
Flexible Aggregation regionalstatistischer Erhebungen – Neue Produkte der Statistik Austria

Thomas PRINZ*, Josef STROBL*, Erich WONKA**

* RSA – iSPACE

** Statistik Austria

Zusammenfassung

Regionalstatistische Informationen aus Demographie, Wirtschafts- und Sozialstatistik werden als Grundlage für planerische Arbeiten, Geomarketing, Standorts-Beurteilung, Mobilfunk - Reichweitenschätzung und Infrastruktur-Marktanalysen benötigt. Adresspunktbezogene Daten können u.a. aus datenschutzrechtlichen Gründen häufig nicht eingesetzt werden, sodass Aggregate auf bauliche (Blöcke), administrative und geometrische Bezugsflächen (Raster) in der Praxis als Alternativen zur Verfügung stehen. Neben der traditionellen Summierung auf topographische und administrative Zonen gewinnt eine flexible und anwendungsorientierte Raster-basierte Aufbereitung an Bedeutung.

Ausgehend von einem nunmehr auch in der Statistik wirksamen Trend zur adressbezogenen Georeferenzierung entwickeln die Autoren eine Methodologie zur anwendungs- und problembezogenen Aggregation. Damit können adressbezogene statistische Masszahlen flexibel bereitgestellt und in z.B. raumplanerische Abläufe integriert werden. Die vorliegende Arbeit entstand aus einer Kooperation des Forschungsstudios iSPACE mit der Statistik Austria, anhand eines Fallbeispiels das mit Unterstützung aus dem Amt der Salzburger Landesregierung umgesetzt wurde.

1 Einleitung und Anforderungen

Die Statistik Austria stellt mit Unterstützung des Bundesamtes für Eich- und Vermessungswesen die Speicherung der Großzählungsdaten (Volks- Gebäude- und Wohnungszählung sowie Arbeitsstättenzählung) 2001 auf generelle Bindung an Adresskoordinaten um. Anhand eines universellen Datenmodells besteht die Möglichkeit, unabhängig von bereits bestehenden statistischen und administrativen Gebietsgliederungen nahezu beliebige Aggregationseinheiten zu definieren (KAMINGER und WONKA 2004).

Wie detailliert aus räumlicher Sicht ein statistischer Sachverhalt dargestellt wird hängt von der jeweiligen Aufgabenstellung ab. Dem Ausgabemaßstab angepasste Repräsentation ist in erster Linie unabhängig von den individuellen topographischen Strukturen des Untersuchungsgebiets auf der Basis von variablen Rasterzellen möglich, da die ‚Planquadrate‘ den jeweiligen Bedürfnissen entsprechend vergrößert oder verkleinert werden können.

Um diesen Anforderungen nachzukommen, hat Statistik Austria sein regionalstatistisches Angebot erweitert und es werden erstmals die Daten aus der Volks- Gebäude- und Woh-

nungszählung sowie Arbeitsstättenzählung 2001 neben den bisherigen statistischen und administrativen Gebietseinheiten auch auf der Grundlage von Planquadraten im Landeskoordinatensystem mit einer Seitenlänge von 250m, 500m, 1.000m, 2.500m, 5.000m und 10.000m zur Verfügung gestellt.

Im internationalen Vergleich stellt diese Aggregation auf Rasterzellen keine Novität dar. Während z.B. in der Schweiz die Gemeinde früher die kleinste räumliche Bezugseinheit für die Auswertung von Volkszählungsdaten bildete, sind seit 1970 auch Daten mit Hektarraster - Auflösung erhältlich (GEOSTAT 1999). Nicht nur für viele Fragestellungen der Raumplanung stellen diese Datengrundlagen auf der Basis administrativ unabhängiger Raumeinheiten eine äußerst interessante Bereicherung dar, sondern insbesondere auch für Marktforschung und Infrastrukturplanung. Die Nutzung und in-Wert-Setzung dieses „neutralen“ Raumbezugs beginnt sich erst durchzusetzen, obwohl die topographisch unabhängige Betrachtung von statistischen Informationen und Raumstrukturen wertvolle, räumlich fein ausdifferenzierte Zusatzinformationen bietet.

