administrativen Einheiten (Größe, Verteilung der Ausprägung…) und die Abgrenzungskriterien

- Im Rahmen der Vergleichbarkeit und leichten Verarbeitbarkeit, bieten sich geometrische Raster (Gitternetze) zur Raumgliederung an

2. Rasterkarten

- Homogene Gliederungseinheiten gewähren eine räumliche





TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung

3. Datenschutzdiskussion

- Datenschutzrechtliche Bedenken bei der Nutzung von Geodaten sind dann angebracht, wenn es sich um personenbezogene Daten nach § 3 Abs. 1 des Bundesdatenschutzgesetzes (BDGS) handelt:

  „Personenbezogene Daten sind Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener)“

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

3. Datenschutzdiskussion

- Diskussion: Wann weisen Geodaten einen Personenbezug auf?
- Abgrenzung zwischen reinem Sachbezug und Personenbeziehbarkeit
- In der Literatur, der Praxis und der Rechtsprechung gibt es keinen Konsens darüber, wann Geodaten einen Personenbezug aufweisen

3. Datenschutzdiskussion

- Ein Personenbezug lässt sich nach § 3 Abs. 6 BDSG durch das Anonymisieren ausschließen:

  „Anonymisieren ist das Verändern personen-bezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können.“

- Somit können die Daten nach gewissen Kriterien auch aggregiert werden

3. Datenschutzdiskussion

- In der Bundesrepublik Deutschland gibt es noch keine verbindlichen Datenschutzgrenzwerte

- Empfehlungen wurden im Rahmen des 3. Geo-fortschrittsbericht der Bundesregierungen gegeben:

  - Karten mit einem Maßstab kleiner als 1 : 5.000
  - Luftbilder mit einer geometrischen Auflösung größer als 20 cm
  - einer größer als auf 100 m x 100 m gerasterten Fläche
  - mindestens auf vier Haushalte aggregierten Informationen

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

3. Datenschutzdiskussion

- Untersuchungen nutzen verschiedene Grenzwerte bspw.:
  - Pilotprojekt „Kleinräumiges Monitoring“ des Regionalverbandes FrankfurtRheinMain erlaubt die Veröffentlichung in Karten, wenn pro Rasterzelle mind. 3 Personen pro 100m x 100m Rasterzelle vorhanden sind
  - Der „Düsseldorfer Kreis“ fasst personenbezogene Daten zu einer Haushaltszelle von mind. 5 Haushalten zusammen (statistisch entspricht dies 9 Personen)
  - Grenzwerte in einer Studie des „Unabhängigen Landeszentrums für Datenschutz Schleswig-Holstein“ (ULD)

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4. Die Erzeugung von Mischrasterkarten

Optionen zur Erzeugung von Datenschutzkonformen Rasterkarten:

- Unterbesetzte Rasterzellen werden gelöscht
- Unterbesetzte Rasterzellen werden als solche gekennzeichnet und verbleiben ohne Werteausprägung im Datensatz
- Unterbesetzte Rasterzellen werden mit benachbarten Rasterzellen aggregiert, um den festgelegten Grenzwert zu erreichen

→ Durch die Aggregation von Rasterzellen wird aus der Rasterkarte eine Mischrasterkarte erzeugt

