Tutorial for Generating Natural Streets and Analyzing their Scaling Structure

Zifang Guo

Faculty of Engineering and Sustainable Development, Division of GIScience
University of Gävle, 801 76 Gävle, Sweden
E-mail: guozifang@msn.cn

I. Introduction

The formation of natural streets is self-organized, naturally from individual road segments, based on the Gestalt principle of good continuity. Each segment joins with its neighbor according to some principles, such as every-best-fit, self-best-fit, and self-fit (Jiang, Zhao, and Yin 2008). The difference between these three principles is whether each segment considers every segment or only itself when choosing one (best) fit neighbor. For joining each segment with its most suitable neighbor segment, it is necessary to repeat the process of choosing the smallest deflection angle until the angle becomes greater than a given value. The result represents streets that are connected to each other and form a network topology. Natural streets differ from named streets that are identified by the unique names. The former (natural streets) has no problems with losses or errors of names, while named streets are difficult to implement due to the incompleteness of names. The concept of natural streets is widely used, such as computing the fewest-turn map directions or routes (Jiang and Liu 2011). Generating natural streets is the first and most important part in this tutorial.

The previous study by Jiang (2009) has illustrated the fact that urban streets demonstrate the scaling property, which refers to far more less-connected than well-connected. This tutorial uses two methods to analyze scaling structure of natural streets. One method is power law distribution which is a type of heavy-tail distribution that can present the scaling pattern. The other is head/tail breaks (Jiang 2013) which is a classification scheme that divides data with heavy-tail distribution into the head (which has very few of large values) and the tail (which has a lot of small values). Data are classified around their arithmetic mean or average, recursively divided into head (above the mean) and tail (below the mean) by the use of head/tail breaks. The rest of this tutorial has four sections. Section 2 provides the download information of software and data resources. Section 3 is a detailed guidance for generating natural streets. Head/tail breaks and the power law distribution are applied in section 4 and 5.

II. Data resource and software support

This tutorial takes a French city Avignon as an example; the associated data of OpenStreetMap can be downloaded from http://download.geofabrik.de/. In this website, select *Europe, France, Provence* under the *Sub-Region* and save the shapefile data.

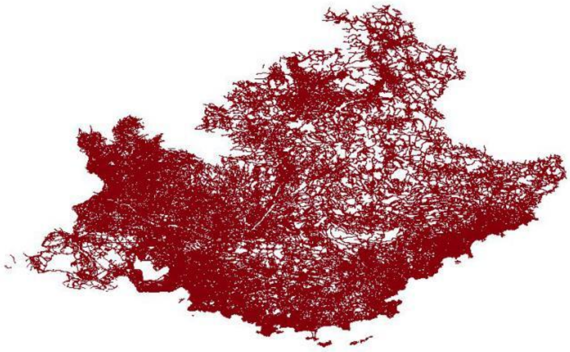
Also crucial to this tutorial is Axwoman software which is used to generate the natural streets, and essentially transforms the basic geometric representation into a topological representation which forms the basis of the analysis performed in the tutorial (Jiang, 2019).

ArcMap 10.2.0, 10.3.1, or 10.4.0 with the Data Interoperability Extension (for users without this extension, an alternative step is available) (http://desktop.arcgis.com/en/arcmap/)
Axwoman 6.3 (http://giscience.hig.se/binjiang/axwoman/)
Head/tail breaks calculator (https://github.com/ChrisdeRijke/HeadTailBreaksCalculator)
Matlab (https://se.mathworks.com/downloads/web_downloads)

- Matlab Powerlaw scripts (http://tuvalu.santafe.edu/~aaronc/powerlaws/)

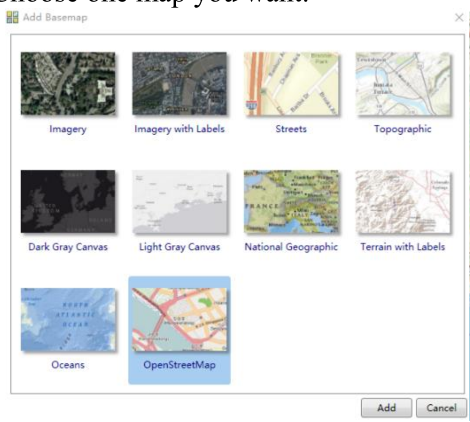
III. Generating natural streets

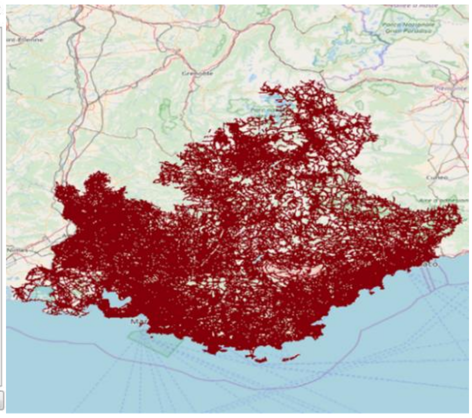
1. Start by opening ArcMap with a *Blank Map*, and import the data you saved before. Click the *Add Data* icon *Or* Click the *ArcCatalog* icon and connect to your data folder by clicking on the *Connect to Folder* icon , then add the *Roads* shapefile.



