

## **Räumliche Disaggregation von sozio-ökonomischen Daten**

Klaus STEINNOCHER, Frederic PETRINI, Tanja TÖTZER, Jürgen WEICHSELBAUM

### **Zusammenfassung**

Im Rahmen des Observatory Spatial Planning des Integrated Project geoland werden Produkte und Services für die räumliche Planung auf europäischer, nationaler und regionaler Ebene entwickelt. Eine der Produktgruppen umfasst räumliche Indikatoren, die sich aus der Kombination von sozio-ökonomischen und real-räumlichen Informationen im GIS ableiten lassen. Zur Verknüpfung dieser Daten wird die Methode der räumlichen Disaggregation eingesetzt. Diese Methode basiert auf der Annahme, dass statistische Informationen, die sich auf administrative Einheiten beziehen, auf Grundlage detaillierter Daten räumlich verfeinert werden können. Die vorliegende Arbeit analysiert die dabei getroffenen Annahmen und die Qualität der räumlichen Disaggregation von Bevölkerungsdaten anhand von Bebauungsdichten. Die Ergebnisse der Untersuchung bestätigen die methodische Vorgangsweise und erlauben eine Abschätzung der zu erwartenden Genauigkeiten.

### **1 Einleitung**

Das Observatory Spatial Planning<sup>1</sup> (OSP) ist eines von sechs Observatories im Integrated Project Geoland, das im 6. Rahmenprogramm der EU gefördert wird. Ziel der Forschungsarbeiten des OSP ist die Entwicklung von Produkten und Services für die räumliche Planung auf europäischer, nationaler und regionaler Ebene. Die eingesetzten Datengrundlagen stammen zum einen aus der Fernerkundung, zum anderen aus räumlichen Datenbanken und statistischen Quellen. Die Arbeiten orientieren sich dabei an den Anforderungen relevanter Politiken und Richtlinien.

Eine grundlegende Gruppe von Produkten stellen räumliche Indikatoren dar, die sich aus der Kombination von sozio-ökonomischen und real-räumlichen Informationen im GIS ableiten lassen. Zu diesen Indikatoren zählen u.a. Flächenverbrauch pro Kopf, Verfügbarkeit von erreichbaren Erholungsflächen pro Einwohner, Flächenproduktivität oder Intensität landwirtschaftlicher Nutzung (STEINNOCHER et al. 2005).

Zur Verknüpfung sozio-ökonomischer mit real-räumlichen Informationen wird die Methode der räumlichen Disaggregation eingesetzt. Diese Methode basiert auf der Annahme, dass statistische Informationen, die sich auf administrative Einheiten wie z.B. NUTS Regionen beziehen, auf Grundlage detaillierter Daten z.B. aus der Fernerkundung räumlich verfeinert

---

<sup>1</sup> OSP wird koordiniert von GeoVille GmbH, die wissenschaftliche Leitung liegt bei ARC systems research GmbH

werden können (CHEN 2002, STEINNOCHER & KÖSTL 2002). Einer räumlichen Disaggregation liegen jedoch gewisse Annahmen zugrunde, die einer näheren Untersuchung bedürfen.

## 2 Räumliche Disaggregation

Räumliche Disaggregation basiert auf der Annahme, dass Daten, die global für eine Region zur Verfügung stehen, mittels räumlich differenzierter Parameter innerhalb der Region verteilt werden können. Die räumliche Aufteilung erfolgt dabei in der Regel über eine gewichtete Summenfunktion. Voraussetzung für diese Vorgangsweise ist allerdings, dass es eine eindeutige Abhängigkeit zwischen dem globalen und dem räumlich differenzierten Parameter gibt. Weiters ist zu klären, welche Faktoren diese Abhängigkeit beschreiben und ob es allgemein gültige Regeln gibt.

