Identifying natural cities and analysing its inner structure

Chris de Rijke

Faculty of Engineering and Sustainable Development
University of Gävle, 801 76, Gävle, Sweden
E-mail: Chris.de.Rijke@hig.se

In this tutorial a step by step guide is provided which shows the process of analyzing urban morphology according to its inherent structure topologically. First natural cities will be explained, whereafter the process of creating them is shown. With the generated natural cities topological analysis is done following the method presented in the paper by Jiang (2018). Finally an indication of how the end result can be interpreted is given.

Introduction

Space syntax, Natural cities, Shaped by us, it shapes us.

Data gathering and preparation

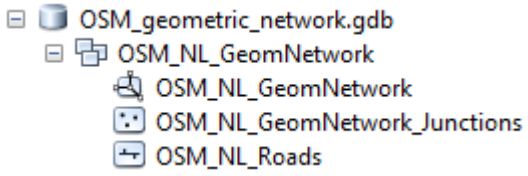
Generating natural cities with large amounts of data (e.g. an entire country) takes a lot of processing time. It can take up to 30 minutes per step/calculation. For this example the Netherlands is used. As basedata OpenStreetMaps roads are used as they are relatively accurate and freely available. At download.geofabrik.de entire OSM datasets can be downloaded. For this tutorial the shapefile from the Netherlands is necessary.

Generating natural cities

Natural cities will be created based on street junctions. A higher concentration of junctions means a higher concentration of streets and can be used as an indication for dense areas or cities. The first step therefore is extracting the junctions from the raw data which is just obtained. Within the dataset downloaded from geofabrik only the osm_*roads.shp* is needed. To extract the junctions from this shapefile it needs to be transformed to a geometric network.

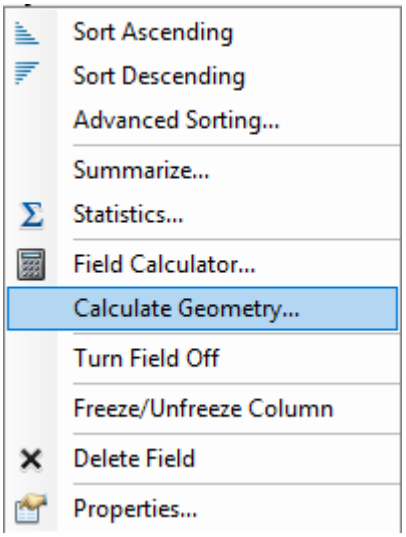
1. Download the OSM data from Geofabrik for the Netherlands (Europe). Open Arcgis and click *Add data -> gis_osm_roads_free_1.shp*
2. Reproject the data by using the *Project (ArcToolbox -> Data Management -> Projections and Transformations)* tool to *RD_New (ESPG: 28992)*.
3. Remove the original OSM shapefile and change the coordinate system of the data frame to *RD_New* as well. (*Data frame properties -> Coordinate system*).
4. Within the *catalog* create a *file geodatabase* within the working folder by *right-clicking* the folder and selecting *new -> file geodatabase*.
5. Within the newly generated *geodatabase* create a *feature dataset* by *right-clicking* the *geodatabase* and selecting *new -> feature dataset*.
6. Now add the *roads.shp* from step 3 to the feature dataset by *right-clicking* the *layer* in the *table of contents* and clicking *data -> export data*. Save it as a *feature layer* within the *feature dataset* created in the previous step.
7. Now you should have a *feature dataset* with a *feature layer*, which is the OSM roads shapefile. The next step is creating a *geometric network* from the OSM roads. Find the *Create geometric network* tool and input the *feature dataset* with the OSM roads as *feature class*.

8. After the creation of the *geometric network* junctions from the OSM roads are generated. These junctions form the basis for the generation of natural cities with OSM data. Open the *Junctions feature class* and export it to a shapefile. This is going to be the file we will be using going forward.



The next part consists of the generation of natural cities within a country. For this a TIN will be constructed which generates lines between the junctions and its nearest neighbors. By calculating the distance of those lines we are able to detect naturally dense areas as the length between junctions will be much shorter there.