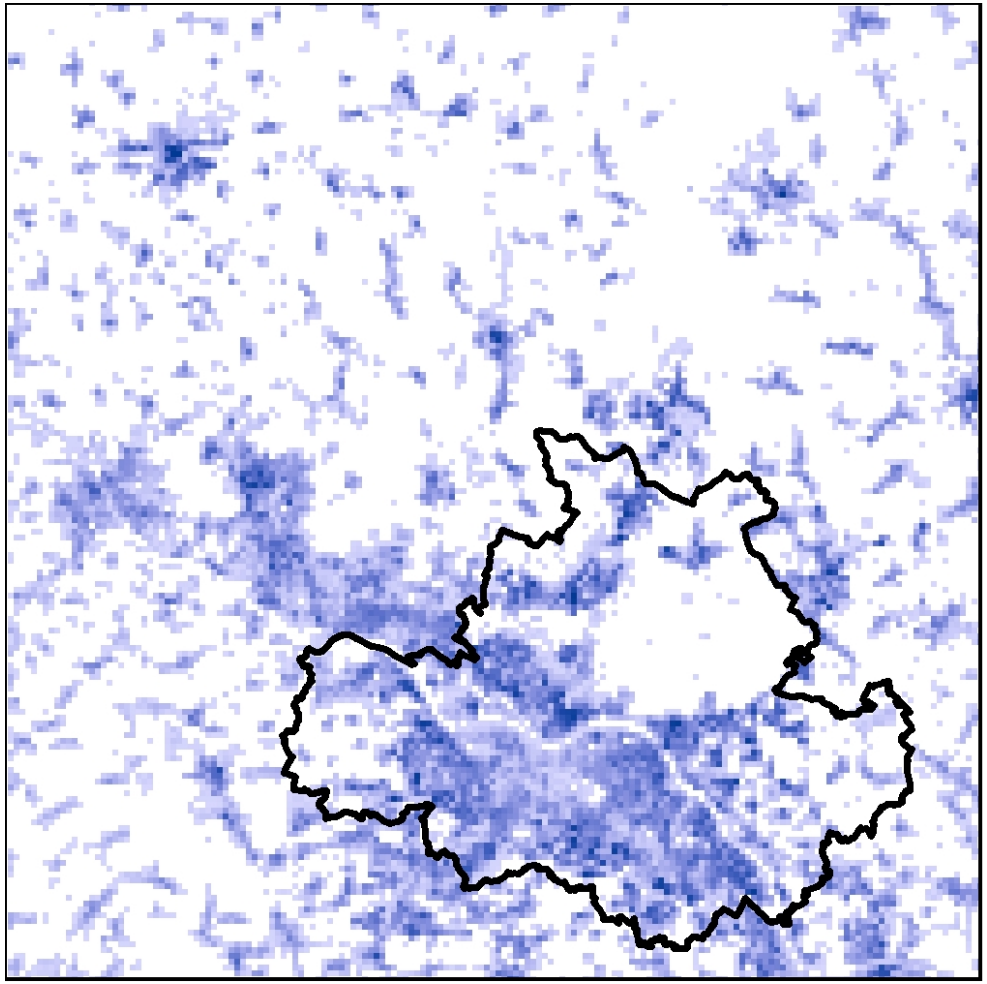
TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4. Die Erzeugung von Mischrasterkarten

Anforderungen an die Erzeugung von Datenschutz-konformen Rasterkarten:

- Die Rasterkarte darf keine Rasterzellen mehr besitzen, deren Fallzahl zu gering ist
- Aggregationskriterium: „So detailliert wie möglich, so stark aggregiert wie (datenschutzrechtlich) nötig“ (Strobl, 2005)
- Die Ansätze sollen sowohl für dicht besiedelte Gebiete als auch für weniger dicht besiedelte Gebiete anwendbar sein
- Es sollen alle thematischen Informationen der Rasterkarte in der Mischrasterkarte enthalten sein

4. Die Erzeugung von Mischrasterkarten



<u>Ausgangsrasterkarte:</u>

256 x 256 Rasterzellen

250m Kantenlänge

Aggregationsgrenzwert >= 4

Das Attribut Gebäudeanzahl wird nur zur Veranschaulichung genutzt, es ist Datenschutz-rechtlich unbedenklich!

Gebäudeanzahl pro Rasterzelle

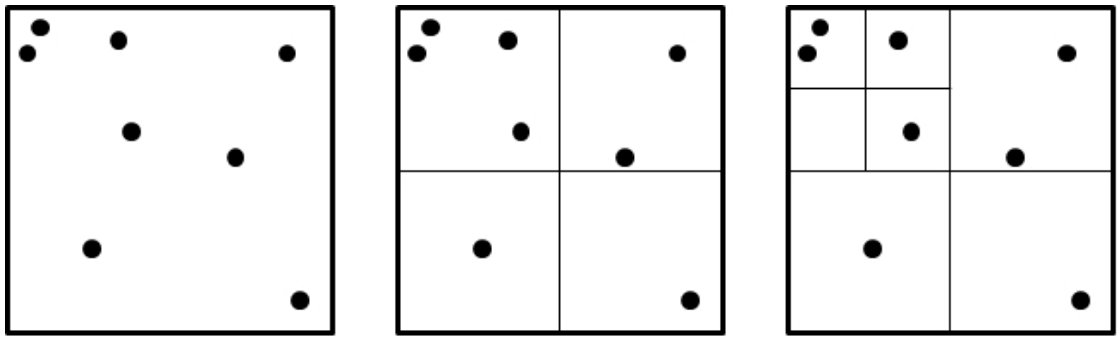
TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung

4.1 Quadtree-basierter Ansatz

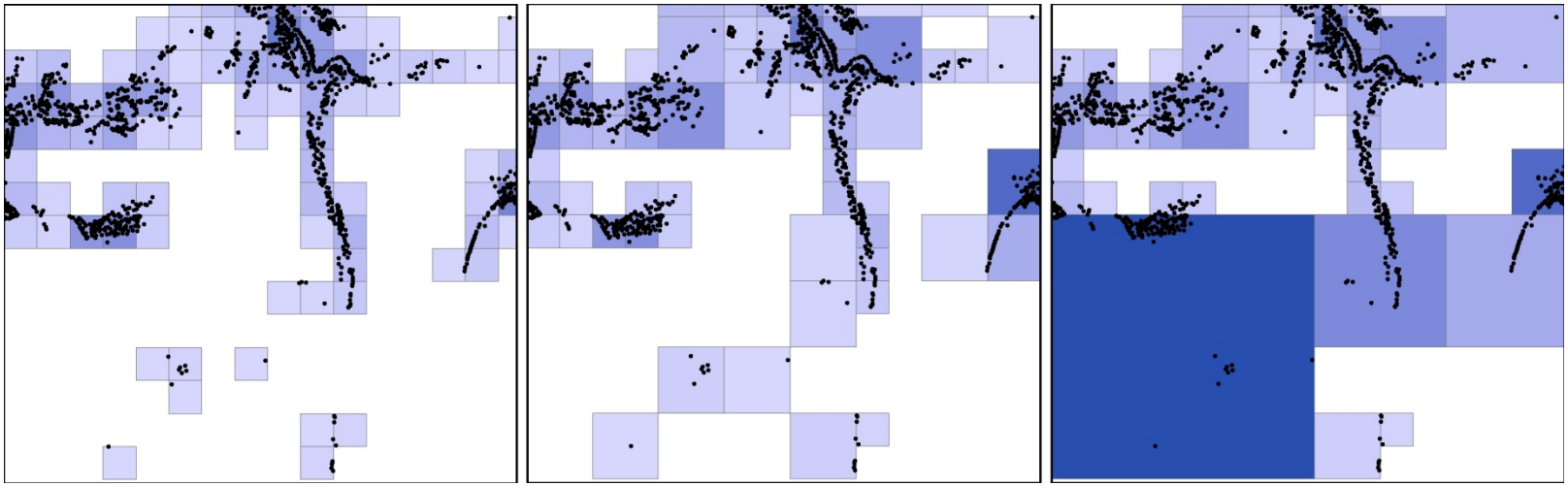
- Von Strobl (2005) als Hierarchische Aggregation vorgestellt
- Er basiert auf dem Quadtree-Prinzip (Region Quadtree)
- Mittels der Quadtrees lässt sich der Raum geometrisch rekursiv unterteilen



TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.1 Quadtree-basierter Ansatz

- Der Quadtree-basierte Ansatz unterteilt den Raum jedoch nicht, er aggregiert ihn