Für die Untersuchungen im Rahmen dieses Beitrages wurden die Parameter Bevölkerungs- und Bebauungsdichte herangezogen, wobei folgende Annahmen getroffen wurden: die Bevölkerungsdichte ( $BevDi$ ) ist proportional zur Bebauungsdichte ( $BebDi$ ), wobei unbebaute Flächen keine Bevölkerung aufweisen. D.h. die Bevölkerungsdichte ist eine lineare Funktion der Bebauungsdichte, wobei diese Funktion durch den Ursprung geht:

$$BevDi = k * BevDi \quad (1)$$

Weiters wird angenommen, dass der Faktor  $k$  innerhalb einer Region konstant ist, zwischen unterschiedlichen Regionen jedoch variieren kann. Um eine räumliche Disaggregation durchführen zu können, muss daher der Faktor  $k$  für die jeweilige Region bestimmt werden. Als bekannte Größen werden die Gesamtbevölkerung der Region ( $Bev$ ) sowie die Flächenanteile der einzelnen Bebauungsklassen vorausgesetzt. Daraus ergibt sich folgende Gleichung:

$$Bev = \sum_i F_i * BevDi_i \quad (2)$$

mit

$F_i$  Flächenanteil der Bebauungsklasse  $i$   
 $BevDi_i$  Bevölkerungsdichte der Bebauungsklasse  $i$   
 $BebDi_i$  Bebauungsdichte der Bebauungsklasse  $i$

Die Gesamtbevölkerung ergibt sich also aus der Summe der Flächenanteile multipliziert mit den korrespondierenden Bevölkerungsdichten. Letztere sind allerdings nicht bekannt, sondern werden über die Bebauungsdichten geschätzt (1):

$$Bev = \sum_i F_i * k * BevDi_i \quad (3)$$

Daraus lässt sich der für die untersuchte Region spezifische Faktor  $k$  ableiten:

$$k = Bev / \sum_i F_i * Bebd_i \quad (4)$$

Die Zuordnung der Bevölkerungsdichten zu den einzelnen Bebauungsklassen erfolgt dann über Gleichung (1).

### 3 Datengrundlagen und Testgebiete

Als Datengrundlage für die Berechnung und die Genauigkeitsanalyse der räumlichen Disaggregation wurden zunächst Daten der Statistik Austria herangezogen. Neben der traditionellen Repräsentation in statistischen Einheiten stehen die Bevölkerungsdaten der Volkszählung 2001 auch in Form eines regelmäßigen Rasters zur Verfügung. Dabei handelt es sich um ein neues Produkt der Statistik Austria, bei dem sozio-ökonomische Daten mittels Adresskoordinaten verortet und – aus Gründen des Datenschutzes – auf Planquadrate mit einer Seitenlänge von 250m aggregiert wurden (KAMINGER & WONKA 2004). Diese Daten umfassen das gesamte Bundesgebiet von Österreich.

Die Einwohnerzuteilung zu den Planquadraten erfolgt über das zentrale Melderegister, in dem jeder Adresse exakte Koordinaten zugeordnet sind. Für eine Plausibilitätskontrolle wurde ein Vergleich der Einwohner pro Bezirk laut Volkszählung (VZ) 2001 und der pro Bezirk summierten Planquadratanzahlen durchgeführt. Es zeigte sich, dass die Einwohnersumme aus den Planquadraten durchgehend unter den Einwohnern aus der VZ 2001 liegt. Dies ist laut Auskunft der Statistik Austria darin zu begründen, dass die Koordinaten nicht ganz vollständig sind und ca. 2-3% der Koordinaten fehlen. Das führt dazu, dass die Einwohner nicht vollständig zugeordnet werden können und sich daher eine Diskrepanz zu den Einwohnerzahlen aus der Volkszählung 2001 ergibt. Um Konsistenz zu wahren, basieren die im Folgenden beschriebenen Untersuchungen auf den Gemeinde- und Bezirkssummen der Planquadratanzahlen.

