Working document

Activity 1.3
Integration of particular land related data sets

Draft
Version 1.0

Prepared by:
ETCLUSI (GISAT, GeoVille)

18.10.2010
Version 1.0

Project manager:
Branislav Olah

Generalitat de Catalunya
Departament de Medi Ambient i Habitatge

ETC Land Use and Spatial Information

Universitat Autònoma de Barcelona
Edifici C – Torre C5
4ª planta
08193 Bellaterrra (Barcelona)
Spain

Contact: andreas.littkopf@uab.es

European Environment Agency

1.2 Population data

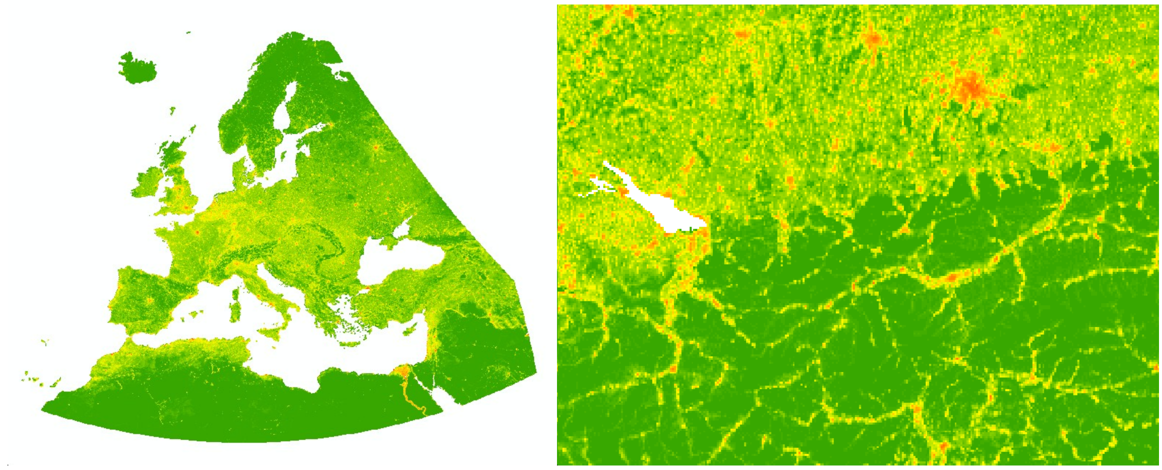
The objective of this task is to match the LandScan Global Population Database 2008 with the 1km² cells of the LEAC grid. The specific challenge here is to use a method that transforms the population values from the LandScan raster (provided in Geographic projection on WGS84 datum) into the LEAC raster.

The output is a table in dBase format, that can be joined to geospatial LEAC representation using a unique reference ID.

Data source:

Input data set is the LandScan Global Population layer, which is a worldwide population database at 30" X 30" (approx. 1 x 1 km) resolution with a single population value for each cell. Data are provided in ESRI grid format and have been clipped to cover entire Europe (EEA-38).

Spatial representation of the LandScan data base is by areas (regular 30" grid) with continuous attributes (population counts).



Left: LandScan population data clipped to Europe. Right: Detailed view of Central Europe (green: low population values; red: high population values)

Target statistics:

The LandScan data base is used to create the following indicators for each LEAC 1km² grid cell:

- Population count per grid cell
- Sealed area/capita (total sealed area per cell divided by population count per cell, [m²/pop])
- Population per sealed area [pop/ha] per grid cell, population density (population count per cell divided by total sealed area per cell)

Reference data (units to be characterized by above information/information to be linked to):

- 1km (LEAC) grid

Coverage:

- 38 countries, ca. 6 mio sqkm (EEA-32 + 6)

Delivery format:

