Land cover modeling in the Travis County

Contents

Prepared for: GIS and Water Resources

By Julie C Faure, December 4 2015

Introduction

In a lot of scientific papers, we can see that a correlation is made between land cover and population characteristics. It has been a preoccupation for several scientists for a few decades. This correlation is namely important for several environmental aspects. For example, land imperviousness that is induced by urbanization can have harmful impacts on water runoff, especially when it is combined to climate change problems. The conservation of this ratio can be essential to the equilibrium of the water resources of the area. For example Urich, Bach et al. [2015] studied the way that urbanization and climate change can affect the runoff of the city, and, by examining different scenarios, gave some ideas of the way the problem could be reduced.

I am currently a research assistant and am studying the effect of population change in urban areas on water infrastructures and sanitation systems. So, my first idea for this final project was to find a correlation between changes in the structure of a city and its water supplies. To do so, I wanted to pick a city that is growing, and look at the population and infrastructure characteristics, to find a link between some of these characteristics and its water supplies. The Travis County was a good fit, as seen in figures 1 and 2.

Figure 1: National Land Cover Database 2006-2011 Percent Developed Imperviousness Change in Texas

The red dots correspond to areas where the percent developed imperviousness changes between 2006 and 2011.

The upper red zone corresponds to Austin and its agglomeration in which is located the Travis county.

Figure 2: National Land Cover Database 2006-2011 Percent Developed Imperviousness Change in the Travis County

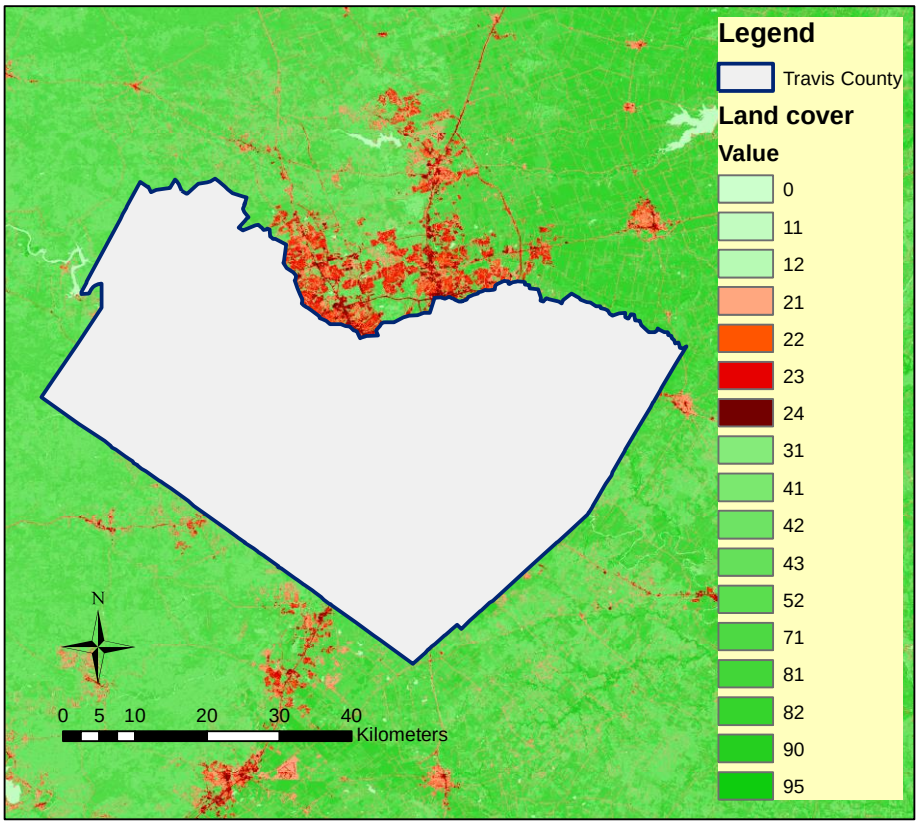
This was actually a lot to do and I decided to focus on the effect of population and urban infrastructures on land cover characteristics of the area.