Der Aspekt des Datenschutzes tritt generell umso stärker in Erscheinung, je kleiner die Aggregationseinheiten sind. So werden etwa für die schweizerischen ha-Raster Zellenwerte bei weniger als drei Einwohnern nicht ausgewiesen (GEOSTAT 1999) um direkte Rückschlüsse auf Einzelpersonen zu verhindern. Je nach Thematik und Verschneidungspotenzial („räumliche Rasterfahndung“) kann dieser Schwellwert natürlich variabel gesetzt werden.

Gerade die Ergebnisse von Markt- und Versorgungsbewertung, Modellierung von Erreichbarkeiten (Infrastrukturplanung, Öffentlicher Personennahverkehr, Schulen) werden umso präziser je kleinräumiger statistische Daten vorliegen. Bevölkerungsverteilung und damit die potenzielle Nachfrage infrastruktureller Leistungen, oder die Prognose der zeitlichen Entwicklung des Bedarfs an Infrastruktureinrichtungen sind Basis der Planung von Kindergärten und Schulen ebenso wie sozialer Dienste, des öffentlichen Verkehrs und anderer Ver- und Entsorgungsdienste.

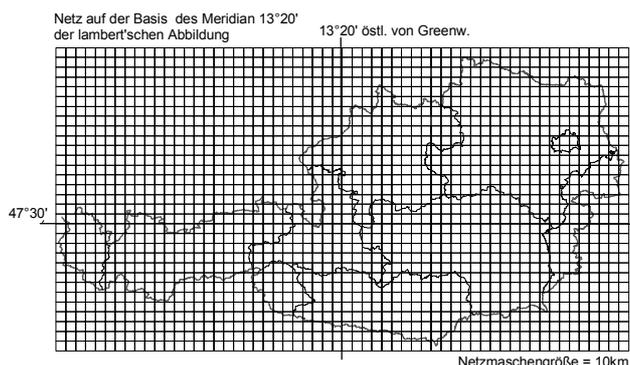
Raumplanung und Marktanalyse bedienen sich heute vielfach der methodischen Grundlagen komplexer Indikatoren. Diese stehen für nicht direkt messbare Faktoren und werden meist aus allgemein verfügbaren Maßzahlen zusammengesetzt, um regionale Vergleiche wie auch zeitliche Längsschnitte zu ermöglichen. Indikatorenentwicklung findet fach- und anwendungsbezogen statt, soll also möglichst wenig seitens der Datenlage eingeschränkt werden. Um unterschiedlichste Indikatoren zu unterstützen, ist eine möglichst flexible Aggregation ohne fixe Vorgabe z.B. von Rasterzellen-Dimensionen erforderlich – die Einhaltung von Datenschutzbestimmungen muss durch entsprechende Algorithmen jedenfalls gewährleistet sein.

2 Statistische Planquadrate der Statistik Austria

Gebietsgliederungen in Planquadrate sind zwangsläufig von gewählten räumlichen Bezugssystemen bzw. Projektionen abhängig. Für die zeitliche Vergleichbarkeit von Daten benö-

tigt man ein fix definiertes und damit von den Koordinaten des jeweiligen Bezugsstems abhängiges Netz. In erster Linie werden dazu die Koordinatensysteme der amtlichen Karten verwendet. In den Österreichischen Karten im Maßstab 1:25.000 (ÖK25V), 1:50.000 (ÖK50) und 1:200.000 (ÖK200) war dies das Gauß-Krüger-Netz bzw. Bundesmeldenetz, wobei zur Zeit ein Umstieg auf das UTM-Netz erfolgt. Für die Übersichtskarte von Österreich 1:500.000 (ÖK500) bietet sich ein Netz auf Basis des Lambert'schen Kegelentwurfs an. Genaue Informationen über die Kartenwerke und Netze können der Homepage des Bundesamts für Eich- und Vermessungswesen (BEV) - www.bev.gv.at unter „Geobasisdaten“ entnommen werden.