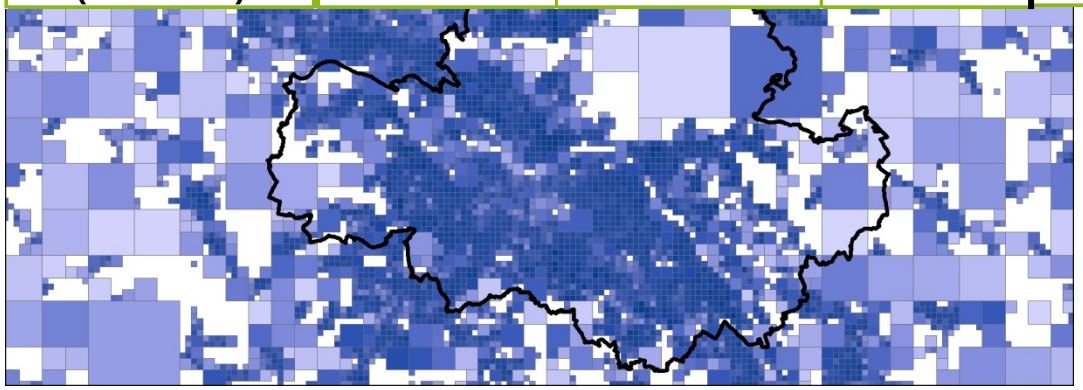


Zusammenfassung von Rasterzellen nach dem Region Quadtree Prinzip (Grenzwert ≥ 4). Von links nach rechts: 250m Rasterzellen, 500m Rasterzellen und 2000m Rasterzellen

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.1 Quadtree-basierter Ansatz

	Dresden Grenzwert ≥ 4			Hamburg Grenzwert ≥ 4		
	Ausgangskarte	Quadtree Mischrasterkarte	Auflösungs-verlust	Ausgangskarte	Quadtree Mischrasterkarte	Auflösungs-verlust
Anzahl Rasterzellen	65536	27751	-37785	65536	35140	-30396
Durchschnittliche Gebäudeanzahl pro Rasterzelle	7,04	16,62	9,58	16,18	30,18	14
Rasterzellengröße (Mittelwert)	250	294,4	44,4	250	287,89	37,89



Gebäudeanzahl pro Rasterzelle

Normiert auf Rasterzellgröße

TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung

4.1 Quadtree-basierter Ansatz

Zusammenfassung:

- Der Quadtree-basierte Ansatz erfüllt die Kriterien zur Erzeugung der Mischrasterkarte
- Relativ leichte Umsetzbarkeit
- Allerdings ist eine Überaggregation in weniger dicht besiedelten Regionen zu erkennen
- Zwingende Aggregation von 4 Rasterzellen
- Einschränkungen bei der Umrissgeometrie der Ausgangsrasterkarte damit der Ansatz durchführbar ist

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

Aus der Bewertung des Quadtree-basierten Ansatzes ergeben sich 2 Optimierungsschlussfolgerungen:

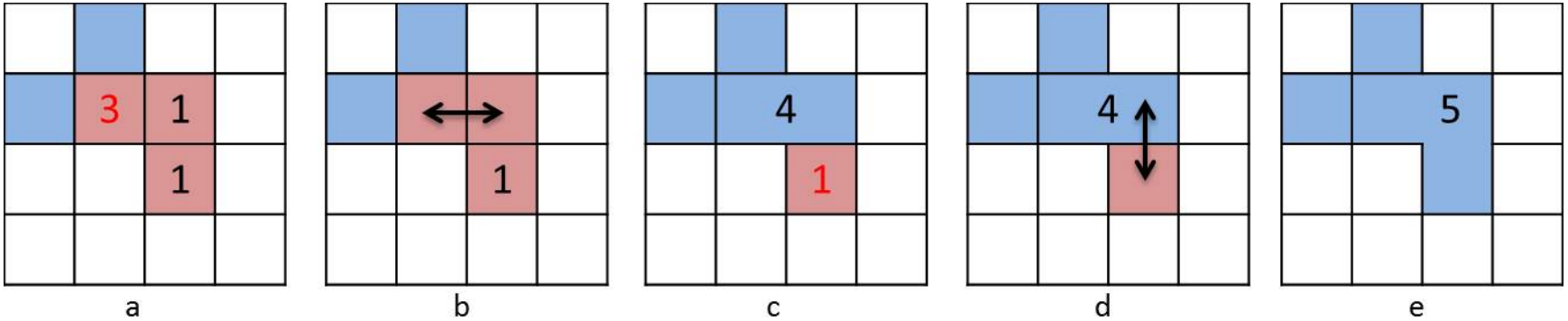
- Aggregationen in der direkten Nachbarschaft verbessern mit dem Ziel, nicht zwingend 4 Rasterzellen zum erreichen des Grenzwertes zu benötigen (4er- und 8er Nachbarschaft)

- Aggregation in weniger dicht besiedelten Gebieten optimieren (mittels Graphen)