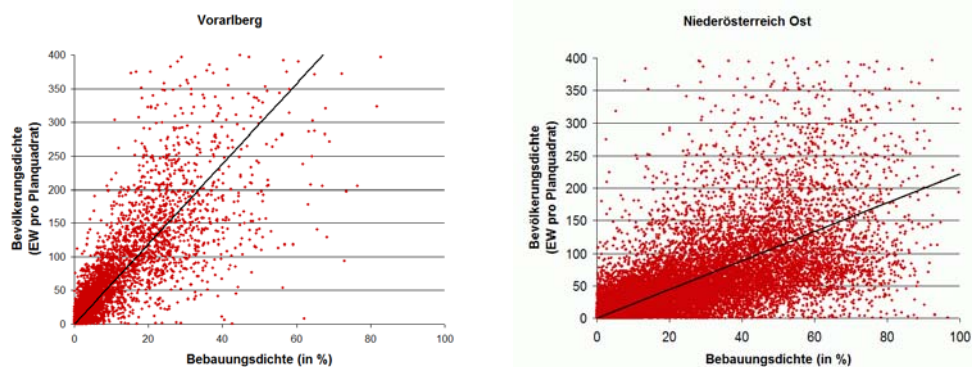
Die Ableitung der Bebauungsdichte erfolgte auf Grundlage von hochauflösenden Satellitenbilddaten. Da die Höhe von Gebäuden dabei nicht berücksichtigt werden kann, wurde die Bebauungsdichte über den Versiegelungsgrad geschätzt. Weiters wurde davon ausgegangen, dass sich in reinen Gewerbe-, Industrie- und Verkehrsflächen keine Wohnbevölkerung befindet, daher wurden diese Gebiete ausmaskiert. Als zusätzlicher Vergleichsdatensatz wurde CORINE land cover (CLC) 2000 von Österreich herangezogen.

Als Testregionen wurden aufgrund der Datenverfügbarkeit Vorarlberg und das östliche Niederösterreich ausgewählt. Vorarlberg ist auf Grund seiner stark variierenden Topographie und der daraus resultierenden unterschiedlichen Siedlungsstruktur eine Herausforderung für die räumliche Disaggregation. Aber auch das östliche Niederösterreich weist signifikante Variationen in den Bebauungsstrukturen auf. Da in der Großstadt Wien deutlich andere Strukturen vorzufinden sind als in ihrem Umland, wurde die Landesfläche von Wien ausmaskiert.

Die Auswertung der Bebauungsdichte in Vorarlberg basiert auf SPOT 5 Daten und weist einen Kartierungsmaßstab von 1:25.000 auf. Drei Klassen wurden ausgewiesen: „lockere Bebauung“ mit 0-50%, „mittlere Bebauung“ mit 50-80% und „dichte Bebauung“ mit 80-100% Bebauungsdichte. In Niederösterreich wurde auf Landsat Daten zurückgegriffen, die einen Kartierungsmaßstab von 1:100.000 erlauben. Die Bebauungsdichte wurde hier in 10% Schritten geschätzt, für die weitere Analyse dann jedoch auf die in Vorarlberg verwendeten drei Dichteklassen aggregiert. Von den CLC Daten wurde in beiden Testregionen die Klasse 1.1 – residential areas – verwendet.

## 4 Analyse

Die Untersuchung konzentrierte sich zunächst auf einen direkten Vergleich der Bebauungsdichten mit der räumlichen Verteilung der Bevölkerung laut den Planquadraten der Statistik Austria. Zu diesem Zweck wurden die Bebauungsdichten mit den Planquadraten verschnitten und eine mittlere Bebauungsdichte pro Planquadrat ermittelt. Abbildung 1 zeigt die Scatterplots der beiden Testregionen. Die Annahme, dass eine Abhängigkeit zwischen Bebauungs- und Bevölkerungsdichte besteht, wird durch die Grafiken bestätigt, – auch wenn vor allem in Niederösterreich die Daten stark streuen. Weiters ist klar ersichtlich, dass der Grad der Abhängigkeit regional verschieden ist, wie durch die unterschiedlich geneigten Trendlinien gezeigt wird. Dieser Effekt ist aber auch auf die unterschiedlichen Datenquellen bei der Erhebung der Bebauungsdichten zurückzuführen.