1. Find the *Create TIN (3D Analyst)* tool and open it up. As coordinate system fill in the previously used *RD_new* and as input feature class insert the *junctions shapefile* from the previous part. Then generate the TIN.
2. Now as the TIN created has no actual height data and because we are only after the junction-junction connection we will transform the TIN into lines. Find the *TIN Edge (3D Analyst)* tool and input the TIN created in step 1 and create the *TIN_edge*.
3. To separate the naturally dense areas from the open areas the distance of each line needs to be calculated. For this open the *attribute table* and *add field* from the *table options*  menu. Name it *length* with *Long Iteger* as its *type*. After you have created the field *right-click* on the header of the new field and select *Calculate Geometry*. Select *length* in *meters* and click *OK*.



4. After the length of each line is calculted we want to select only the lines which are shorter than the average distance. For this we need to know the average distance, which can be found by *right-clicking* the header of *length* again and this time selecting *Statistics*. In the case of the Netherlands the (first) mean is about 180m. Remember this number as we will need it to select all lines with a distance shorter than this, these are presumed to be in denser areas and they will form the natural cities. Click the *select by attributes*  button and select everything where *Length < 180*. Export the new selection to a new shapefile by *right-clicking* the layer and going to *data -> export data*. Make sure you are exporting the selected data only.

   Depending on the dataset used and especially the overall absolute size of the dataset used it might be preferable to take the second mean. Calculate the average of the first selection and select everything below this average and continue on. This is because with larger countries the length of the lines between dense areas are so large it significantly affects the average which means that also lines in open areas will be selected. In already dense areas, like the Netherlands, this effect is not noticable and the first mean is sufficient.

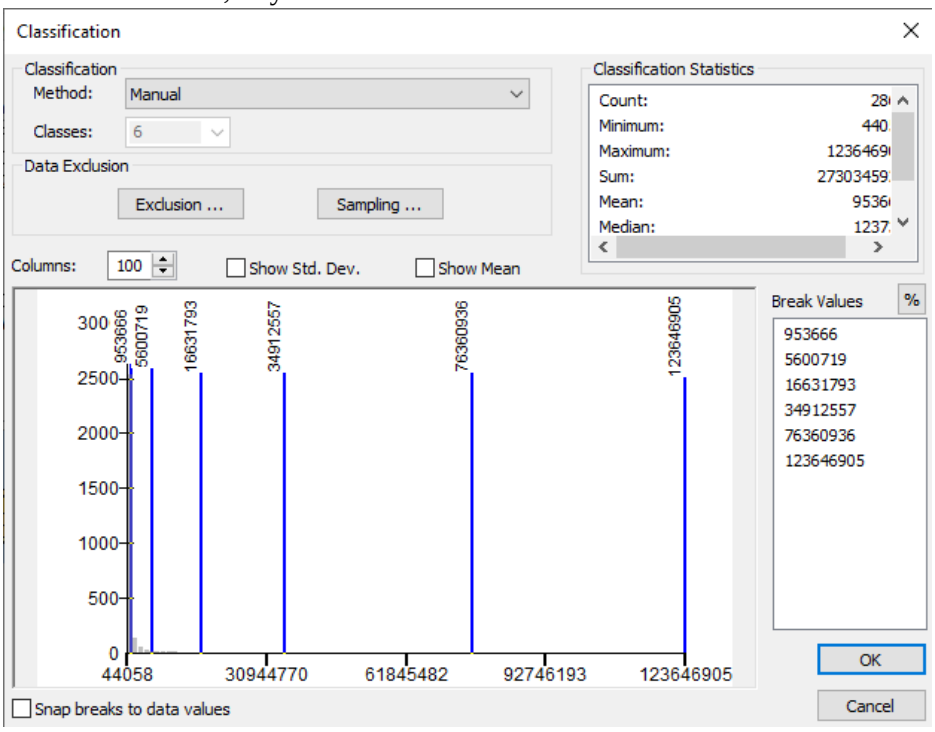
5. Now with the slimmed down dataset we can generate the natural cities. By using the *feature to polygon* tool and using the dataset created in step 4, polygons are created.
6. The final step in the creation of the natural cities is the removal of extra lines within the polygons. This is done with the *dissolve* tool. Open the tool and as input use the polygon shapefile created in the previous step. Make sure that *create multipart features* is ticked OFF.
7. Now to be able to work with the natural cities we will have to create some sort of value to differentiate with. This can be either area or wholeness for example. The calculation of area

follows the same principle as step 3. Create a new field for the area and *calculate geometry* and select *area*.

