Comparing the GEOSTAT population grid 2006 data with municipal data in Germany

Dr. Udo Maack KOSIS-Verbund, Germany



EUROPEAN FORUM FOR GEOSTATISTICS SOFIA CONFERENCE 2013 23 - 25 October, Sofia, Bulgaria







GISCO Meeting March 2009:

The Geostat Hypotheses;

In search of a high resolution foundation for Geostatistics in Europe; an interactive response to operational user needs in a time of transition.

Lars H. Backer

Statistics Sweden Nordic Forum for GeoStatistics (NFGS) European Forum for GeoStatistics (EFGS) European GridClub

a Geostat 1 Backgroundpaper

7. "ESS_grid"

- We have a well functioning statistical system of statistics aggregated to administrative areas designed and built serve "accounting" purposes.
- For building a global infrastructure of spatial information these traditional statistical systems represents an obstacle. Their data are **not suitable for the** "engineering" purposes required.
- 3. The current global challenge will a statistical systems that allows for integrated spatial analysis that involves qualified statistics from any available source. This calls for a shared of statistics that covers the whole earth, is liberated from administrative areas, and that allows for hierarchical analysis of datasets on all scale ranges from global to local.



Project Investigations:

- A. Population by Pan European Population Grid and Population registered
- B. Relation between Grids and Small Area Classifications (SAC's) used in German municipalities
- C. Tools and use cases
- D. Conclusions and Advises



Grid-Data Project



Roadmap (revised)

		2012								2013																						
bis Se	ptember		Okto	ober		No	over	nbe	r	C)eze	mbe	r		Ja	nua	ar		Fe	bruar	·		/lärz	Ар	ril	Ī	Vov	emb	ber	Dez	emb	ber
Woche	39	40			4	4				49				1					6		1(0		14			46			49	50	
Projekt-Vorbereitun	g																															
Entscheidung für da	s Projekt																															
Pilotphase																																1
Dateneinsammeln																																
Vorbereiten Daten																										_						
Workshop																										_						
Erstellen / Abstimm	en Berich	nt																														1
Ergebnispräsentatio	n	[



Grid-Data Project







Spatial Data Base per City (Subcity-Classification)

									Nbg /		
Spatial Units	DD	F	FR	GE	HB	L	М	MA	Fü	S	WOB
1kmRaster	394	365	204	139	393	362	371	185	311	257	253
100mRaster	33.662	25.682	15.893	10.844	32.695	30.524	31.786	14.924	25.252	21.321	21.036
EuroPopGrid	345	263	149	135	308	302	357	150	285	230	154
Citv	1	1	1	1	1	1	1	1	2	1	1
Stadtbezirke		121	42	5		10	27	17	105	23	
Stadtbereiche			6	18							
Stadtbezirksteile							108				
Distrikte								387			
Stadt-/Ortsteile	64		29		90	63				152	42
Stadtviertel										319	
Stat. Bezirke	388		161			310					
Rege/GBL/				40/84							
MBL				370							
Baublöcke	11.622	5.943	1.636	1.768	5.957	4.013	10.348	2.803	4.693	6.048	1.998
Baublockseiten						17.504					
Urban Atlas	Х	Х	Х		Х	Х	Х		Х	Х	Z









9





How to produce comparable Data?

City	Grids within City boundaries	Grids popu- lated	Population in Grids	Grids truncated	Pop within City Limits	registered Population
DD	394	345	529.613	114	496.288	487.199
F	365	263	724.738	97	660.502	627.826
FR	204	149	229.039	70	215.189	194.137
GE	139	135	344.740	64	269.015	256.990
HB	393	308	578.833	113	537.378	558.018
L	362	302	524.477	85	499.762	489.099
М	371	357	1.313.226	113	1.243.148	1.283.855
MA	185	150	332.436	50	306.038	295.876
NF	311	285	634.230	163	598.040	608.346
S	257	230	628.963	74	586.096	556.710
WOB	253	154	124.950	41	120.706	120.709













disaggregated Population







disaggregated Population







- Population not represented <u>by</u> Disaggregation
- Population distribution in 100x100m Grids
 - disaggregated Population





Blockareas overlaying disaggregated populated Grids (only blocks with buildings)







Differences between aggregated and disaggregated values

860,938478 - 2900,891113







Error estimation:

 $\frac{\text{Difference per Grid:}}{R_s = \dot{P}_s - P_s}$

City	relErrorTAE
DD	37%
F	33%
FR	31%
GE	22%
HB	33%
L	45%
MA	39%
М	40%
NF	39%
S	29%
WOB	37%

Acceptable ?





From Geostat 1A Report:

A quality assessment of the disaggregated population grid carried out by AIT showed that the relative error at national level for those countries where reference register data are available is in the order of 30% but can exceed 50% for sparsely populated countries

such as NO, SE and FI.

How was this performed?