- dBase tables with reference-ID allowing linkage to LEAC grid

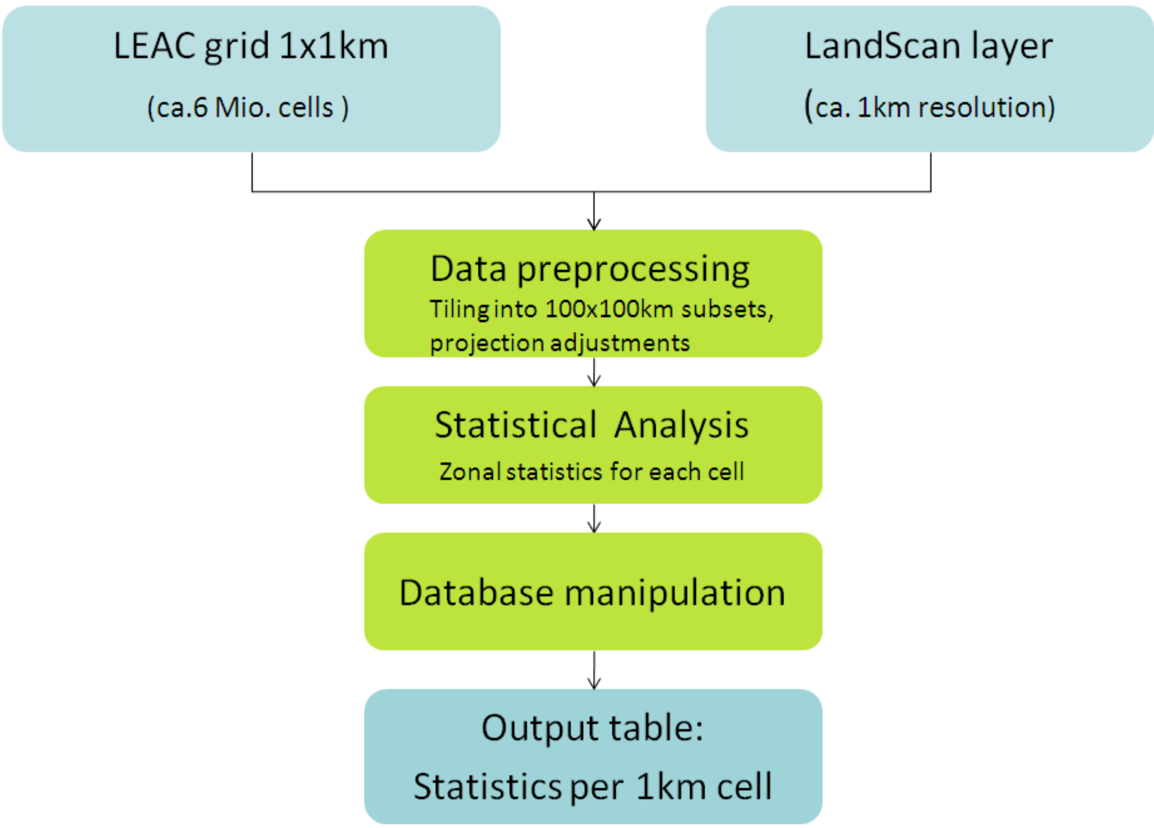
1.2.1 Data processing

Basically the data integration method for the LandScan data is a transfer of the population counts from one grid (LandScan) into another grid (LEAC) with slightly different geographic registration. While the LEAC grid in LAEA projection has regular 1km x 1km grid cells, the LandScan grid shows slightly different extents of its cells due to its geographic projection.

In any case, the original population counts from LandScan must be preserved to the highest possible extent, and resulting population sums for entire regions must be identical.

In order to correct the geometric mismatch of the grid origins and orientation, and to preserve the original counts as much as possible, some adjustments have to be made. Data processing is carried out in the original geographic projection (WGS84) of LandScan, as a reprojection of LandScan to LAEA and subsequent resampling would alter the population values significantly. Therefore the LEAC grid is reprojected to geographical projection and the population counts from the LandScan data are transferred using the "zonal statistics as table" function.

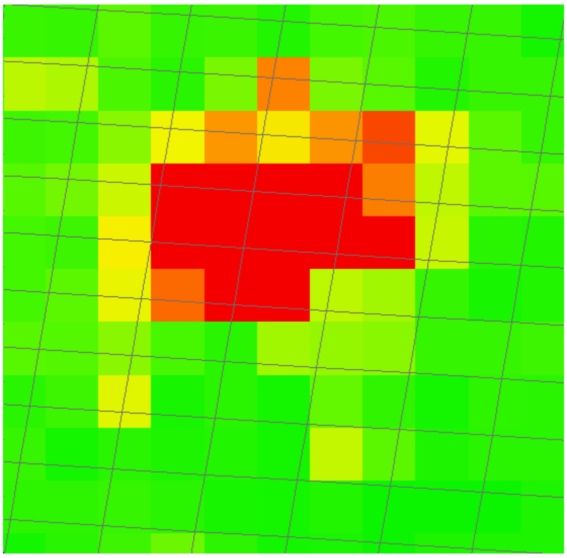
For data processing, the same 100x100km tile structure has to be established as for the imperviousness layer (see chapter 2.1). An overview of the processing steps of the workflow is given below.