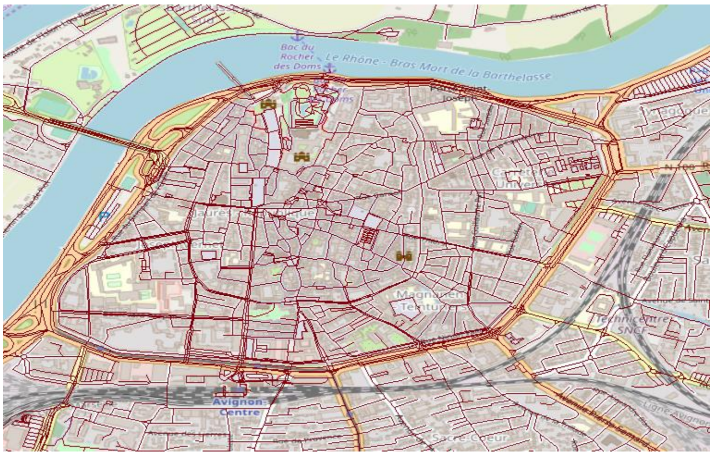
Note: In this tutorial, we selected Avignon as an example, and so it needs to be extracted from the data.

2. Avignon should be located from the whole of Provence. Adding a base map when looking for an area is helpful because it can display the boundaries, names and other information. (If you do not need to use a base map, go to step 3 directly.)
   a. Click *Add Data > Add Basemap*.
   b. Choose one map you want.

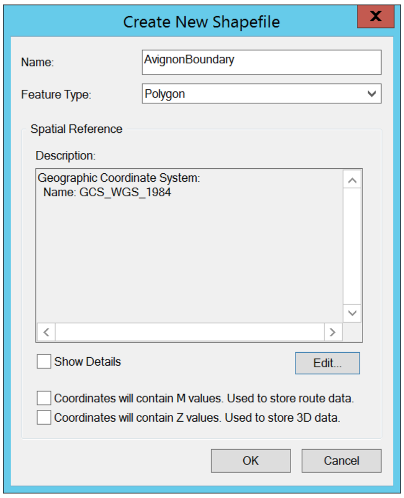




   c. Locate Avignon by using the pan and zoom tools to enlarge and center the city of Avignon on the screen

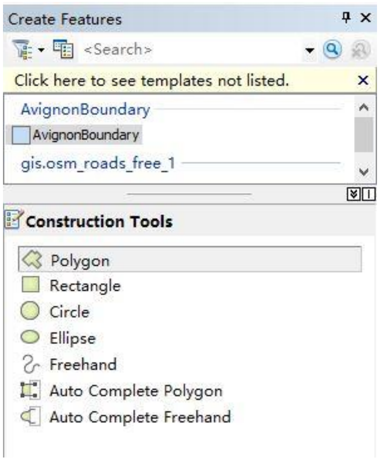


3. Create a boundary for Avignon so that it can be clipped.
   a. Click the *ArcCatalog* icon, *Connect to Folder* you are working with.
   b. Right click the folder, select *New > Shapefile...*
   c. In the Create New Shapefile Dialog, type *AvignonBoundary* as the *Name*. For the *Feature Type*, choose *Polygon*.
   d. Click *Edit* and select *Geographic Coordinate Systems, World,* and *WGS 1984.* Click *OK*.

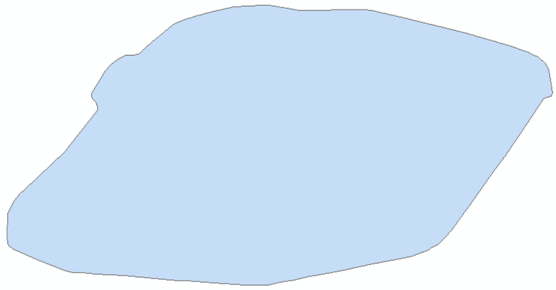


4. Edit the layer you created in the previous step.
   a. Go to *Editor > Start editing* (If necessary, click *Customize > Toolbars > Editor* to active the tool), choose AvignonBoundary, click *OK*. (Click *Continue* if you meet the warning window.)