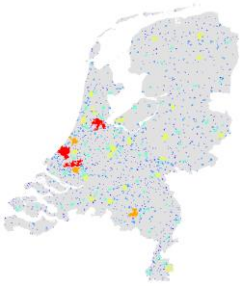
Visualizing natural cities

Now that the natural cities of the country or area have been created, a way to visualize them is necessary. Head/tail breaks is an excellent way of visualizing natural cities as there will be many more smaller natural cities than large ones. Head/tail breaks will be able to show this pattern correctly. See https://github.com/dingmartin/HTCalculator for a calculator which is able to calculate head/tail breaks automatically based on the set division rule. You can also calculate it yourself by following the instructions provided at https://en.wikipedia.org/wiki/Head/tail_Breaks.

1. The first step is reducing the amount of created natural cities. There are many very tiny polygons which have been created. These are obsolete for the visualization of one entire country as they decrease the clarity of the image. Therefore some have to be excluded. For this we can drop the first head/tail breaks class as this will select all the smallest natural cities. Similar to step 4 of the previous section we will look at the statistics of *area* of the natural cities. Open its *attribute table* and *right-click* the *area* column. Then select *statistics*. Remember the mean area which is given. Following this *select by attributes* and select everything with *area* > [the mean] (44052m^2 for the Netherlands). After the selection *export* the selected data only to a new shapefile.
2. To classify the newly created shapefile according to head/tail breaks there are several methods. You can use ArcGIS and manually calculate each class, you can export the *attribute table* and calculate it in excel, or you can use the calculator.



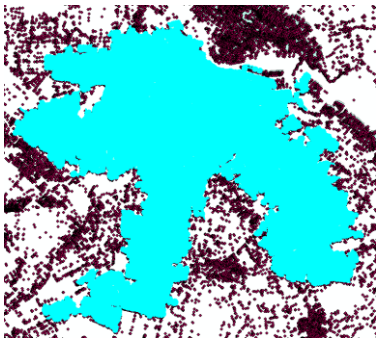
3. After you have applied head/tail breaks to the *area* values you can use underlying administrative data from the country (which is available online) to provide a background.



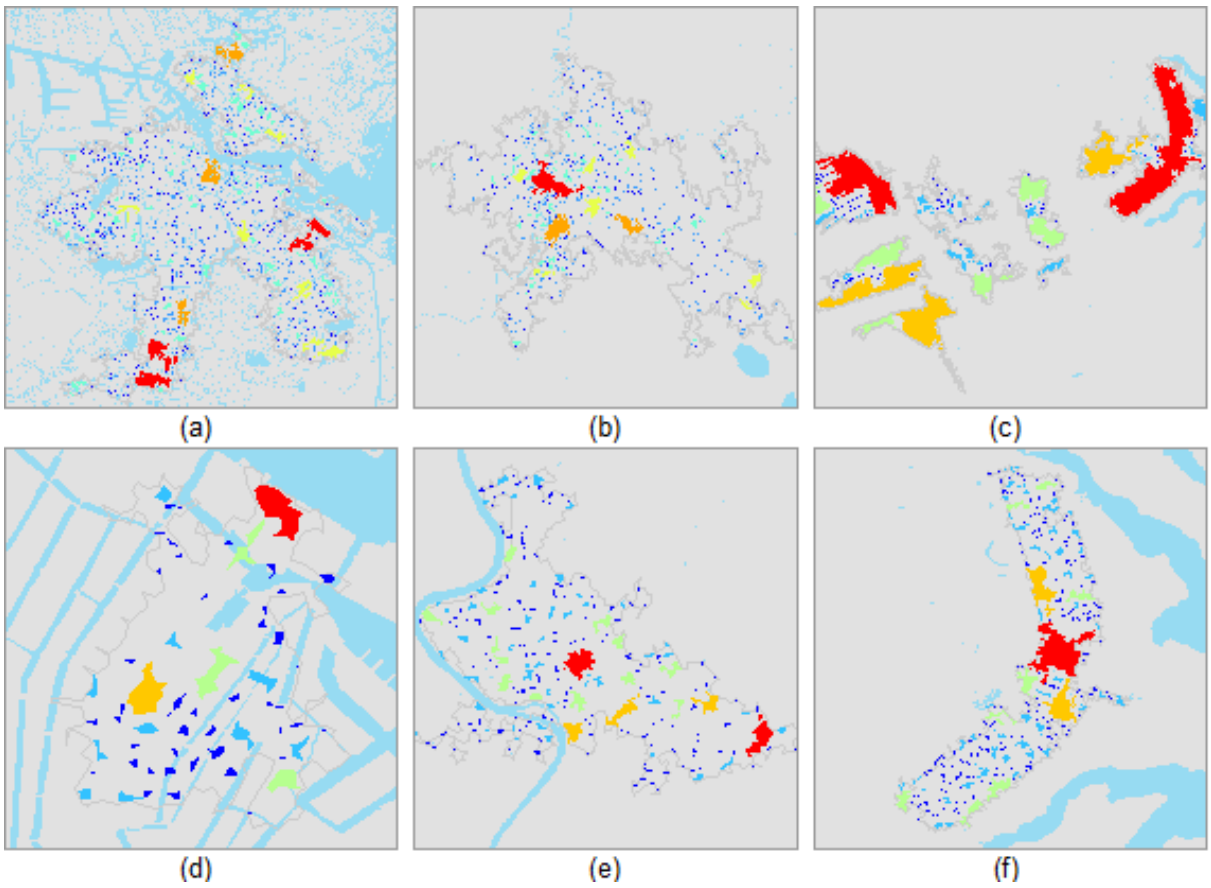
At this moment it becomes visible where the largest natural cities of a country are located, and by looking at the ht-index, which is 7 for the natural cities of the Netherlands, an impression of the inner structure of a whole country or area can be given.