Workflow for integration of LandScan data into the LEAC grid.

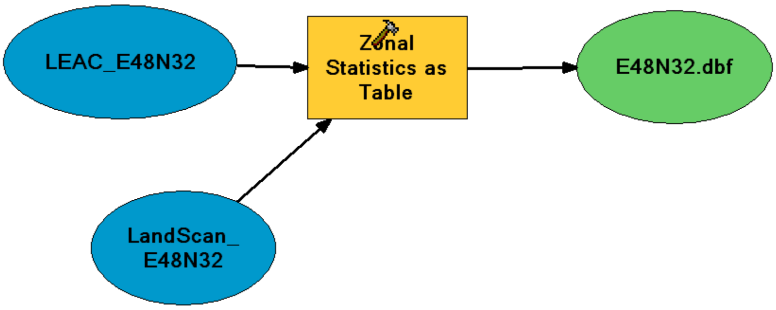
In a first step the global LandScan data are clipped to European coverage. For better performance the same tiling procedure as for the imperviousness layer is performed in a second step using the previously generated 880 tiles of the LEAC grid (100x100km, cf.

chapter 2.1). These tiles have been reprojected from ETRS_1989_LAEA to GCS / WGS 84 to use them as a clip extent for the LandScan data. The results are 880 tiles of the LandScan data, which intersect one by one with the 880 tiles of the LEAC 1km grid. Now further analyzing steps are performed using a batch process.



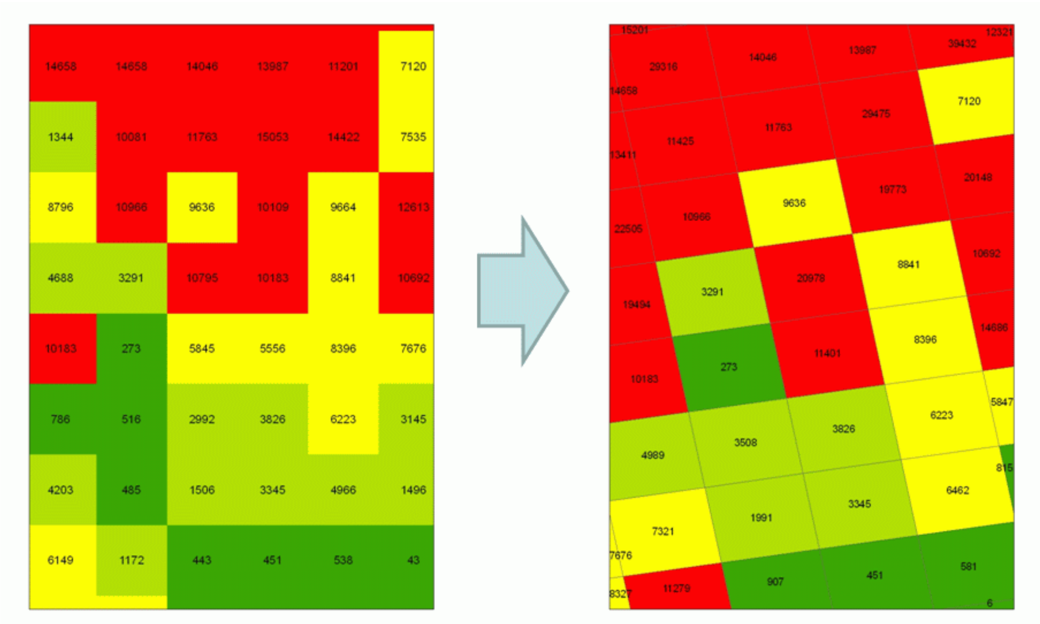
Overlay of the reprojected LEAC grid (black outlines) on the LandScan raster

Based on the above considerations for data integration the ArcGIS tool **"zonal statistics as table"** is executed, using each 1km LEAC grid cells as a zoning feature. NoData values are ignored. All processing steps are repeated 880 times using the batch tool of the ArcGIS Model Builder.



ArcGIS model for integrating the LandScan population values with the LEAC grid.

