

A TOOLKIT FOR MEASURING LOCALIZED COLOCATION QUOTIENT (LCLQ)

First Released on Jan 25, 2016

Updated on May 31, 2016 (sample data set is updated)

Two toolkits (Localized_CLQ.exe and Global_CLQ.exe) are developed to automate the computations of Localized Colocation Quotient (LCLQ) and Global Colocation Quotient (GCLQ), respectively. The outputs also include results on corresponding statistical significance tests. Please cite:

- Wang, F., Hu, Y., Wang, S., and Li, X. 2016. Local Indicator of Colocation Quotient with a Statistical Significance Test: Examining Spatial Association of Crime and Facilities. *The Professional Geographer*, DOI: 10.1080/00330124.2016.1157498.

1. Preparing the Data:

Before running the tools, the user needs to prepare an ASCII text file for the input point data (*.txt) including points of all categories of interest. For example, the sample data set provided (MotorTheft_allfacilities.txt) is for exploring the colocation behavior between motorcycle theft crime and land use facilities that include school, retail shop, and entertainment establishment. The text file consists of all points of these four categories. The text file should have these fields delimited by tab: unique id in the whole list (e.g., OBJECTID), *x* coordinate (e.g., X), *y* coordinate (e.g., Y), unique id in each category (e.g., ID), and category name (e.g., Category). See the example data in Figure 1 for more details.

OBJECTID	X	Y	ID	Category
1	1389415.334670000000	3478269.064910000000	1	School
2	1388808.653190000000	3475445.899300000000	2	School
3	1389882.116200000000	3475867.629550000000	3	School
4	1389756.083340000000	3479327.149210000000	1	Retail_shop
5	1389077.563390000000	3477863.620740000000	2	Retail_shop
6	1389826.415010000000	3474852.721700000000	3	Retail_shop
7	1388817.259190000000	3475215.497040000000	4	Retail_shop
8	1389798.293190000000	3474807.178580000000	1	Entertainment
9	1389781.421030000000	3474072.820380000000	2	Entertainment
10	1390143.539420000000	3474456.859880000000	3	Entertainment
11	1389120.733500000000	3479525.403600000000	1	Motorcycle_theft
12	1388444.567600000000	3471700.553900000000	2	Motorcycle_theft
13	1389573.528400000000	3473088.454900000000	3	Motorcycle_theft
14	1373416.099900000000	3491929.135600000000	4	Motorcycle_theft

Figure 1. An example of input data

2. Executing the Programs “Localized_CLQ” or “Global_CLQ”

To run the Localized Colocation Quotient tool, simply double-click the executable program Localized_CLQ.exe. Some users may be asked to download and install the “.Net” framework depending on its availability in your system. Similarly, double-click the executable program Global_CLQ.exe to run the Global Colocation Quotient tool.

The interface looks the same for the two tools as in Figure 2.