Data collection

Land cover data

The land cover data I used were taken from the National Land Cover Database (NLCD). The available databases were 1992, 2001, 2006 and 2011 databases (it would then be interesting to use my model to predict the 2015 data for example). In these databases, I could find the land cover distribution in the Travis County, where each value corresponds to a land cover characteristic (e.g. 'highly urbanized', 'open water',…). This land cover reparation is represented on figure 3.

Figure 3: NLCD 2011 Land cover in the Travis County



The red values correspond to the urbanized areas.

I also found the percent imperviousness for each 30m*30m raster cell of the county, this value is the value I actually used for my model, because I assumed that this was the value I could actually model, using the population and urban infrastructure characteristics.

Block Groups data

A block group is small subdivision of a Census tract, which is a subdivision of a County. There are 220,000 census block groups in the United States

I could download these block group polygon coordinates from the American Census Bureau. This gives a table like Table 1, and can be visualized on Figure 4.
The unique number of each block group is the GeoID. For example, the first line of the previous Table 1 corresponds to the GeoID: 48 453 000305 3

48: Texas
453: Travis County
000305: Census tract
3: the block group number inside the 000305 census tract

Table 1: Block Groups characteristics

Table

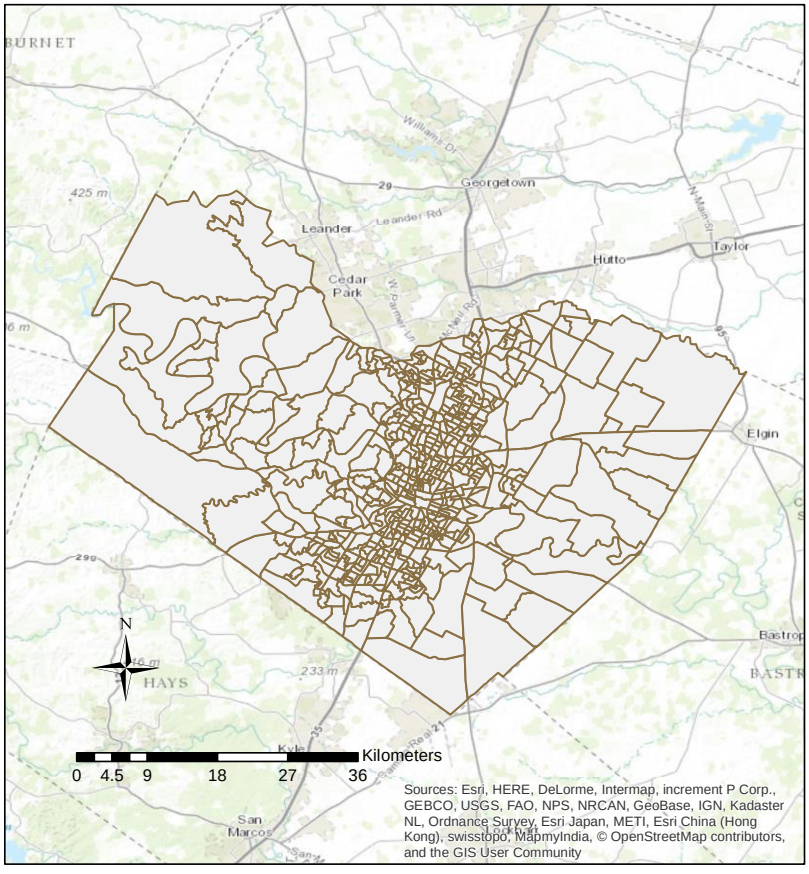
Block GroupsTravis County

OBJECTID *	Shape *	STATEFP	COUNTYFP	TRACTCE	BLKGRPCE	AFFGEOID	GEOID	NAME	LSAD	ALAND	AWATER
1	Polygon ZM	48	453	000305	3	1500000US484530003053	484530003053	3	BG	432264	0
2	Polygon ZM	48	453	001820	3	1500000US484530018203	484530018203	3	BG	742706	0
3	Polygon ZM	48	453	002312	2	1500000US484530023122	484530023122	2	BG	831259	0
4	Polygon ZM	48	453	001774	1	1500000US484530017741	484530017741	1	BG	583906	0
5	Polygon ZM	48	453	002413	1	1500000US484530024131	484530024131	1	BG	234241	0
6	Polygon ZM	48	453	002109	1	1500000US484530021091	484530021091	1	BG	856954	0
