

# Automatisierte Auswahl von Methoden zur Überlagerung von Geodaten mit unterschiedlichen Bezugsgeometrien

MAIKE SCHOTTEN<sup>1</sup> & JOCHEN SCHIEWE<sup>1</sup>

*Zusammenfassung: Durch die Verknüpfung verschiedener Datensätze können oft vielfältige neue Informationen gewonnen werden. Zu dieser Überlagerung gehört nicht nur die geometrische Verschneidung, sondern auch die Neuverteilung der Attributwerte.*

*Da es trotz einer Reihe existierender Überlagerungsalgorithmen an einer kompakten und homogenen Bereitstellung der unterschiedlichen Methoden fehlt, werden in diesem Paper die Erstellung eines aufgabenorientierten und datenbasierten Regelapparates zur Methodenauswahl sowie die Dokumentation und Darstellung der erzielten Überlagerungs-Genauigkeiten beschrieben.*

## 1 Problemstellung

Oft können erst durch die Verknüpfung von unterschiedlichen Datensätzen interessante und räumlich differenzierte Zusammenhänge erkannt und daraus weiterführende Ergebnisse gewonnen werden. Bei solchen Überlagerungen tritt allerdings häufig das Problem auf, dass die zugrunde liegenden Daten für unterschiedliche Bezugsgeometrien (z. B.: Landkreisebene vs. PLZ-Bereiche) vorliegen und somit die Zuordnung der jeweiligen Attributwerte nicht ohne Weiteres möglich ist.

Es existiert bereits eine Reihe von Überlagerungs-Algorithmen (vgl. FLOWERDEW & GREEN 1994; LI et al. 2007; TOMIO et al. 2016), allerdings fehlt es den Nutzern an einer leicht handhabbaren Software, die die notwendigen Methoden kompakt zusammenstellt, sowie an wissenschaftlich fundierten Empfehlungen zur fallspezifischen Methoden-Auswahl. Ferner findet in der Regel auch keine Dokumentation der resultierenden Interpolationsgenauigkeiten statt, sodass Bewertungen des Überlagerungs-Ergebnisses nicht möglich sind.

## 2 Methodik

Die beiden Basisoperationen einer Überlagerung sind einerseits die Zusammenführung kleinerer zu größeren Raumeinheiten (Aggregation, Aufsummierung) und andererseits der umgekehrte Prozess, die Zerlegung (Disaggregation) in kleinere Einheiten. Bei den meisten Überlagerungen findet eine Kombination der Operationen statt.

In den gängigen Softwareprodukten ist für diesen Zweck eine einfache Flächengewichtung implementiert. Über die Einführung von Zusatzwissen können die Ergebnisse jedoch an die realen Bedingungen (z. B. Bevölkerungsverteilung) angepasst und damit verbessert werden.

---

<sup>1</sup> HafenCity Universität Hamburg, Labor für Geoinformatik und Geovisualisierung, Überseeallee 16, D-20457 Hamburg, E-Mail: [Maike.Schotten, Jochen.Schiewe]@hcu-hamburg.de

Die verschiedenen Methoden setzen unterschiedliche Ausprägungen des Zusatzwissens voraus. Daneben gibt es Zusammenhänge zwischen der Interpolationsgüte und (geo-)statistischen Parametern der Eingangsdatensätze. Je stärker der Grad der Disaggregation ist, desto größer ist der zu erwartende mittlere Fehler der Interpolationsergebnisse, da im Gegensatz zur reinen Aggregation eine erhöhte Notwendigkeit zur Interpolation zu erwarten ist. Hierzu wird die Homogenität zwischen Quell- und Zielzonen sowohl über die Anzahl der Polygone wie auch über die Standardabweichung der Flächengrößen der Polygone betrachtet.

Für die Bestimmung der internen Genauigkeit werden die Ergebnisse aller Interpolationsmethoden aus mehreren Testgebieten verglichen und auf Abweichungen untersucht. Durch die Analyse zusätzlicher synthetischer Datensätze ist es zudem möglich, systematische Auffälligkeiten der auftretenden Differenzen zwischen den Methoden zu entdecken.

Im Rahmen der zunehmenden OpenData-Ansätze in Bund und Kommunen liegen für viele Testgebiete (z. B. Hamburg oder Berlin) ausgewählte Statistiken auf mehreren Bezugsebenen vor. Nach Anwendung aller Methoden kann die externe Genauigkeit durch direkten Vergleich mit der ‘ground truth‘ bestimmt werden.

Im Zuge einer besseren Nutzbarkeit werden die erhobenen (geo-)statistischen Kennwerte dokumentiert und daraus (möglichst parametrisierte) Empfehlungen abgeleitet.

### **3 Ergebnisse und Ausblick**

Dieser Werkstattbericht zeigt, dass die daten- und aufgabenspezifische Methodenempfehlung sowie die Dokumentation und Darstellung der Interpolationsungenauigkeiten dem Nutzer eine transparente Entscheidungsgrundlage zur Bewertung und Weiterverwendung der Ergebnisse bieten. Dadurch ist es möglich, den wachsenden Bestand an Geodaten explorativ zu erforschen um neue Erkenntnisse zu gewinnen.

### **4 Literaturverzeichnis**

- FLOWERDEW, R. & GREEN, M., 1994: Areal interpolation and types of data. S. Fotheringham & P.Rogerson (eds.), *Spatial Analysis and GIS*, London: Taylor & Francis, 73-89.
- LI, T., PULLAR, D., CORCORAN, J. & STIMSON, R., 2007: A comparison of spatial disaggregation techniques as applied to population estimation for South East Queensland (SEQ), Australia. *Applied GIS*, **3**(9).
- TOMIO, B., SCHIEWE, J., & TRÖGER, J., 2016: Die Kombination von geographischen Daten verschiedener Bezugsräume mittels Flächeninterpolation – Methoden für Datenjournalisten und Datenjournalistinnen. *Kartographische Nachrichten*, **66**(3), 123-128.