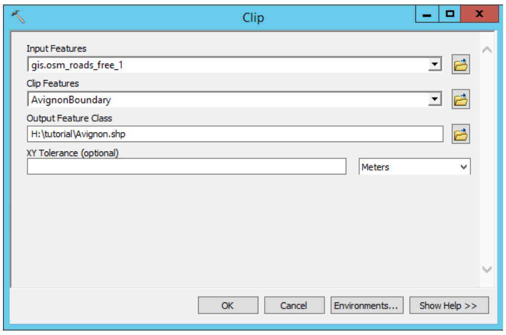
   b. From the *Create Features* window (If the window isn't displayed, click the button ), click *AvignonBoundary* and choose *Polygon* in the *Construction Tools*.

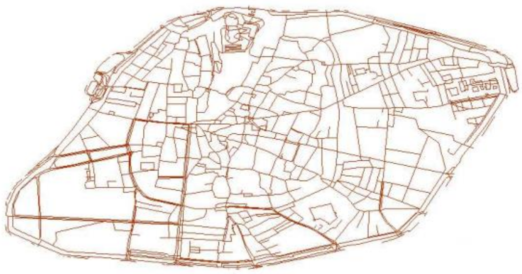


   c. Draw the boundary along the edge of the city in the layer. Avignon is surrounded by walls, which can be used as a reference for the boundary.
   d. When finished, click *Editor > Save Edits*, then click *Editor > Stop editing*.

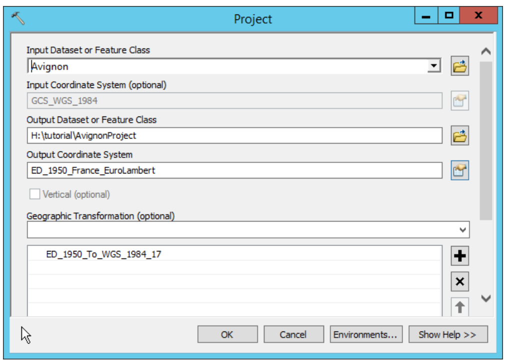


5. Clip Avignon from the whole map.
   a. Use the tool by clicking on *Geoprocessing* > *Clip.*
   b. In the Input Features, select the Roads shapefile.
   c. In the Clip Features, select *AvignonBoundary*.
   d. Save the file as a shapefile in your working directory by clicking the  button next to Output Feature Class. Navigate to your working directory and save it with an appropriate name (e.g. *Avignon.shp*) and make sure the file is saved outside of a geodatabase. (This is important as feature classes will not work, only shapefiles will).





6. This step is to make the map projection. It is necessary to choose a more suitable projected coordinate system for the layer.
   a. In *ArcToolbox* , double click *DataManagement Tools* > *Projections and Transformations* > *Project*.
   b. Input *Avignon* and name Output file as *AvignonProject.shp*. (similar to step 5d, save it outside of a geodatabase as a shapefile)
   c. For the Output Coordinate System, choose *Projected Coordinate Systems, National Grids, France, ED 1950 France EuroLambert* and click *OK*.
   d. The shapefile generated here should be added to a new dataframe, as all other data is not necessary anymore and has conflicting coordinate systems. Go to *Insert* > *Data Frame.* Check if the data frame has the same Coordinate System; right click the *data frame* > *properties* > *Coordinate System.* (Set it to *ED 1950 France EuroLambert* if it is not set to that already)

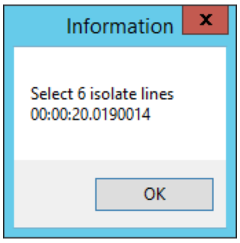


7. In the process of generating natural streets, it is important that all streets have a connection to each other (i.e. it must be able to reach all points regardless of starting position). Therefore edit disconnected streets so that they are attached to other neighboring streets or delete them directly if they consist of a very small proportion.

   a. Click *Customize* > *Toolbars* > *Axwoman 6.3*. (If necessary, *Customize* > *extensions* > tick *Axwoman 6.3*）

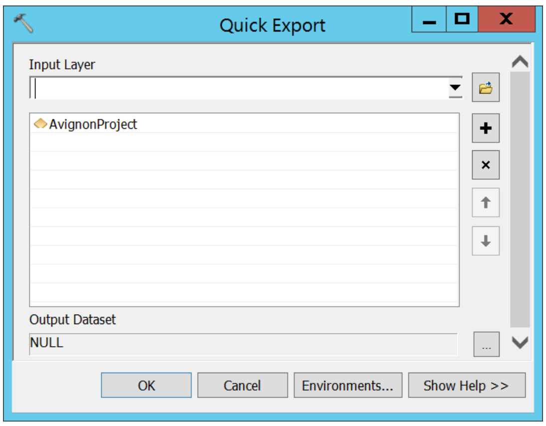
   b. Highlight the *AvignonProject* layer and click *Select Features* button .
   c. Select one line in the present layer and click the button *Get isolate lines*  in Axwoman toolbar.