Data handling of border grids

Pan European Population Grid Data:

GE	OSTAT_grid_EU_P	POP_2006_1K	(_V1_1_1	.csv			×
	GRD ID	POP TOT	YEAR	METHD CL	CNTR CODE	DATA SRC	
	1kmN1440E4725	3620	2006	D	МТ	АЛТ	
	1kmN3154E3858	26	2006	М	NL(8):BE(18)	AIT:NL	
	1kmN1440E4726	1070	2006	D	MT	AIT	

INSPIRE conform - Grid Data:

А	В	С	D	E	F	G
GridCode	PercentageOfArea	POP_TOT	YEAR	METHD_CL	CNTR_CODE	DATA_SRC
RES1000mN3154000E3858000	45	8	2006	А	NL	NL
RES1000mN3154000E3858000	55	18	2006	D	BE	AIT
RES1000mN1440000E4726000	100	1070	2006	D	MT	AIT





Municipal Situation:



10 Stadtbezirke / Hektarraster 2.616 von 30.524 - 27.908 freie Raster



4.013 Baublöcke / Hektarraster 16.427 von 30.524 - 14.097 freie Raster





First-Conclusion (for Block A)

- Grids are useful for an easy harmonization of reference areas
- Grids are useful for a certain class of applications
- Grids not able to provide statistical data on user territorial classifications without loosing quality
- Grids <u>can not</u> replace the vector representation of statistical / planning / administrative units at least on municipal level







Two-Pillar-Model of municipal georeferencing





How to characterize the relation between the two pillars?

- Quantitative How to measure?
- Visual By which rules?
- Advantages and Disadvantages?





Visual comparison



Population density on Baublock-Level



Visual comparison





Visual comparison

Population density on **Baublock-**Level (only Areas < 13 ha) and 100mx100m Grids (contrasted colors)





Mixed referenced Units







SAC vs. Grids

Evaluation:

Advantages of Grids

Disadvantages of Grids

Advantages of SAC-Units

Disadvantages of SAC-Units

Results will be published at beginning of 2014





Displaying

Displaying Single values per 100mx100m Grids







Displaying

Smoothing with neighboring grids



Weightingfactors: 1,0 / 1,0 / 1,0







Displaying

Smoothing with neighboring grids



Weightingfactors: 1,0 / 0,5 / 0,2







Calculation of potentiality









accessible potential and barriers

	Leg	gen	de
0,2	unt	ter	0,5
1		0,5	- 1
1,2		2	- 3
2,5		3	- 4
3,5		4	- 5
4,5		5	- 6
5,5		6	- 7
7		7	- 8
8,5		8	- 9
9,5		9 -	10
10			10

0,4	0,4	1	1	1	1	1	1	1	1	1	1	1	0,4	0,4
0,4	1	1	1	1,2	1,4	1,6	1,7	1,6	1,4	1,2	1	1	1	0,4
1	1	1	1,3	1,8	2,4	2,9	3,1	2,9	2,4	1,8	1,3	1	1	1
1	1	1,3	2	3,1	4,5	5,9	6,4	5,9	4,5	3,1	2	1,3	1	1
1	1,2	1,8	3,1	5,4	8,3	9,9	10	9,9	8,3	5,4	3,1	1,8	1,2	1
1	1,4	2,4	4,5	8,3	10	10	10	10	10	8,3	4,5	2,4	1,4	1
1	1,6	2,9	A ?	9,9	10	10	10	10	10	9,9	5,9	2,9	1,6	1
1	1,6	2,9	5,5	6,5	10	10	10	10	10	10	6,4	8,1	1,7	1
1	1,3	2,1	2,9	2,9	2,1		10	10	10	9,9	5,9	2,9	1,6	1
1	1	1,3	1,6	1,6	1,3	1	1	3	10	8,3	4,5	2,4	1,4	1
1	1	1	1	1	1	1	1	1	1,3	54	3,1	1,8	1,2	1
0,4	1	1	1	1	1	1	1	1	1,3	2	2	1,3	1	1
0,4	0,4	0,4	0,5	0,5	0,4	0,4	1	1	1	1,3	1,3	1	1	1
0,3	0,3	0,3	0,4	0,4	0,3	0,3	0,4	1	1	1	1	1	1	0,4
0,2	0,2	0,3	0,3	0,3	0,2	0,3	0,4	0,4	1	1	1	1	0,4	0,4
	Po	ten	zia	l (S	um	me	e üb	er	alle	e Ze	eller	ר):	59	95
								da	s ei	ntsp	oric	ht:	77	%
			im \	/erg	leic	h zı	ur R	lech	nun	g "o	ohne	e Ba	arrie	re",
	1,7	%-	Pun	ikte	da∖	on t	tür d	lie b	eid	en g	jete	Iter	ı Ze	llen
											© 20	13,⊦	lans l	Meng



European Forum for Geo Grid Statistics





Thank you