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

Aggregationsschritte in der direkten Nachbarschaft:



a: Unterbesetzte Rasterzelle

b: Suchen in der 4er Nachbarschaft

c: Aggregation mit ausreichend besetzter Rasterzelle

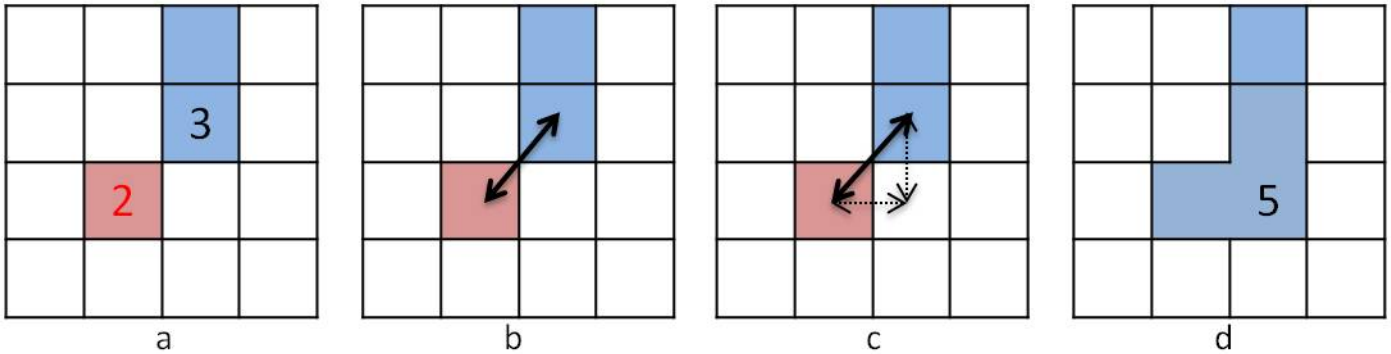
d: Das Aggregat liegt nun in der 4er Nachbarschaft einer unterbesetzten Rasterzelle

e: Einbezug der Rasterzelle in das Aggregat

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

Aggregationsschritte in der direkten Nachbarschaft:



a: Unterbesetzte Rasterzelle

b: Suchen in der 8er Nachbarschaft

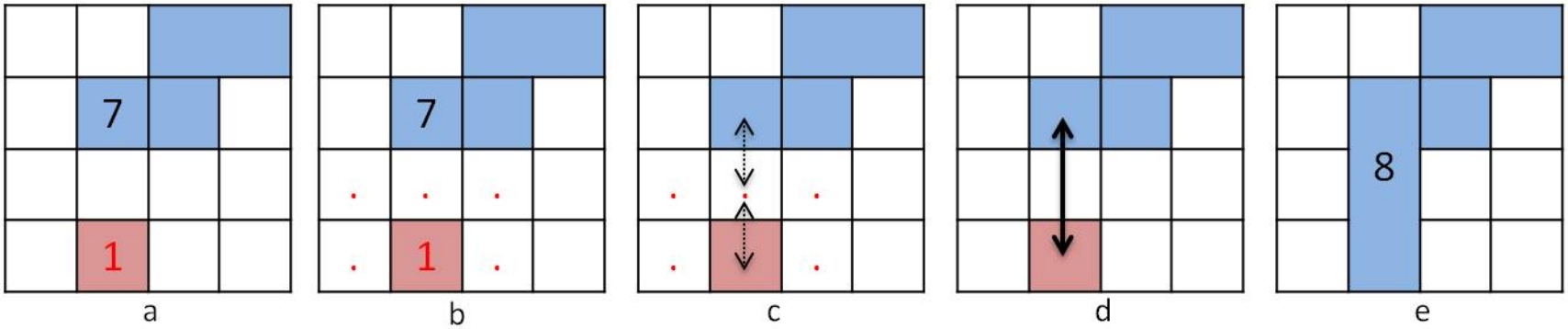
c: Zu Hilfenahme einer 3. Rasterzelle

d: Aggregation

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

Aggregationsschritte in der direkten Nachbarschaft:



a: Unterbesetzte Rasterzelle

b: Auflisten der Rasterzellen der 8er Nachbarschaft

c: Ausgehend von dieses Rasterzellen wird erneut in einer 8er Nachbarschaft gesucht

d: Das Aggregat liegt nun in der 4er Nachbarschaft einer unterbesetzten Rasterzelle

e: Einbezug der Rasterzelle in das Aggregat

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

- Rasterzellen, die jetzt noch unterbesetzt sind:

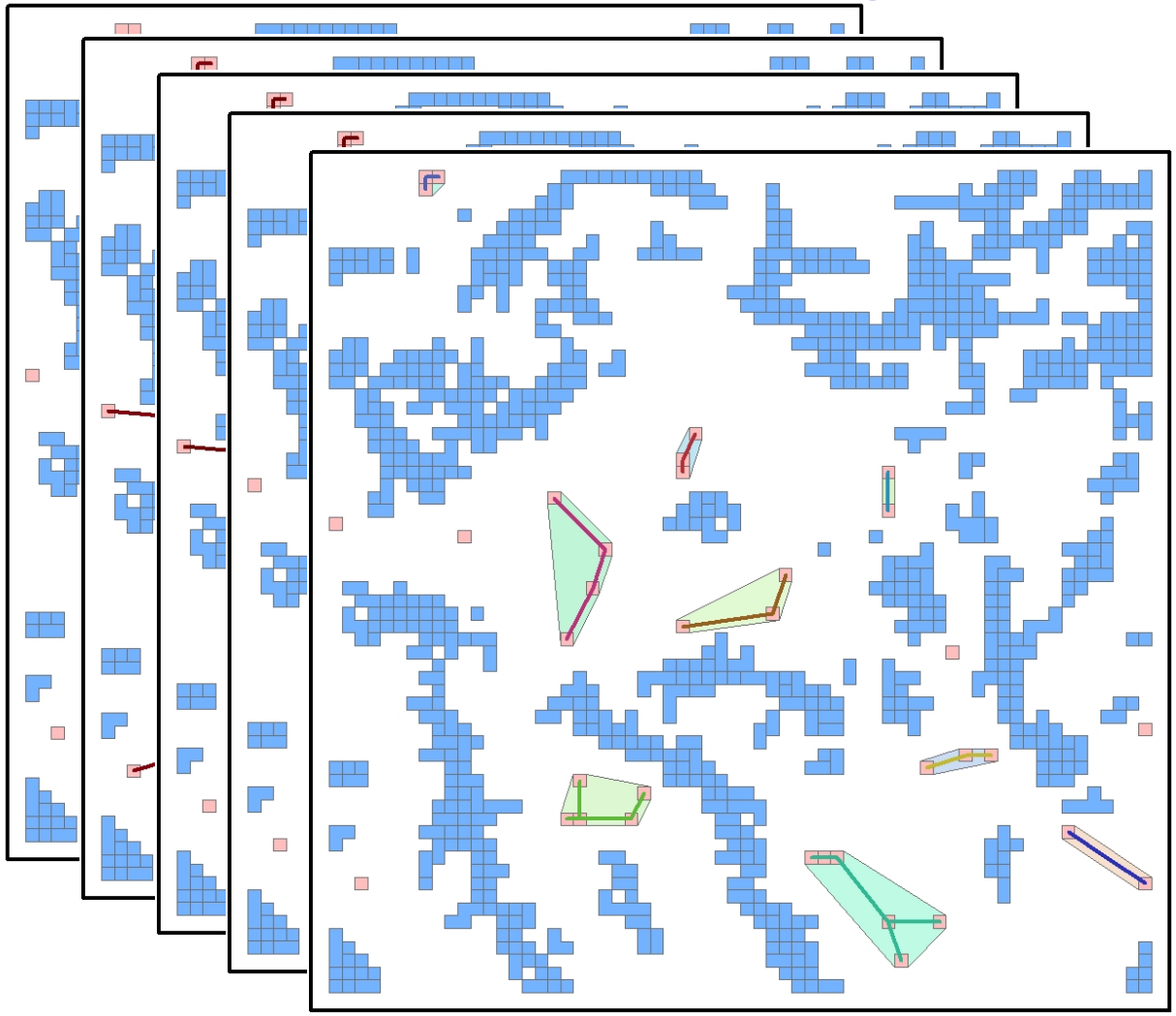
  - weisen entweder zu große Distanzen zu ausreichend besetzten Rasterzellen auf