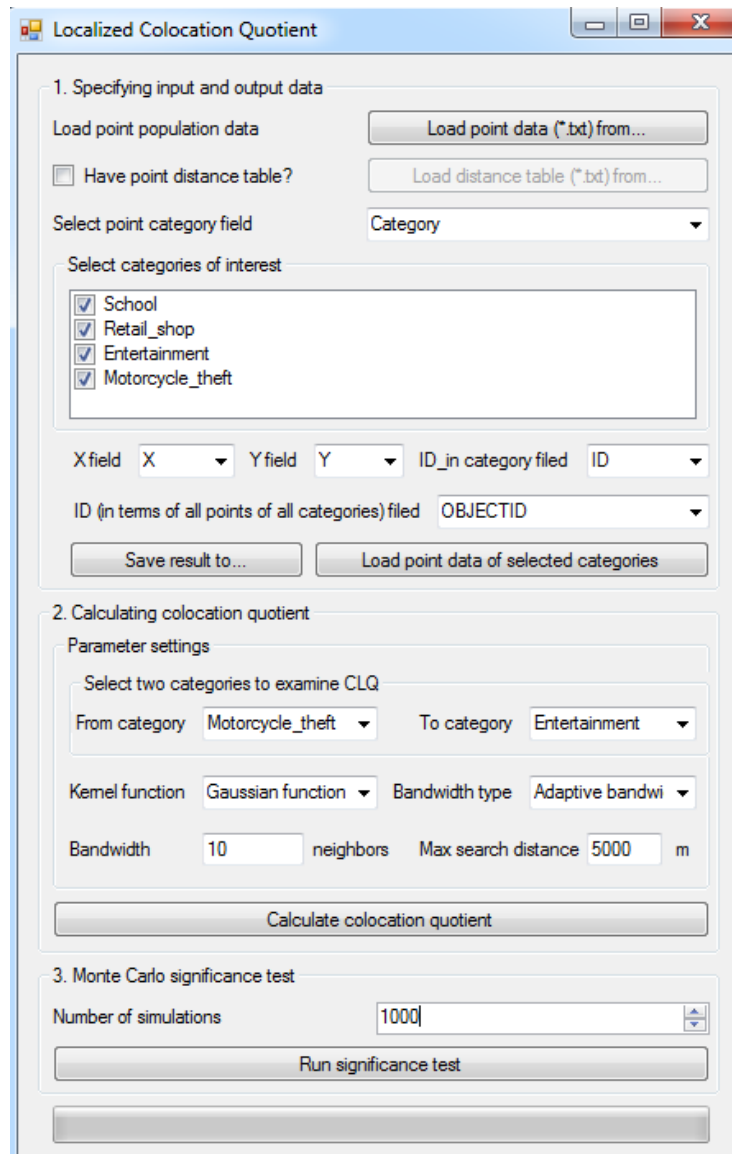


Figure 2. Interface of the Localized_CLQ (or Global_CLQ) tool

3. Defining Components of the Interface:

The interface consists of three major components:

(1) *Specifying input and output data*: Click the button “Load point data (*.txt) from...” to load the point data for measuring LCLQ (or GCLQ) from disk. By default, this program measures LCLQ (or GCLQ) based on a built-in Euclidean distance calculator. Some users, however, may want to use other distance metrics such as the street network distance, which is already provided by other programs or software such as ESRI ArcGIS. With a distance

table in place, simply check “Have point distance table?” and then click the button “Load distance table (*.txt) from...” (the button is disabled when “Have point distance table?” is unchecked). See Figure 3 for a snapshot of an example distance table.

```
source_id,destination_id,distance
1,1,4346.96620457
1,2,4560.68605026
1,3,4772.16548437
1,4,4002.11882843
1,5,3910.48545044
1,6,4202.01851399
1,7,4160.43158302
```

Figure 3. Distance table (example)

Next, select point category field (e.g., Category) from the list, and all the category names (e.g., School, Retail_shop, Entertainment, and Motorcycle_theft) will be automatically listed in the area under “Select categories of interest” as shown in Figure 2. One may check any categories to examine the pairwise LCLQ (or GCLQ) based on your own preference. By default, all categories are included in the calculation. Similar to the category field, other four field names (e.g., X, Y, ID_in category, and ID) will be automatically populated as well. Then choose a destination path to save the result by clicking the button “Save result to...”. Given that GCLQ reports only one value for the entire study area, the tool Global_CLQ.exe reports the GCLQ estimate right after the computation without writing it into a file; therefore, this “Save result to...” function is not available in Global_CLQ.exe. After that, click the button “Load point data of selected categories” to load the data.

(2) *Calculating localized (or global) colocation quotient*: As illustrated in Figure 2, select two categories of interest to measure LCLQ (or GCLQ), e.g., Motorcycle_theft and Entertainment as From category and To category, respectively. Choose a desired kernel function (e.g., Gaussian function) and bandwidth type (e.g., Adaptive and Fixed bandwidth). After bandwidth type, say adaptive, is selected, specify a bandwidth such as 10 nearest neighbors. The Max search distance is designed to save computation by only searching neighbors that locate within a predefined distance (e.g., 5000 m), and it should be set according to the bandwidth value. Note that the Max search distance should be

determined before loading data in step 1. Finally, click the button “Calculate colocation quotient” to measure the LCLQ (or GCLQ).

(3) *Testing the statistical significance of the obtained CLQ measures*: Specify the number of simulations and then click the button “Run significance test” to proceed.