7	Polygon ZM	48	453	000402	3	1500000US484530004023	484530004023	3	BG	619672	0
8	Polygon ZM	48	453	001501	4	1500000US484530015014	484530015014	4	BG	1330394	0
9	Polygon ZM	48	453	001851	4	1500000US484530018514	484530018514	4	BG	2359891	0
10	Polygon ZM	48	453	001823	1	1500000US484530018231	484530018231	1	BG	510749	0
11	Polygon ZM	48	453	000902	3	1500000US484530009023	484530009023	3	BG	531272	0
12	Polygon ZM	48	453	001304	4	1500000US484530013044	484530013044	4	BG	473234	0
13	Polygon ZM	48	453	001503	1	1500000US484530015031	484530015031	1	BG	641563	0
14	Polygon ZM	48	453	001605	1	1500000US484530016051	484530016051	1	BG	706807	0
15	Polygon ZM	48	453	002434	1	1500000US484530024341	484530024341	1	BG	5516058	6
16	Polygon ZM	48	453	002431	2	1500000US484530024312	484530024312	2	BG	5986896	0
17	Polygon ZM	48	453	002104	2	1500000US484530021042	484530021042	2	BG	537632	0
18	Polygon ZM	48	453	001737	1	1500000US484530017371	484530017371	1	BG	851708	0
19	Polygon ZM	48	453	001848	3	1500000US484530018483	484530018483	3	BG	1887135	0
20	Polygon ZM	48	453	001776	2	1500000US484530017762	484530017762	2	BG	625181	0
21	Polygon ZM	48	453	002112	1	1500000US484530021121	484530021121	1	BG	394668	0

1 (0 out of 580 Selected)

Block GroupsTravis County

Figure 4 : Block groups in the Travis County



Legend

Block GroupsTravis County

Population data

The greatest source of data about population in the United States seems to be the American Community Survey from the American Census Bureau.

The only American Community Survey data that are available on the Census Bureau website at a block group level are the 2013 and 2012 five years estimates. The reason is that a block group is so small that they need time to have enough population information to be precise.

So I decided to build my land cover model for the Travis County using the 2011 National Land Cover Database and the American Community Survey 2013 five years estimates.

I downloaded several different population characteristics for each Block Group from this database, for example:
- total population
- median household income
- mean age
- number of person per housing
- number of housing units