Zur nahtlosen und überlappungsfreien Darstellung des Bundesgebiets erfolgt die Abbildung durch winkeltreue Lambert'sche Schnittkegelprojektion mit Parallelkreisen in 46° und 49° nördlicher Breite. Als Ursprung des Gitternetzes dient der Schnittpunkt des Meridians $13^\circ 20'$ mit dem Breitenkreis $47^\circ 30'$.



Alternative Planquadrate sind durch das Netz der Gauß-Krüger Projektion oder der UTM-Projektion definiert. Die resultierenden Rasterzellen sind mit einer einheitlichen Nomenklatur zur eindeutigen Identifikation der Zuordnung von Adressen zu Zellen versehen (KAMINGER und WONKA 2004). Für diese koordinatengebundenen Planquadrate in der jeweils gewählten Auflösung erhält der Datenbankbenutzer die folgenden statistischen Fallzahlen: Zahl der Hauptwohnsitzer, Zahl der Nebenwohnsitzer, Zahl der Haushalte, Zahl der Wohnungen, Zahl der Gebäude und Zahl der Arbeitsstätten. Merkmale zu diesen Fallzahlen dürfen aus Datenschutzgründen an externe Datenbankbenutzer nur in aggregierter Form weitergegeben werden. Wenn es sich z.B. um Merkmale aus der Volkszählung handelt, müssen in einem Planquadrat mindestens 31 Hauptwohnsitzer sein, bei Merkmalen aus der Gebäude- und Wohnungszählung mindestens 4 Gebäude.

Da sich bei Rasterzellen die Daten auf gleiche Flächen beziehen, stellen die Absolutwerte zugleich auch die Dichtewerte dar – ein klarer Vorteil bei der kartographischen Visualisierung! Wie aus dem Vergleich der Abbildungen 1 und 2 hervorgeht stellt die implizierte Betrachtung von statistischen Informationen und z.B. Siedlungsstrukturen einen signifikanten planerischen Mehrwert dar.