"Zonal statistics as table" calculates the statistics for each zone (each LEAC cell) resulting in a table where each record contains the population sum of all LandScan cells that have their centroid inside the overlaying LEAC cell. This value represents the new calculated population count for each LEAC cell. The data are potentially altered at grid cell level, but the overall sum for larger areas always remains identical.



Example of LandScan data integration. Original LandScan data (left), LEAC grid cells with integrated data (right)

In a final step the tables are integrated into a MS Access data base. The individual tables are merged automatically and joined to the final result table created for the imperviousness layer using the unique LEAC ID (chapter 2.1).

Three columns are added:

- POPCOUNT: Population value
- AREA_POP: Sealed area/capita (total sealed area per cell divided by population count per cell, [m²/pop])
- POP_AREA: Population per sealed area [pop/ha] per grid cell, population density (population count per cell divided by total sealed area per cell)

In case of null values either for population or for sealed area, cells are marked with a value of 99999. LEAC grid cells exceeding the coverage of the LandScan layer also have a value of 99999.

	A	B	C	D	E	F	G	H	I	J	K	L
1	CELLCODE	POPCOUNT	POP_AREA	AREA_POP	COUNT	AREA	MIN	MAX	MEAN	STD	MEDIAN	S9_COUNT
2	E4300N3300	6	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
3	E4300N3301	4	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
4	E4300N3302	94	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
5	E4300N3303	333	0,00190	526,05405	2500	175176,00	0	100	17,52	34,79	0	547
6	E4300N3304	395	0,00221	452,93165	2500	178908,00	0	100	17,89	37,34	0	476
7	E4300N3305	660	0,00150	664,86061	2500	438808,00	0	100	43,88	45,43	31	1262
8	E4300N3306	778	0,00182	549,16710	2500	427252,00	0	100	42,73	43,80	34	1302
9	E4300N3307	452	0,00163	615,05310	2500	278004,00	0	100	27,80	37,89	0	943
10	E4300N3308	103	0,00374	267,65049	2500	27568,00	0	100	2,76	14,01	0	100
11	E4300N3309	26	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
12	E4300N3310	3	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
13	E4300N3311	0	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
14	E4300N3312	147	0,01992	50,20408	2500	7380,00	0	100	0,74	8,05	0	22
15	E4300N3313	13	0,00038	2639,38462	2500	34312,00	0	100	3,43	16,44	0	112
16	E4300N3314	3	0,00008	11961,33333	2500	35884,00	0	100	3,59	16,33	0	127
17	E4300N3315	6	0,00014	6976,00000	2500	41856,00	0	100	4,19	19,15	0	118
18	E4300N3316	9	0,00262	382,22222	2500	3440,00	0	100	0,34	5,53	0	10
19	E4300N3317	5	0,00470	212,80000	2500	1064,00	0	75	0,11	2,67	0	4
20	E4300N3318	2	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
21	E4300N3319	1	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
22	E4300N3320	3	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
23	E4300N3321	3	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
24	E4300N3322	6	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
25	E4300N3323	15	0,00403	248,00000	2500	3720,00	0	99	0,37	5,63	0	11
26	E4300N3324	491	0,00198	504,30957	2500	247616,00	0	100	24,76	34,17	0	931
27	E4300N3325	39	0,00178	562,15385	2500	21924,00	0	100	2,19	13,11	0	72
28	E4300N3326	15	0,00082	1221,33333	2500	18320,00	0	100	1,83	12,39	0	55
29	E4300N3327	15	0,00251	398,40000	2500	5976,00	0	100	0,60	7,48	0	16
30	E4300N3328	96	0,00392	255,16667	2500	24496,00	0	100	2,45	14,00	0	81
31	E4300N3329	155	0,00639	156,43871	2500	24248,00	0	98	2,42	12,90	0	91
32	E4300N3330	6	0,00268	373,33333	2500	2240,00	0	82	0,22	3,68	0	10
33	E4300N3331	5	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
34	E4300N3332	4	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
35	E4300N3333	108	0,00387	258,59259	2500	27928,00	0	100	2,79	14,93	0	92
36	E4300N3334	37	0,00000	0,00000	2500	0,00	0	0	0,00	0,00	0	0
37	E4300N3335	3	0,00022	4650,66667	2500	13952,00	0	100	1,40	10,72	0	45
38	E4300N3336	140	0,00559	179,02857	2500	25064,00	0	100	2,51	13,87	0	84

Final output table showing combined indicator information from population and imperviousness layers.