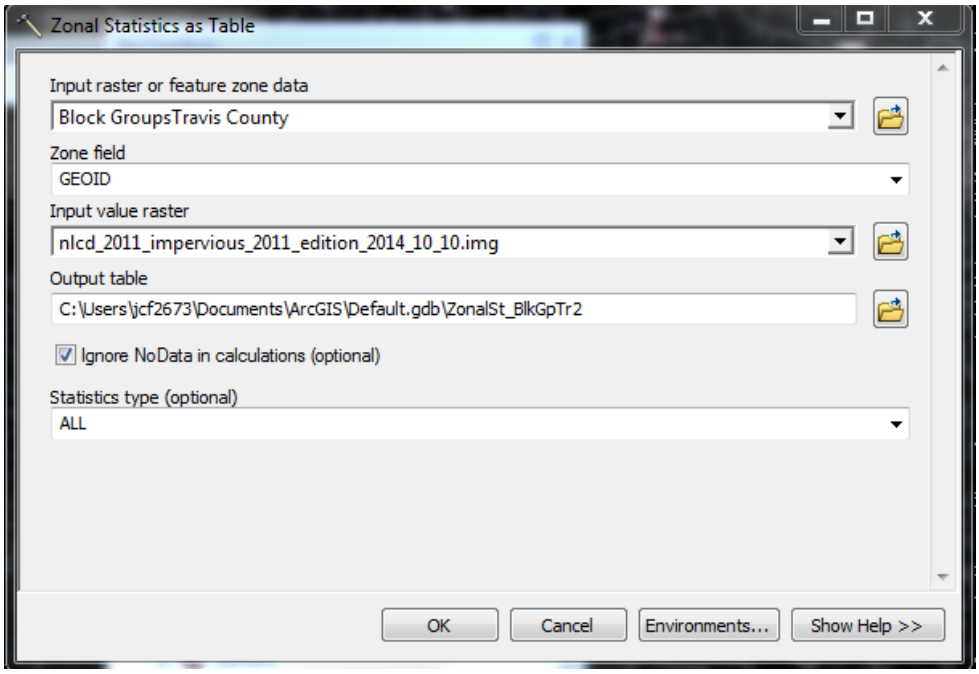
Data pre-processing

To use all these data, I had to pre-process them. I mainly created a giant table in which each line corresponds to a block group, with all the interesting population and land cover characteristics.

1) I joined the interesting population data to Travis County Block groups
2) I used the Spatial Analysis tool *Zonal statistics as Table* to create a table that gives the mean percent imperviousness for each GeoID (i.e. for each Block Group). The two next pictures show the way I used the tool and the result. In the table, MEAN is the mean percent imperviousness for each block group.

I could then obtain several maps (some of them are shown in Figure 5 to 8). And it enabled me to assume the population factors that were the most important for the percentage of imperviousness for each group.

I finally assumed that the percentage of imperviousness in each block was a function of:
- the total population per km^2
- the median income
- the number of housings per km^2