Note: After waiting for the process, we can see the number of isolated arc segments and they are marked. If there are problems running this function, it may be a solution to restart ArcGIS completely and then load in only the AvignonProject.shp shapefile

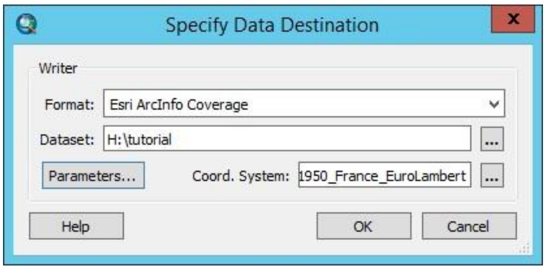


   d. Go to *Editor* > *Start Editing*, Click *Edit Vertices*  to edit isolated streets in order to connect with other segments. Or press *Delete* on the keyboard if the number is small. *Save Edits* and *Stop Editing* before continuing.
   e. Repeat step c. and d. to check for isolated lines again until there are no isolated lines left. (This is important as natural streets cannot be generated if there are no disconnected streets)

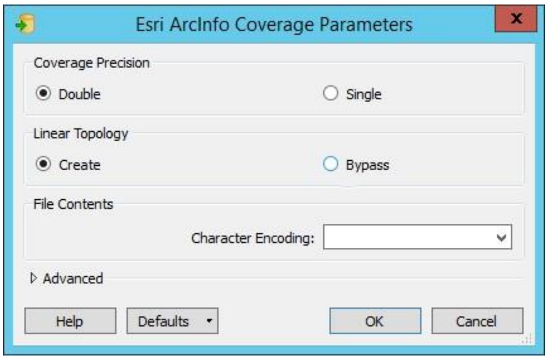
8. This step is to chop streets at each junction into segments. It is an important step to choose the most suitable neighboring street segments to form a good continuity. (If ArcGIS Data Interoperability Tools is not installed see the Appendix for an alternative step 8).
   a. Go to *ArcToolbox* > *Data Interoperability Tools* > *Quick Export.*
   b. Input *AvignonProject* layer.



c. Click  button to activate more options for the output dataset. Choose *ESRI ArcInfo Coverage* for Format and choose a *folder for the ARCINFO Coverage*. Select the same coordinate system as the layer.

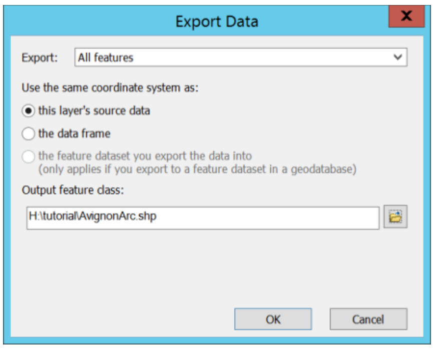


d. Click the *Parameters* button, tick *Double* for Coverage Precision and *Create* for Linear Topology.

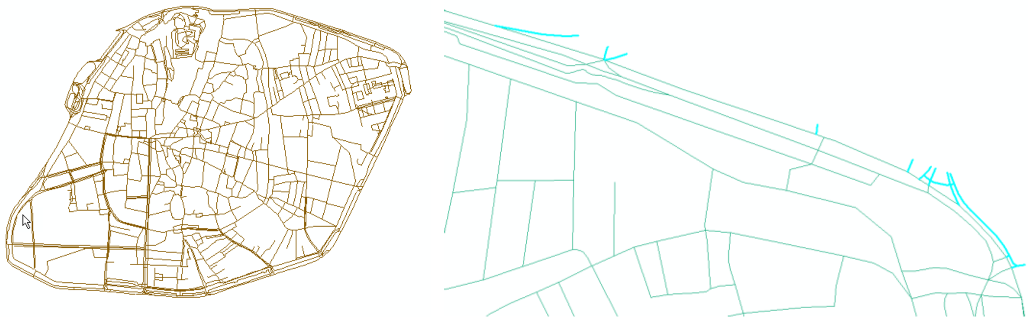


e. Click *OK* in each window to finish Quick Export.