  - oder eine Aggregation mit ihren benachbarten Rasterzellen würde den Grenzwert nicht überschreiten

  -> Anwendung von Graphen zur Clusterbildung

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz



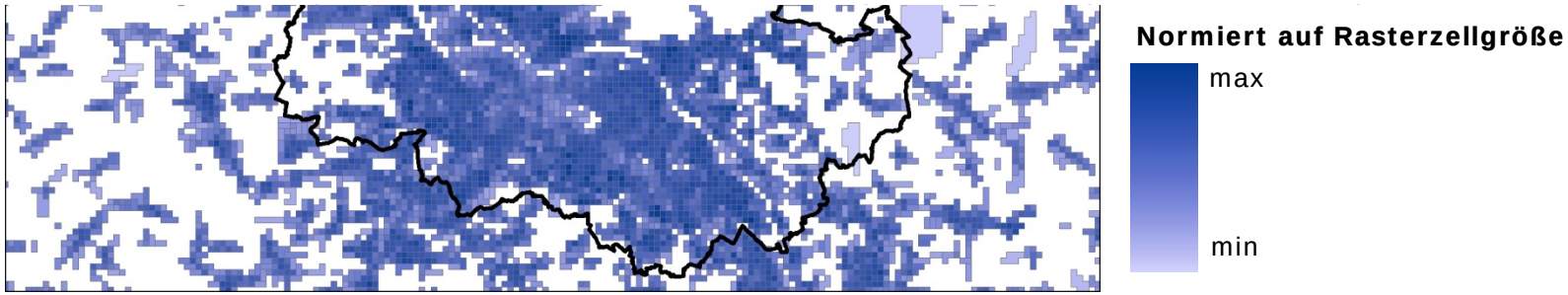
TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung

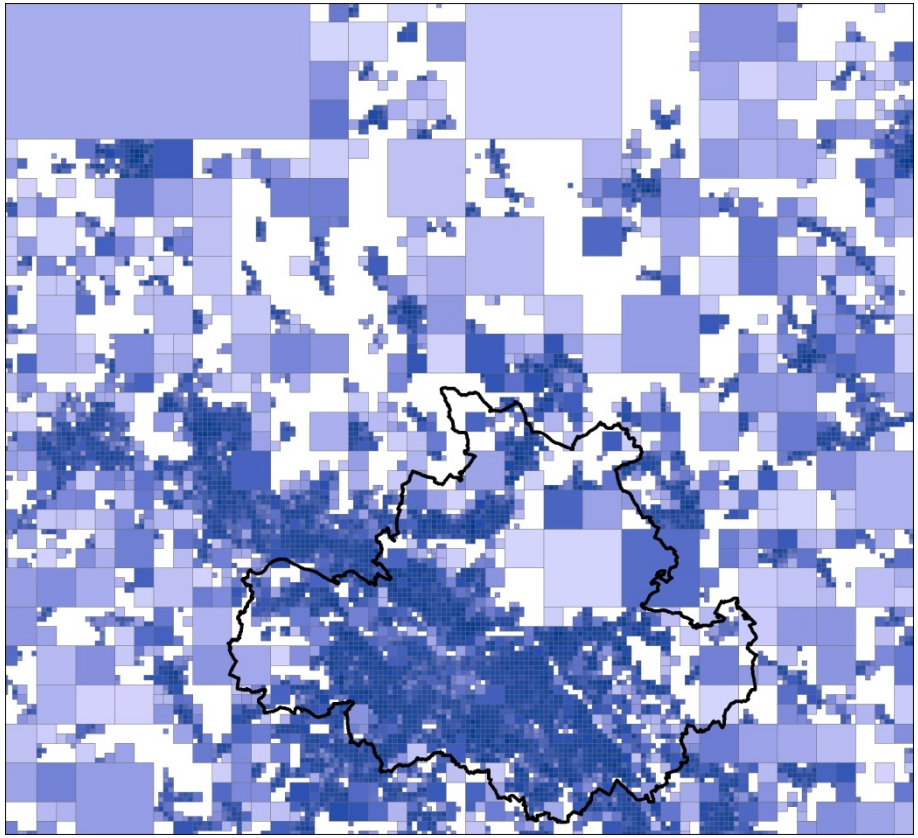
4.2 Formflexibler Aggregationsansatz

		Dresden Grenzwert ≥ 4	
		Formflexibler Ansatz	
	Ausgangskarte	Mischrasterkarte	Auflösungs-verlust
Anzahl Rasterzellen	65536	59441	-6095
Durchschnittliche Gebäudeanzahl pro Rasterzelle	7,04	7,76	0,72
Rasterzellengröße (Mittelwert)	250	260,59	10,59