**Abb. 1:** Scatterplots und Trendlinien der Bebauungs- und Bevölkerungsdichte für die beiden Testregionen

Die Bestätigung der eingangs getroffenen Annahmen durch die empirische Untersuchung erlaubt nun die Anwendung der in Kapitel 2 beschriebenen räumlichen Disaggregation. Als Ausgangswerte wurden für beide Testregionen die Gesamtbevölkerung, sowie die Flächenanteile der drei Bebauungsklassen herangezogen. Gemäß Gleichung (4) konnte daraus der regionsspezifische Parameter  $k$  geschätzt werden. Die geschätzte Bevölkerungsdichte ergibt sich dann aus Einsetzen des Parameters  $k$  in Gleichung (1). Tabelle 1 zeigt die relevan-

ten Größen für beide Testgebiete. Die unterschiedliche Ausprägung des Faktors  $k$  und die daraus resultierenden Unterschiede in der Bevölkerungsdichte sind sowohl auf regionale Unterschiede als auch auf die unterschiedlichen Datenquellen zurückzuführen.

**Tab. 1:** Schätzung von Bevölkerungsdichten aus Bebauungsdichten

	EW gesamt	Bebauung [ha]			k	Bevölkerungsdichte [EW/ha]		
		locker	mittel	dicht		locker	mittel	dicht
NÖ Ost	974.106	42.005	20.482	2.217	0,377	9,44	24,53	33,97
Vbg	350.316	7.828	1.231	412	1,119	27,99	72,77	100,76

## 5 Analyse

Die geschätzten Bevölkerungsdichten wurden den korrespondierenden Bebauungsklassen zugewiesen und daraus eine räumliche Bevölkerungsverteilung berechnet. Für die numerische Analyse wurden die disaggregierte Bevölkerungsverteilungen auf räumliche Einheiten zusammengefasst und den Daten aus der Statistik Austria gegenübergestellt. Zusätzlich wurde eine flächenproportionale Verteilung der Bevölkerung auf die CLC Klasse 1.1 (residential areas) und auf die gesamte Fläche der statistischen Einheiten durchgeführt. Daraus ergibt sich die Möglichkeit, den Effekt der räumlichen Disaggregation auf verschiedenen Stufen zu zeigen. Als räumliche Referenzeinheiten wurden für Niederösterreich die Bezirke und für Vorarlberg die Gemeinden herangezogen.

Tabelle 2 zeigt den mittleren Fehler (rms) der Abweichungen der einzelnen Disaggregationsergebnisse von den statistischen Referenzdaten. Die Angaben in Prozent bezeichnen das Verhältnis des mittleren Fehlers zur durchschnittlichen Bevölkerung der Bezirke oder Gemeinden. Während die Verteilung auf Bezirks- oder Gemeindeflächen erwartungsgemäß stark von den Referenzdaten der Statistik Austria abweicht, bringt die Verteilung auf Grundlage der CLC Klasse 1.1 schon ein deutlich besseres Ergebnis. Die beste Übereinstimmung liefert eindeutig die räumliche Disaggregation basierend auf den Bebauungsdichten. Je kleiner die Referenzeinheiten sind, desto größer wird auch der relative Fehler, d.h. die Ergebnisse für Gemeinden sind relativ „ungenauer“ als die Ergebnisse für Bezirke.

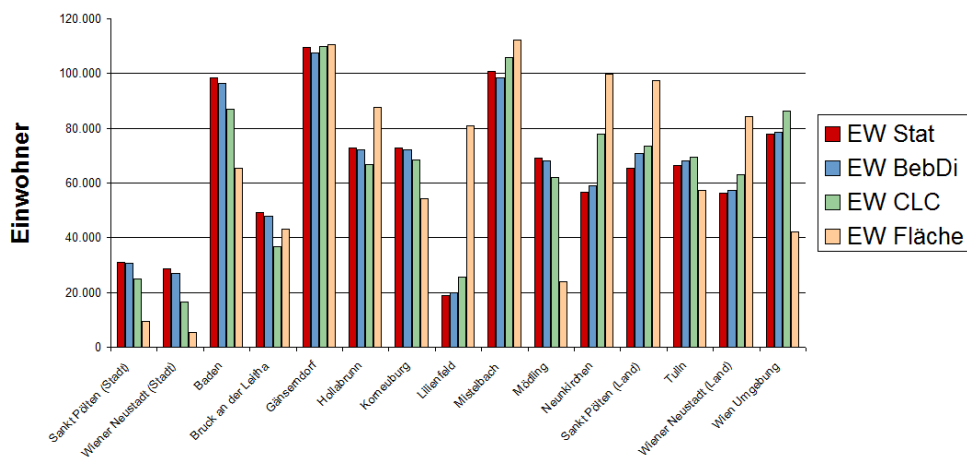
**Tab. 2:** Qualität der Ergebnisse der räumlichen Disaggregation