Analysing inner city structure

As scaling law is universal and also recursive, this analysis can also be done on individual cities. A city can be seen as a small country with an inner structure of places with high concentration of streets and places with lower concentrations of streets. Therefore we can look into hotspots and even microspots of hotspots to see how a cities inner structure is organized. For this the process is very similar to the creation of natural cities, the difference is that instead of using OSM junctions of an entire country we will be using the OSM junctions of a natural city and subsequently of a hotspot within the natural city.



1. By using the OSM junctions obtained at the beginning of the tutorial and by using the natural cities obtained in the previous part we are able to select a city for further analysis. For this tutorial we will be using Amsterdam. The first step in the process after a city has been selected is to select and isolate only the junctions within the boundary of the natural city. You can either clip *OSM_Junctions* with the selected natural city or you can *selection -> select by location* and use the *intersect* method. Export only the junctions within the cities boundary to a new shapefile.
2. By using only the extracted junctions from within the boundary of a natural city the same analysis can be done again. Follow all the steps from the second part of *generating natural cities* again but now for a much smaller area. In this way the *hotspots* of the city are generated.
3. After the hotspots are generated the analysis can be performed again within a *hotspot* area with exactly the same procedure. Depending on which city is selected a choice can be made which hotspot is chosen to analyse further as the choice is less obvious than it was compared to an entire city.



Interpreting results

In the previous figure the true hotspots are shown classified according to its size. The hotspots in Rome (b) indicate a good structure, there is a clear center with supporting hotspots around them, and within the city center this pattern continues. For Amsterdam it is a bit more complicated. After the Second World War the city expanded very quickly to be able to deal with the increasing population and economic activities. The result was that lots of already neighboring cities, with or without a big historic background, were integrated into Amsterdam. Some of the largest developments taking place was the development of the southern part called Amstelveen. This part is very artificial and planned (a, south). This explains the biggest hotspots in the south and northeast. As these modernist parts are disjointed and disconnected from the living structure of Amsterdam, they are ignored during the recursive analysis and the city center instead is chosen as an indicator for living structure. In section 5.6 this is analysed further in a historical development analysis. Finally Brasilia is already disconnected as a whole, which results in very large unnatural hotspots throughout the city. This violates prerequisites for living structure as there is no scaling or support between classes.

References

Jiang B. (2018), A topological representation for taking cities as a coherent whole. *Geographical Analysis*, 50(3), 298-313.

Page L. and Brin S. (1998), The anatomy of a large-scale hypertextual Web search engine. *In: Proceedings of the seventh international conference on World Wide Web*. Elsevier: Amsterdam, 107–117.