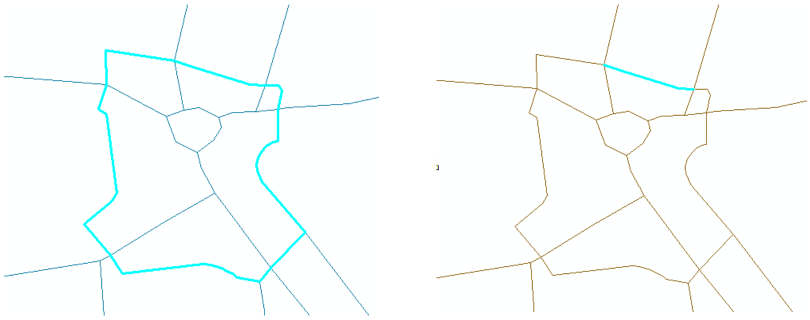
f. When it shows Quick Export Completed in the results window, add this layer which is automatically named as *avginonprojec arc*. Right click the file and choose *Data* > *Export Data*. Select the storing folder and named as what you want and click *OK*. Add the exported data to the map as a layer.



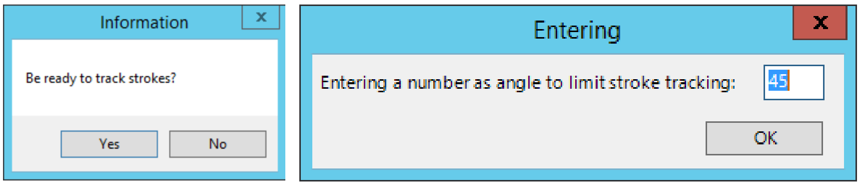
g. There may be some unimportant arcs outside the area that can be deleted directly. (Look for small segments connected to the outside boundary, see the next figure) Go to *Editor* > *Start editing*, highlight the layer and click S*elect Features* icon to select the arcs that should be deleted. Press *Delete* on the keyboard. *Save edits* and *Stop Editing.*



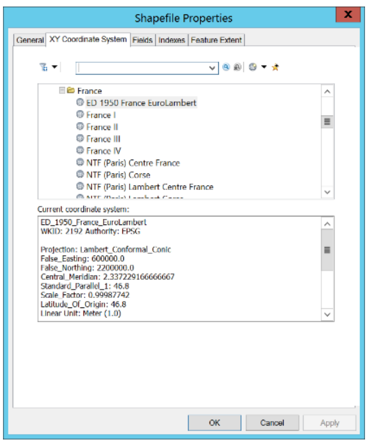
Note: When you finish step 8, zoom in the layer and check the arcs whether be chopped like the picture below. (The left figure is original, the right one is chopped.)



9. You can use Axwoman to generate natural streets automatically.
   a. Check isolated arcs again by repeating step 7 to ensure that there are no isolated segments.
   b. Highlight the layer and click *Tracking strokes at junction by a limited angle*  in the Axwoman toolbar. Choose *Yes* and *OK*, as you can keep the default angle. *Save* the natural streets to a shapefile and name it as *AvignonNR.shp*.



   c. Now you can see that the natural streets layer has been added. Check its coordinate system and change it if it is not the same as the layer. In Catalog window, right click *AvignonNR* shepefile > *Properties*. In Shapefile Properties window, click *XY Coordinate System* tab, then choose the coordinate system and click *Add*. Then click *OK*.



d. Examine the newly generated natural streets by using the *select feature* function and compare them to the previously examined segments in step 8. You should be able to see how the natural streets are generated from the segmented file previously.

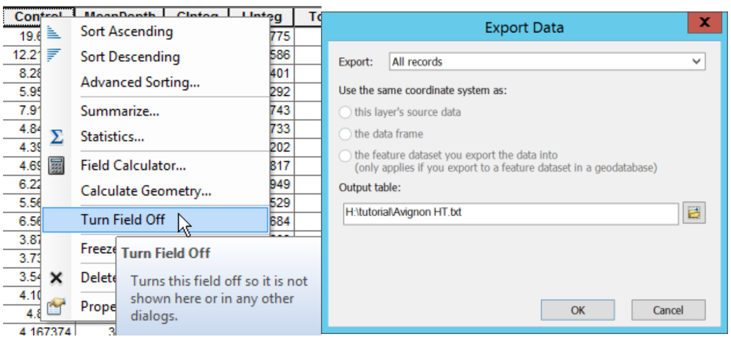
Note: The generation of natural streets for a small city like Avignon will probably take around 30-60 seconds. The generation of larger cities will take longer up to 5 to 10 minutes. It is possible that a lack of proper licenses will slow the process dramatically. Also clicking the *start editing* button twice may speed up the process if this is done before *tracking strokes*.