Table

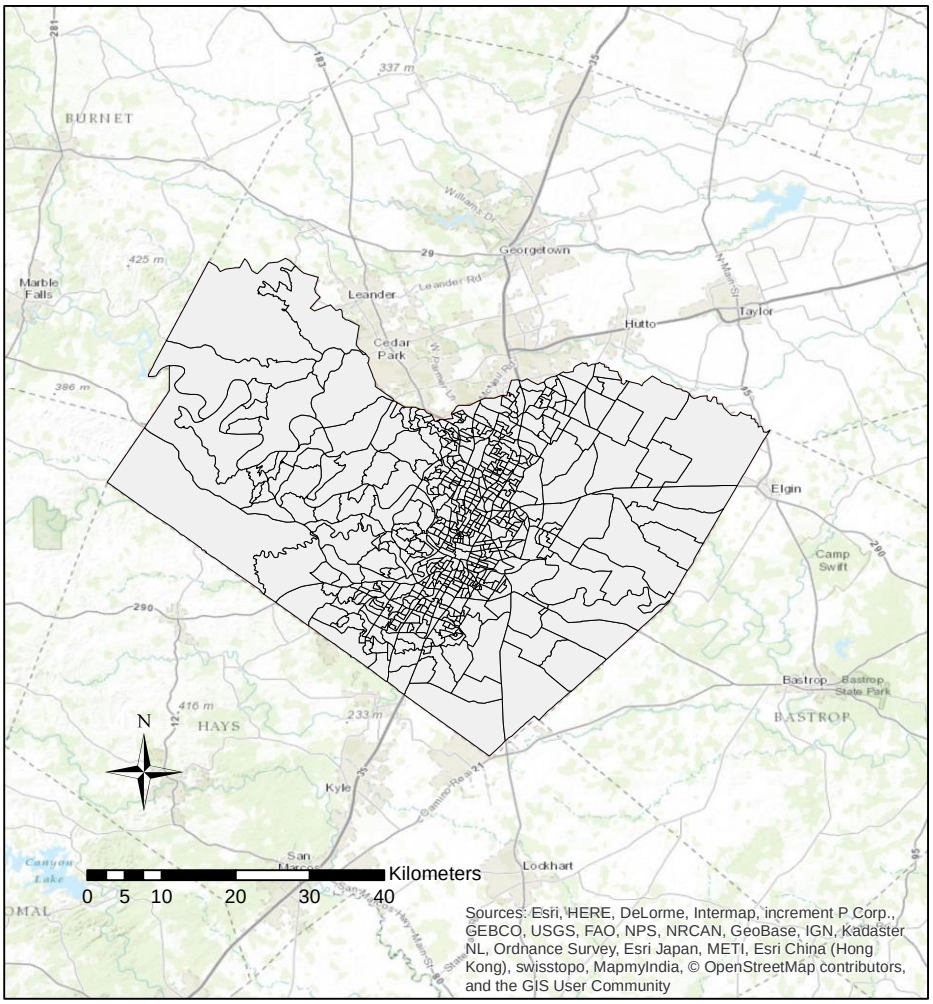
ZonalSt_BlkGpTr1

OBJECTID *	GEOID	ZONE_CODE	COUNT	AREA	MEAN
198	484530018192	198	244	219600	59.668033
199	484530021062	199	486	437400	22.903292
200	484530023121	200	829	746100	32.615199
201	484530023071	201	818	736200	28.701711
202	484530004012	202	1515	1363500	51.733333
203	484530003041	203	652	586800	48.552147
204	484530019173	204	1326	1193400	8.426094
205	484530017562	205	1013	911700	32.654492
206	484530018181	206	1088	979200	63.230699
207	484530018561	207	40662	3659580	4.724313
208	484530024291	208	2931	2637900	13.109519
209	484530017842	209	9098	8188200	7.78182
210	484530013042	210	383	344700	20.360313
211	484530017802	211	11843	1065870	5.153677
212	484530022085	212	1253	1127700	24.138867
213	484530017062	213	1026	923400	24.139376
214	484530005003	214	431	387900	24.146172
215	484530020022	215	528	475200	15.818182
216	484530018291	216	1107	996300	29.483288
217	484530018043	217	795	715500	65.738365
218	484530021102	218	679	611100	20.776141
219	484530018064	219	518	466200	57.322394
220	484530014021	220	641	576900	20.684867
221	484530021063	221	825	742500	25.917576
222	484530018555	222	1916	1724400	31.286534
223	484530017332	223	2866	2579400	9.144452
224	484530017502	224	3563	3206700	26.801572
225	484530017292	225	468	421200	24.141026

1 (0 out of 580 Selected)