4. Results:

For `Localized_CLQ.exe`, the result (i.e., a text file saved) includes the LCLQ scores and corresponding p-values for each record. Other fields such as the unique id in the whole list (i.e., `id_alldata`), unique id in each category (i.e., `id_incategory`), and category name (i.e., `category`) are also provided to facilitate subsequent analyses (e.g., thematic mapping in ArcGIS). See Figure 4 for more details.

```
id_alldata,id_incategory,category,LCLQ_value,p_value
810,1,Motorcycle_theft,0.822891973806735,0
811,2,Motorcycle_theft,0.00518164686683363,0
812,3,Motorcycle_theft,0.000350176992651069,0
813,4,Motorcycle_theft,2.78808611081278,0.22
814,5,Motorcycle_theft,0.00340805658821656,0
815,6,Motorcycle_theft,1.32014149233712E-06,0
816,7,Motorcycle_theft,0.239305351699218,0
817,8,Motorcycle_theft,1.13789291084213,0
818,9,Motorcycle_theft,4.46142291198935,0.92
```

Figure 4. Illustration of LCLQ results

For `Global_CLQ.exe`, the result includes one GCLQ score and corresponding p-value that are shown in a message box right after the computation. See Figure 5.

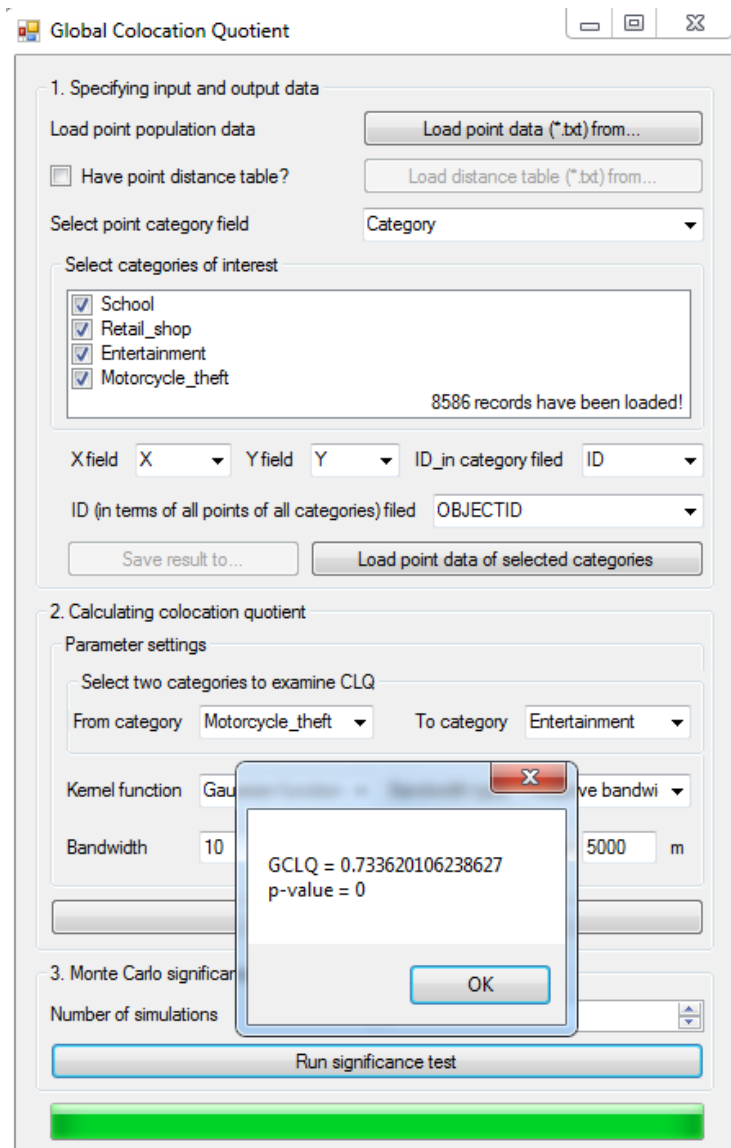


Figure 5. Illustration of GCLQ results