IV. Conducting head/tail breaks

1. You may need to calculate the spatial syntax parameters of the street network for more research. This step is also achieved through the function of Axwoman.
   a. Highlight the layer *AvignonNR* and click *Calculate parameters in case of lines with lines*  and click *OK*.
   b. This process takes a few minutes and the results are like the figure below.

The resulting image of Avignon shows roads based on connectivity. Connectivity means the number of roads which have a shared intersection. A street has a direct connection when another street is directly attached. Warmer colors (red) show roads which are very well connected while colder colors (blue) show less connected roads. The image should portray much more colder colors, which reflects that there are much more smaller things than large ones in many spatial situations. For example the streets shown in red have high connectivity values, with 54 to 71 streets intersecting them. In contrast, the streets in blue have the lowest connectivity, where only 1 to 4 streets are intersecting them.

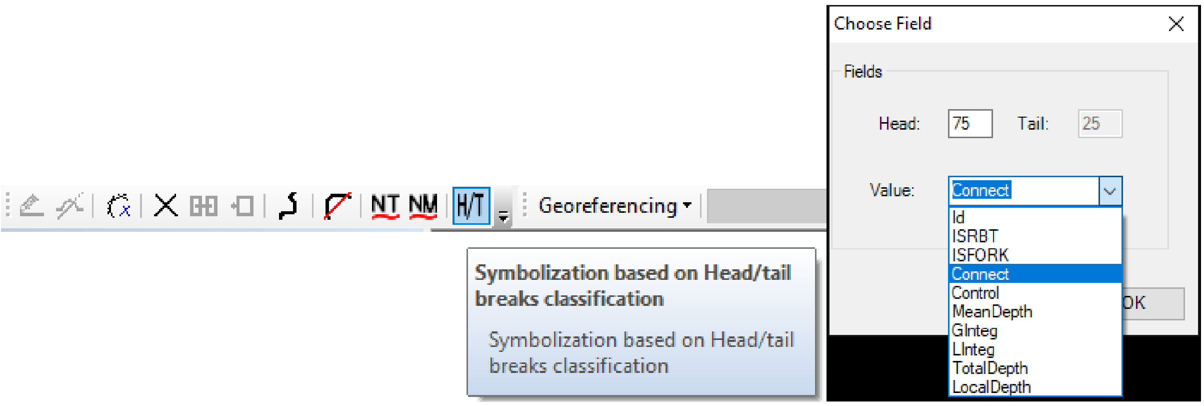
2. This is step intended to apply the head/tail breaks method to classify natural streets by using head/tail breaks calculator.
   a. In ArcMap, right click the *AvignonNR* layer and open attribute table.
   b. In the table windows, click *Table Options*  and *Export* the *Connect* column as a text file. One way to export only the Connect column is to turn off the other fields. (To turn a field off, right-click on the field name and select *Turn Field Off*.)



c. Open the head/tail breaks calculator (ExcelHTcalculatorOneColumn.xlsx) and *copy* the data from the text file to the A column in the “Results” sheet. The result will show up to the left. There are two results, one based on head/tail breaks 1.0 (Jiang, 2013), and one based on head/tail breaks 2.0 (Jiang, 2019). Note that axwoman uses head/tail breaks 1.0 by default, you can use the calculator to see if the classification in ArcGIS is done correctly. Do note that your results may differ from the screenshot depending on which roads/segments were removed in step 8g from the previous section. Also note that in your case there may be a difference in the results between head/tail breaks 1.0 and 2.0, in that case use the classification shown in head/tail breaks 2.0.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Connect	<--- Insert here		Results													
2	2			Head/tail breaks 1.0							Head/tail breaks 2.0						
3	5			# rows	Mean	# head	# tail	% head	% tail		# rows	Mean	# head	# tail	% head	% tail	avg head
4	11			341	4.111437	71	270	21%	79%		341	4.111437	71	270	21%	79%	21%
5	42			71	10.23944	26	45	37%	63%		71	10.23944	26	45	37%	63%	29%
6	1			26	17.03846	9	17	35%	65%		26	17.03846	9	17	35%	65%	31%
7	3			9	24.11111	3	6	33%	67%		9	24	3	6	33%	67%	31%
8	17			3	33	1	2	33%	67%		3	33	1	2	33%	67%	32%
9	3																
10	3																
11	2																
12	17																

Note: In Axwoman you can also classify using the H/T button (“*Symbolization based on Head/tail breaks classification*”) to automatically generate a classification based on Head/tail breaks. For head/tail breaks 1.0 you can leave the default settings for the head value, for the *value* field select *Connect*. For head/tail breaks 2.0 you can increase the head value, but you will have to use the head/tail breaks calculator sheet or manually calculate if the overall head% does not exceed 40%. It may happen that the head value resets to 50%, therefore manually type in the box the higher value (e.g. 90%) and immediately click OK.