ZonalSt_BlkGpTr1

Figure 5 : percent imperviousness in the Travis County



Legend

Block GroupsTravis County

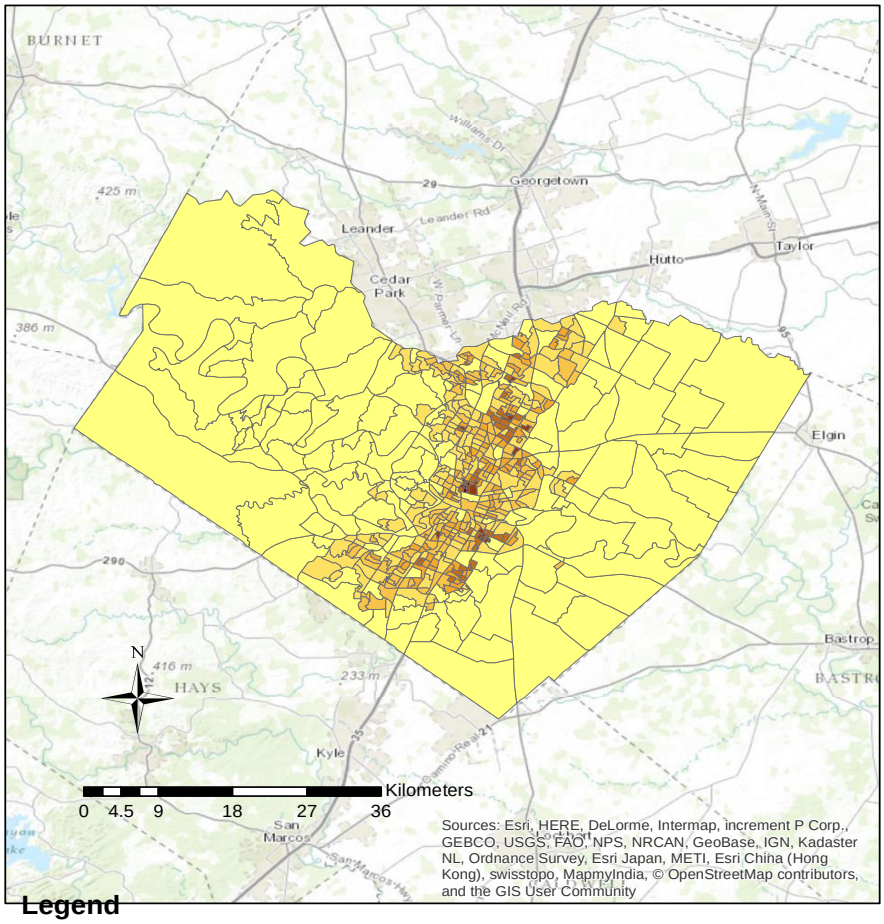
Mean Percent Imperviouness

Value

High : 80.4608

Low : 0.175074

Figure 6 : population per km^2

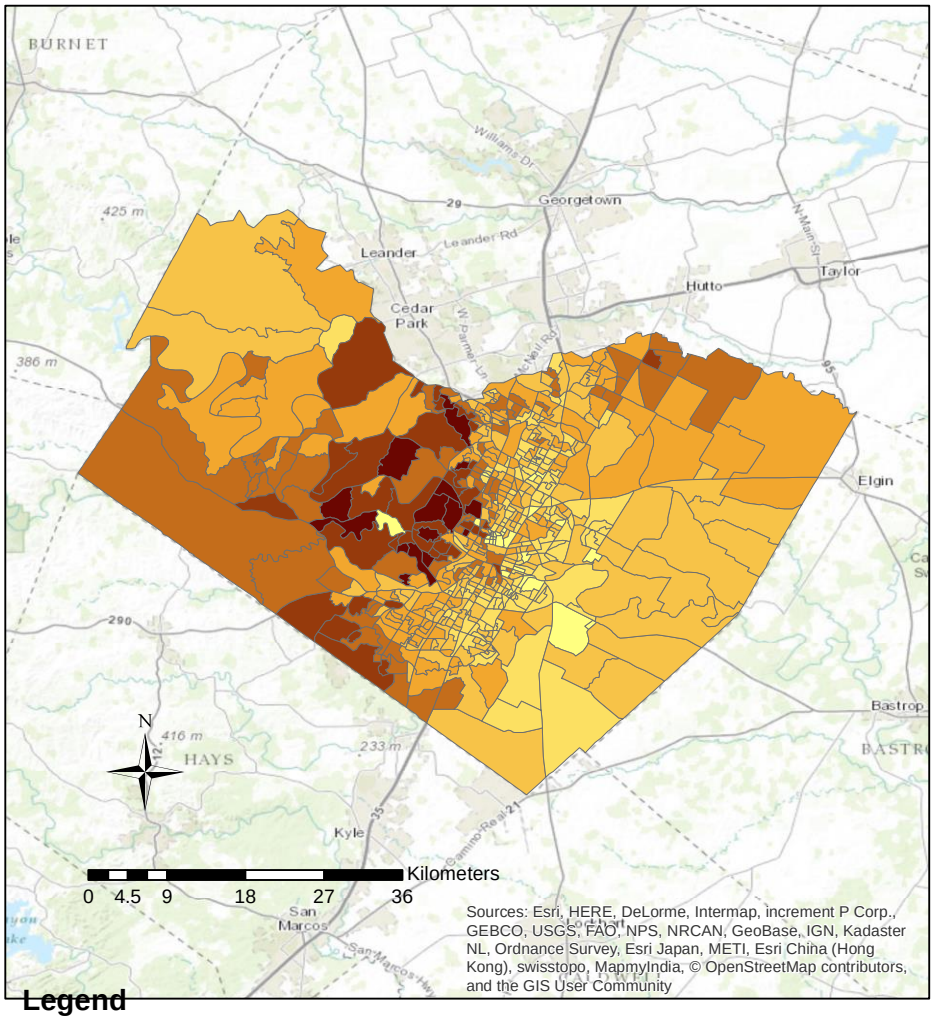


Legend

Population per km2

- 11.1369324 - 785.857484
- 785.857485 - 1610.99315
- 1610.99316 - 2441.31078
- 2441.31079 - 3669.51161
- 3669.51162 - 6054.68046
- 6054.68047 - 11399.6204
- 11399.6205 - 18839.9756