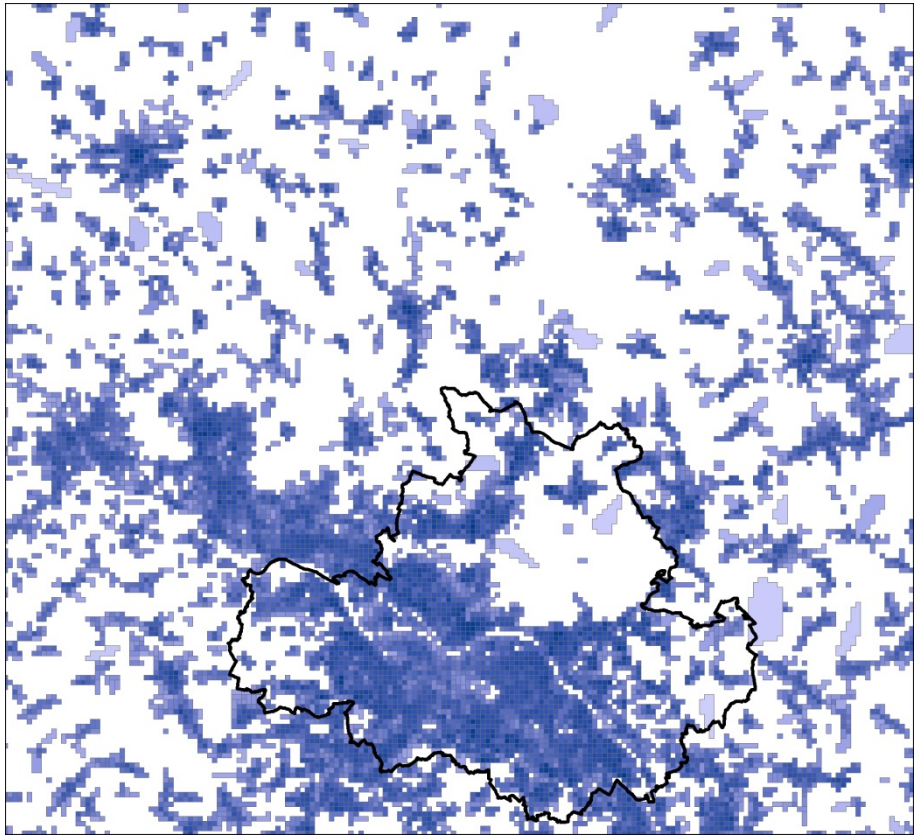


4.2 Formflexibler Aggregationsansatz

Quadtree-basierter Ansatz



Formflexibler Ansatz



TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

4.2 Formflexibler Aggregationsansatz

Zusammenfassung:

- Der Ansatz der Formflexiblen Aggregation erfüllt die Kriterien zur Erzeugung der Mischrasterkarte
- Kleinräumige Statistik besser umsetzbar
- Relativ komplex in der Umsetzung → höhere Fehleranfälligkeit
- Vorgaben auch mit anderen Parametern möglich
- Verlust der quadratischen Form der Aggregate

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

5. Fazit

- Rasterkarten bieten eine gute Grundlage, um den Datenschutz mittels Aggregation zu erreichen
- Die vorgestellten Ansätze eignen sich für dicht und geringer besiedelte Gebiete und sind für verschiedene Grenzwerte anwendbar
- Weiterhin können sie für andere Zielstellungen genutzt werden
- Auf Grund der unterschiedlichen Zellgrößen und- formen können jedoch Interpretationsschwierigkeiten auftreten

TECHNISCHE UNIVERSITÄT DRESDEN
DRESDEN concept
Leibniz-Institut für ökologische Raumentwicklung

Vielen Dank für die Aufmerksamkeit

TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Leibniz-Institut für ökologische Raumentwicklung