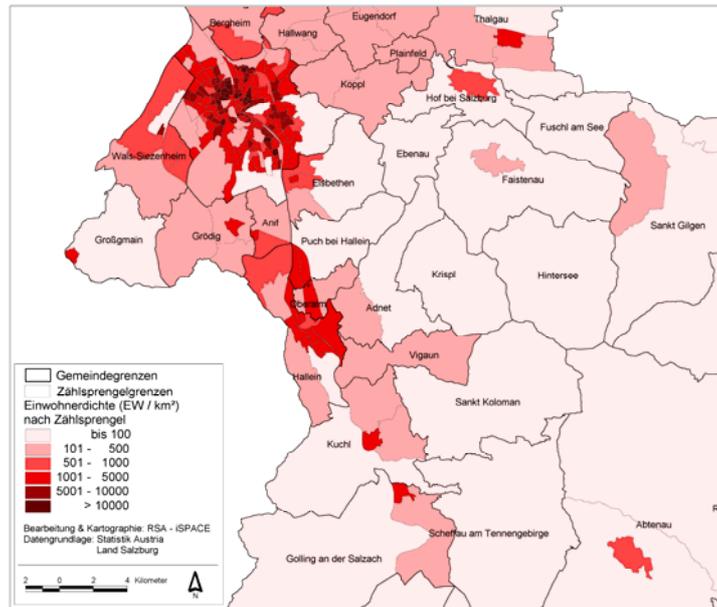


Abb. 1: Einwohnerdichte bezogen auf Zählsprengele

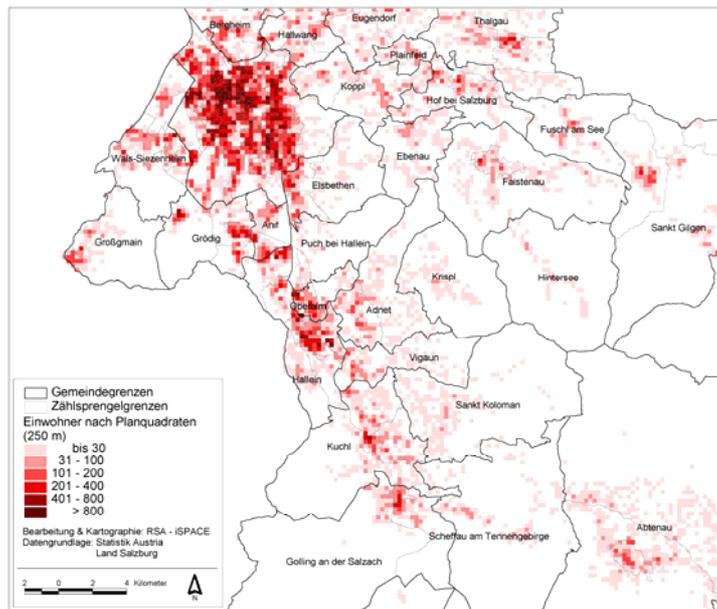


Abb. 2: Einwohnerdichte bzw. -verteilung bezogen auf Raster (250m)

Für bestimmte regionalstatistische Aufgabestellungen vor allem in dicht verbauten Gebieten (z.B. für die Infrastrukturplanung) ist auch die Planquadratgröße von 250m vielfach noch zu grob. Aus diesem Grund hat sich Statistik Austria entschlossen, in Verdichtungsgebieten die Fallzahlen (z.B. nur Hauptwohnsitze oder Gebäude) auf der Basis von Planquadraten auch mit einer Seitenlänge von 125m abzugeben.

Dadurch hat der externe Datenbankbenutzer die Möglichkeit aggregierte Merkmale von größeren Planquadratinheiten (z.B. 250m oder 500m) entsprechend der Größe der Fallzahl nach einem von ihm festgelegten Algorithmus proportional aufzuteilen. Diese Vorgangsweise hat den Vorteil, dass räumliche Detailuntersuchungen gemacht werden können, ohne dass Rückschlüsse auf individuelle Merkmale von Gebäuden, Wohnungen oder Personen möglich sind.

3 Flexible Aggregation regionalstatistischer Erhebungen

Wie oben bereits erwähnt, steht die Aussagekraft eines jeglichen räumlichen Indikators in Abhängigkeit von der Wahl eines geeigneten räumlichen Bezugssystems. Dieses ist daher anwendungsbezogen zu wählen und wird somit sinnvollerweise nicht seitens der Datenquelle festgelegt – eindeutig ein Änderungs-Wunsch an die im vorangehenden Abschnitt charakterisierte Vorgangsweise im Rahmen der Statistik Austria.

Von essentieller Bedeutung ist für flexible Planungsanwendung die Möglichkeit der Definition anwendungsspezifischer Aggregationseinheiten, ebenso wie die freie Festlegung von räumlichen Bezugssystemen. Dies wiederum bedeutet dass statistische Aggregationseinheiten oberhalb der Adress-Ebene nach Möglichkeit nicht fix vordefiniert sein sollen, sondern dass eine Aggregation auf beliebige administrative oder geometrisch-regelmäßige Einheiten möglich bleiben muss. Um diesen Anforderungen zu entsprechen, muss GIS-Funktionalität zur räumlichen Aggregation letztlich im „Statistik-Server“ integriert sein, sodass (Adress-) Punkt-bezogene Angaben unter Wahrung von einschränkenden Regeln (Datenschutz!) im Zuge von Abfragen ad hoc räumlich aggregiert werden können.