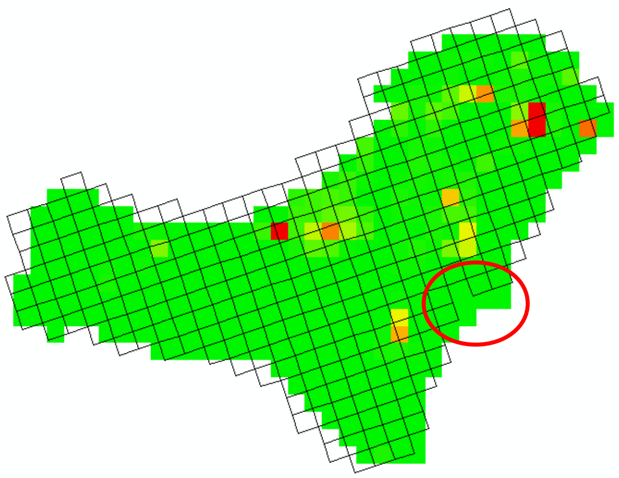
1.2.2 Problems and mitigation

Again limitations with regard to processing time and capacity have to be considered. It is not possible to operate the process with one single model run. Therefore input data have been tiled again based on the EEA 100km reference grid.

In a first try the spatial join function of ArcGIS was tested. There a documented bug of the spatial join tool causes problems, as in batch mode not all attribute tables are populated correctly. This problem could be solved by using the most recent ArcGIS version 10.0, where it seems that this bug has been eliminated. However, by using this method the LandScan population counts are substantially altered due to resampling. For this reason it was decided work in the original LandScan projection and to use the "zonal statistics as table" function for getting more accurate results.

One aspect with the "zonal functions" procedure is that this operation results in a sum of the population counts from those LandScan cells which centroids' lie inside the LEAC grid cell, and not in an area-weighted population count. Introducing an area-weighting procedure for entire Europe would again lead to problems with processing capacity.

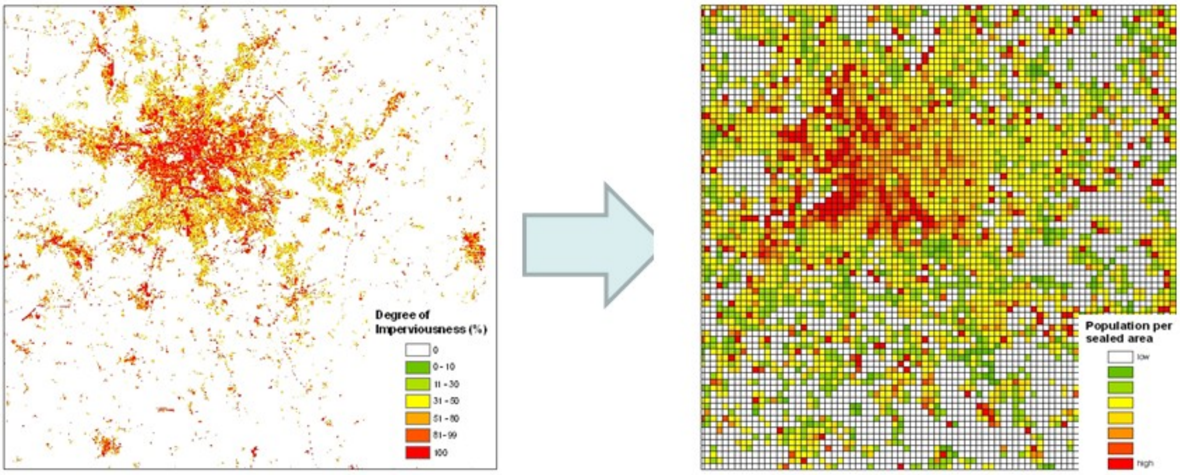
Another issue is illustrated in the figure below. Some populated LandScan cells are not covered by LEAC 1km grid cells, and vice versa. These values are obviously lost in the integration process. The reason for these missing LEAC cells should be clarified.



Overlay of the LEAC grid (black outlines) on the LandScan raster with missing LEAC cells (circle)

1.2.3 Output

The resulting attribute table has been joined with the output table created for the imperviousness layer and exported to DBF format. This table is joined to the LEAC grid using the reference ID for further analysis and map making. The same limitations apply as described for the imperviousness data base (see chapter 2.4.1).



Result example for the city Berlin: Degree of imperviousness in original 20m resolution (left) and resulting population density indicator (right)