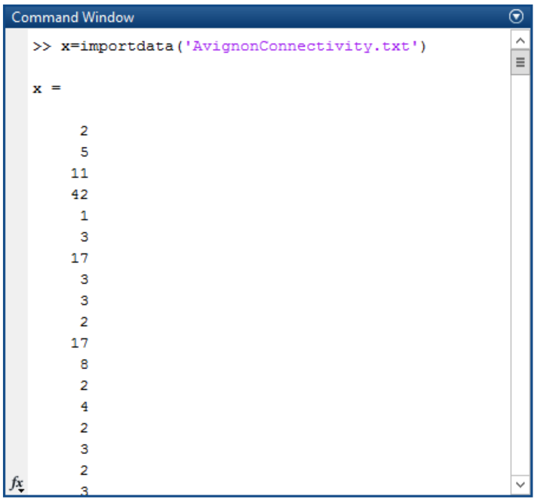
V. Examining power law distribution

1. MATLAB software is used to examine if the degree distribution of the street network fits a power law distribution. For this the connectivity calculated in the previous section can be used. If you have already exported the Connectivity column for the head/tail breaks calculation you can use this again, otherwise follow step 2b from the previous section to export the column.
2. Start MATLAB and click *Browse for folder* to connect the folder you are working with.

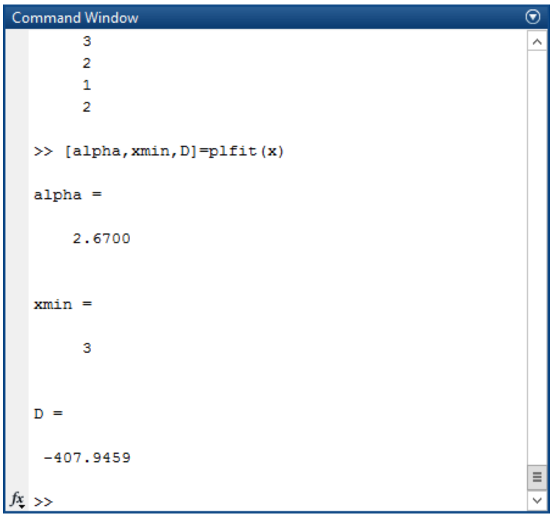
Note: Make sure the degree vector file and Clauset MATLAB code files (Clauset, 2009) plfit.m, plplot.m and plpva.m are in this folder. (Go to http://tuvalu.santafe.edu/~aaronc/powerlaws/ to download these three MATLAB code files.)

a. Use Notepad to open the text file and delete the first row and save it as a text file. So that there are only connectivity numbers.

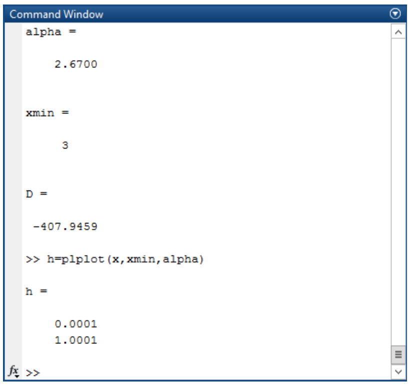
b. In the command window, type the code: *x=importdata('AvignonConnectivity.txt')* and press *Enter* on the keyboard to import the degree data of Avignon street network.

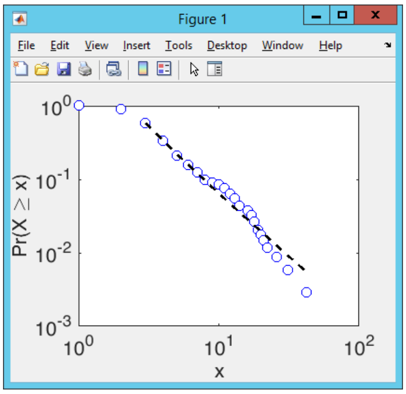


c. Type: *[alpha,xmin,D]=plfit(x)* and press *Enter.* This step is intended to get the alpha and minimum value using a power law fit function.