Figure 7: median annual household income

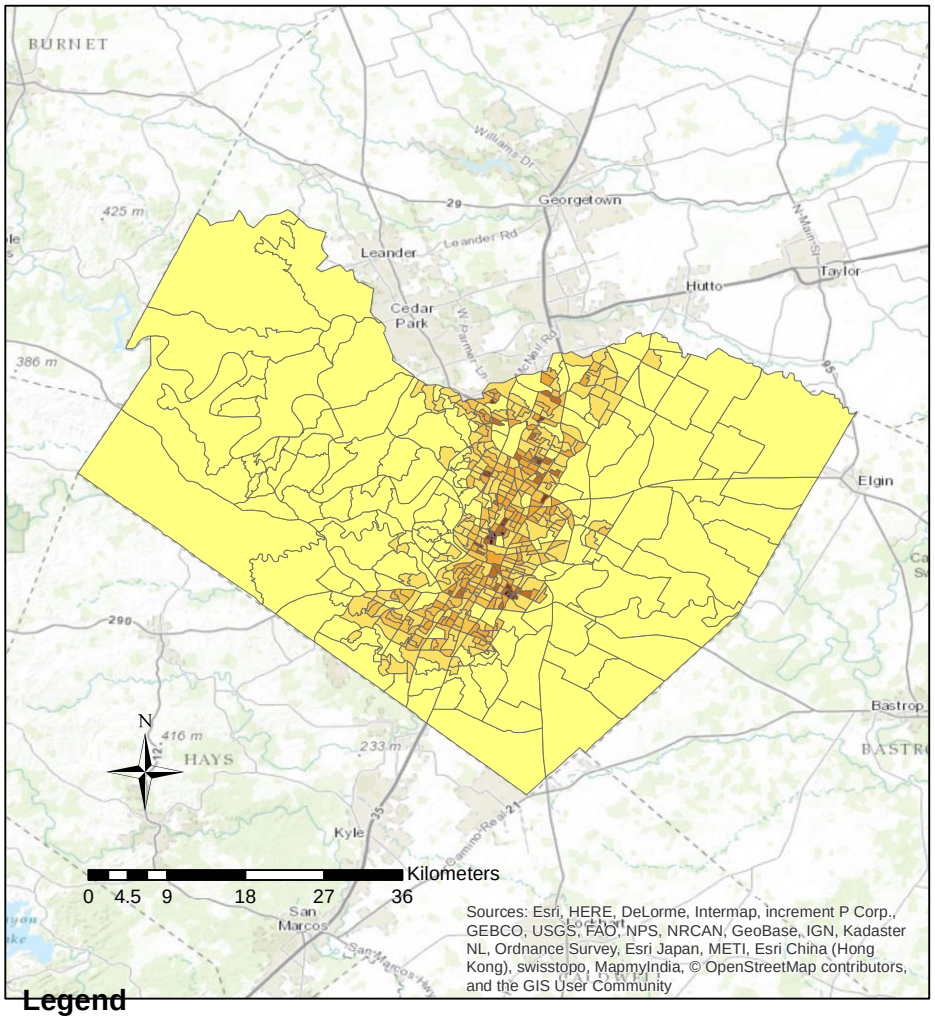


Legend

Median annual household income

- 0.00000000 - 21445.0000
- 21445.0001 - 40820.0000
- 40820.0001 - 59167.0000
- 59167.0001 - 83672.0000
- 83672.0001 - 114500.000
- 114500.001 - 154750.000
- 154750.001 - 247625.000

Figure 8: Number of housings per km^2



Legend

Housing units per km2

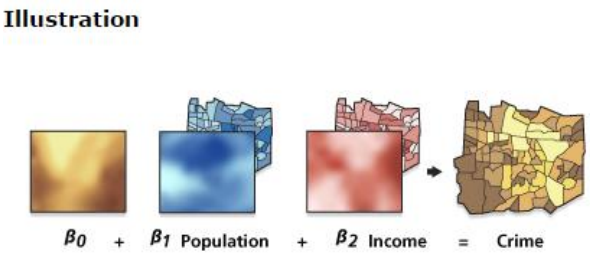
- 0.00000000 - 354.550843
- 354.550844 - 756.259005
- 756.259006 - 1179.08112
- 1179.08113 - 1763.19272
- 1763.19273 - 2564.95214
- 2564.95215 - 3917.10718
- 3917.10719 - 5631.91475

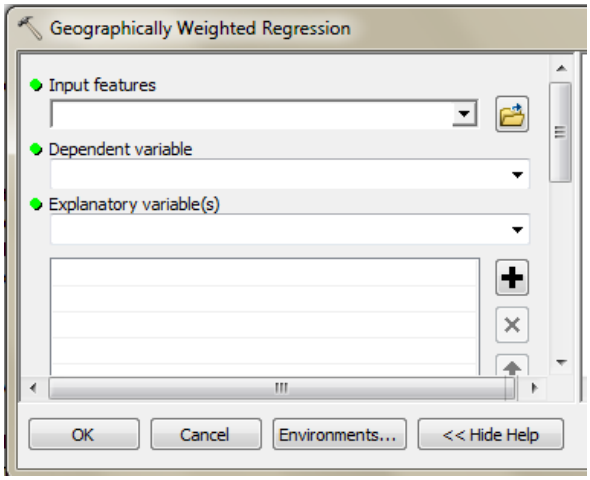
Land cover modeling

After having obtain all these data and having chosen the parameters that I wanted to use to model the percent imperviousness in each block group, I could finally run my model.

I first wanted to use the *Geographically Weighted Regression* tool from Spatial Statistics: mathematical model for 2011:

Figure 9: illustration taken from Arcmap help and *Geographically Weighted Regression* tool





I could then found a relation: ***b0 + b1*population per km^2+ b2* housing units per km^2 + b3*median annual income per housing = number of impervious cells***

But I did not manage to run this tool correctly since I always had a "Background processing error".

However, I used the solver after having exported the data, and the values that best fit for b0, b1, b2 and b3 are:

b0	1.999841
b1	0.000135
b2	0.000106
b3	0.000251

Final considerations

By comparing the values that are theoretically found using my model, and the real values, I concluded that my model was coherent but not precise at all: there is a margin of error with this model of about 40%, which is a lot.

To reduce this problem, I could try to use a more complex formula, or run this model for more counties. for example by adding the Bexar county that corresponds to the urbanized areas of San Antonio.