	BebDi		CLC		Fläche	
	rms [EW]	%	rms [EW]	%	rms [EW]	%
NÖ Ost (Bezirke)	2.049	3,2	9.296	14,3	30.312	46,7
Vbg (Gemeinden)	857	23,5	1.726	47,3	6.744	184,8

Abbildung 2 zeigt die Ergebnisse für die Bezirke im östlichen Niederösterreich im Detail. Die erste Säule des Diagramms repräsentiert die Einwohner laut Statistik Austria (EW Stat)

und damit den Referenzwert. Die räumliche Disaggregation auf Grundlage von drei Bebauungsdichteklassen ist in der zweiten Säule dargestellt (EW BebDi). Die dritte Säule zeigt das Ergebnis der räumlichen Disaggregation auf Grundlage einer einzigen Bebauungsklasse aus dem CLC Datensatz (EW CLC). Die vierte Säule ist das Ergebnis der Bevölkerungsverteilung proportional zu den Bezirksflächen (EW Fläche) und entspricht damit der Ausgangssituation von einem Bevölkerungswert für die gesamte Region.

Die Ergebnisse bestätigen die Aussagen aus Tabelle 2. Zusätzlich ist zu erkennen, dass es in Bezirken mit einem hohen Bebauungsanteil wie St. Pölten (Stadt) oder Wr. Neustadt (Stadt) tendenziell zu einer Unterschätzung der Bevölkerung kommt, während in ländlichen Bezirken wie St. Pölten (Land) oder Neunkirchen die Bevölkerung eher überschätzt wird. Das dürfte darauf zurückzuführen sein, dass ab einer gewissen (hohen) Bebauungsdichte die Bevölkerungsdichte stärker zunimmt als die Bebauungsdichte. Ein Grund dafür könnte die zunehmende Bebauungshöhe in Stadtzentren sein, die nicht erfasst werden kann.



**Abb. 2:** Vergleich der räumlichen Disaggregation von Bevölkerungszahlen mit Daten der Statistik Austria auf Bezirksebene in Niederösterreich

Die Untersuchungen haben bestätigt, dass die räumliche Disaggregation von Bevölkerungsdaten sinnvolle Ergebnisse liefert, wobei die Genauigkeit der resultierenden räumlichen Bevölkerungsverteilung vom Detailgrad der Bebauungsdichtedaten abhängt.

## Acknowledgment

Das Projekt geoland wird im Rahmen der GMES (Global Monitoring for Environment and Security) Initiative im 6. RP der EU gefördert (Contract No SIP3-CT-2003-502871) und von Infoterra GmbH, Friedrichshafen, und Medias France, Toulouse, koordiniert.

## **Literatur**

- Chen K. (2002): An approach to linking remotely sensed data and areal census data. *Int. J. Remote Sensing*, Vol. 23, No. 1, pp. 37-48.
- Kaminger I. & Wonka E. (2004): Von einer Österreichgliederung nach Gemeinden zu Planquadraten: Statistik Austria erweitert sein regionalstatistisches Angebot. In (Manfred Schrenk Hrsg.): *CORP 2004 Proceedings*, pp. 549-554.
- Steinnocher K. & Köstl M. (2002): Verdichtung oder Zersiedelung? Eine Analyse des Flächenverbrauchs im Umland von Wien. In (Manfred Schrenk Hrsg.): *CORP 2002 Proceedings*, pp. 193-200.
- Steinnocher K., Gangkofner U., Hoffmann C., Köstl M., Petrini-Monteferrri F., Tötzer T. (2005): Spatial planning indicators – the geoland approach. In (Manfred Schrenk Edt.): *CORP 2005 Proceedings*, pp. 185-191.