Die Planquadratgröße muss den anwendungsbezogenen Bedürfnissen entsprechend gewählt werden können. Bei kleinen Rasterzellen sind statistische Daten leichter mit detaillierten raumbezogenen Daten aus Topographie, Planung, Naturraum und sonstigen Thematiken zu verknüpfen. Gerade die Erstellung von Grundlagen für multi-thematische Stadt- und Regionalplanungs erfordert die Zusammenführung von Informationen aus unterschiedlichen Quellen.

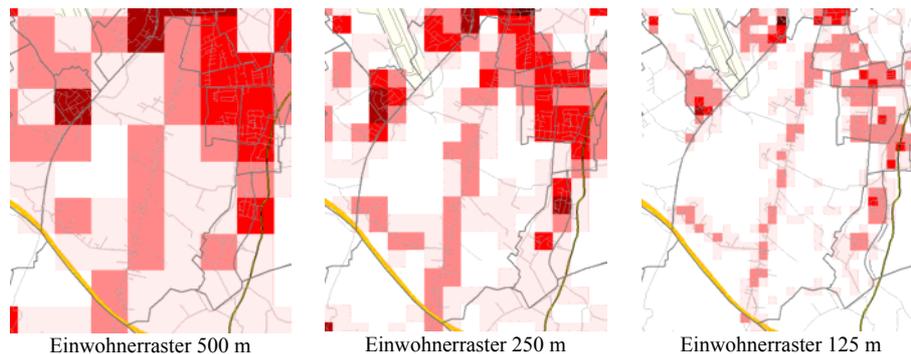


Abb. 3: Variable Raster-Aggregation regionalstatistischer Originaldaten

Mit einer jeweils kleinräumige angepassten Planquadrat-Aggregation im Bereich dichter verbauter städtischer Gebiete werden neue Voraussetzungen geschaffen, um eine sozial-räumliche Stadt- und Regionsgliederung unabhängig von den seitens der Statistik Austria veröffentlichten Bezugseinheiten durchzuführen.

4 Ausblick

Als Zukunftsperspektive werden Erfordernis und Ansätze zur ‚on the fly‘ Erzeugung flexibel definierter polygonaler und zellbasierter Bezugseinheiten auf der Basis von map und feature services und zur Präsentation von „beliebig“ aggregierten adressbezogenen Statistikdaten unter gleichzeitiger Kontrolle datenschutzrechtlicher Bestimmungen diskutiert. Letztere werden u.a. durch hierarchisch höhere Aggregationsniveaus bei Zellfrequenzen unterhalb kritischer Werte realisiert, woraus sich eine inhaltsgesteuerte variable Auflösung von Rastern ergibt (dazu siehe STROBL 2004).

5 Literatur

- GEOSTAT (1999), *Die Servicestelle des Bundes für raumbezogene Daten.*- Online: <http://www.statistik.admin.ch/service-stat/geostat/pdf/geostat-99d.pdf>
- KAMINGER I. und E. WONKA (2004): Von einer Österreichgliederung nach Gemeinden zu Planquadraten: Statistik Austria erweitert sein regionalstatistisches Angebot. In: Schrenk, M. (Hrsg.): *Computergestützte Raumplanung – Beiträge zum Symposium CORP 2004.*- Wien.
- PRINZ, T. (2003): *GIS-gestützte Bewertungsverfahren in einer zukunftsorientierten Stadt- und Regionalplanung*, In: Strobl, J., T. Blaschke und G. Griesebner (Hrsg.): *Angewandte Geographische Informationsverarbeitung XV, Beiträge zum AGIT-Symposium Salzburg*, S. 424 – 429
- STROBL, J. (2004, in Druck): *Hierarchische Aggregation: Detailinformation versus Datenschutz am Beispiel adressbezogen georeferenzierter Datensätze.*