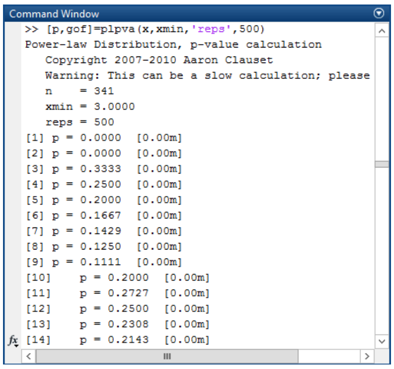


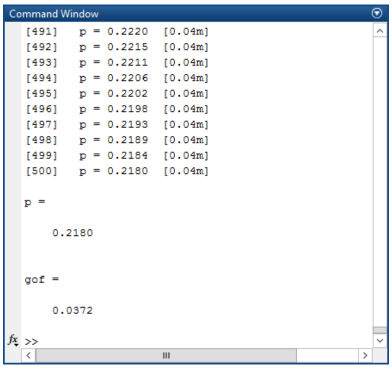
d. Next, use power law plot function, type *h=plplot(x,xmin,alpha)* and press *Enter.* A figure will display that shows a power law distributional model of Avignon street. The value along the x axis are degree values from the natural streets degree file (connectivity). The values along the y axis show the cumulative probability of obtaining those values. You can see specific X and Y values by hovering your mouse cursor over a blue dot.





e. Calculate a p-value by using power law pva function. Type: *[p,gof]=plpva(x,xmin,'reps',500)* and press *Enter*.





Note: If the p-value is greater than 0.05, the data fits a power law distribution. For Avignon, the p-value is equal to 0.2180, which is greater than 0.05. (In contrast to traditional hypothesis testing in statistics, here we want high p values.) Note: Depending on how Avignon was prepared in pre-processing results may be different, however if the selected area is representative of Avignon it should represent a p-value greater than 0.05.

After completing above exercises, you have learned how to generate natural streets, conduct head/tail breaks and examine power law distribution for data. When using ArcMap to generate natural streets, it is recommended that you check the coordinate system at every step, because this may cause some errors. In addition, installing and activating the software and tools correctly is an important step in getting the results.

Acknowledgement:
I would like to thank Prof. Bin Jiang for his guidance and help in preparing this tutorial. This tutorial was updated by Chris de Rijke following discussions with Prof. Terry Slocum. Thanks to Prof. Slocum's detailed comments the quality of this tutorial and its guidance through the steps could be improved benefitting all future readers.

References:

Clauset A., Shalizi C. R., and Newman M. E. J. (2009), Power law distributions in empirical data, *SLAM Review*, 51, 661–703.

Jiang B., Zhao S., and Yin J. (2008), Self-organized natural roads for predicting traffic flow: a sensitivity study, *Journal of Statistical Mechanics: Theory and Experiment*, July, P07008.

Jiang B. and Liu X. (2011), Computing the fewest-turn map directions based on the connectivity of natural roads, *International Journal of Geographical Information Science*, *25*(7), 1069–1082.

Jiang B. (2009), Street hierarchies: a minority of streets account for a majority of traffic flow, *International Journal of Geographical Information Science*, *23*(8), 1033–1048.

Jiang B. (2013), Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution, *The Professional Geographer*, 65(3), 482–494.
Jiang B. (2019), *Axwoman in a Nutshell*, https://www.researchgate.net/publication/337656401_Axwoman_in_a_Nutshell
Jiang B. (2019), A recursive definition of goodness of space for bridging the concepts of space and place for sustainability. *Sustainability*, *11*(15), 4091.
Miao Y. (2012), *A Tutorial for Transforming a Line Shapefile to Pajek's Netfile, and then back to the shapefile*, University of Gävle, Sweden.
Sun X. (2012a), *Tutorial for Automatically Generating Axial Lines*, University of Gävle, Sweden.
Sun X. (2012b), *Small-world and Scale-free Properties of Urban Street Networks: A Tutorial*, University of Gävle, Sweden.
Wu S. (2014), *Tutorial for Applying Head/tail Breaks in ArcGIS*, University of Gävle, Sweden.

Appendix

Generating Natural Streets Alternative Step 8 (Advanced License)

1. To separate streets at every junction the *Feature to Line* tool can be used.
   a. Go to *ArcToolbox > Data Management Tools > Features > Feature to Line.*
   b. Choose *AvignonProject* as Input Layer.
   c. Choose an Output Feature Class by clicking the  button and select a location outside of a geodatabase and name the file appropriately (e.g. *AvignonSegment.shp*).
   d. Leave *XY Tolerance* blank and untick *Preserve attributes* and run the tool.
   e. Continue at step 8g