GeoMasker 1.0

Fundamentals and User Guide

Institution: Center for Geographic Information Science, Research Center for Humanities and Social Sciences, Academia Sinica, Taipei, Taiwan

Team Members:

Dr. Ta-Chien Chan (Principal Investigator)
Mr. Chien-Min Wang (GIS Programmer)
Ms. Yung-Mei Lee (GIS Analyst)
Ms. Li-Ting Feng (GIS Analyst)

Acknowledgement:

This application was supported by a grant from National Science Council, Taiwan, R.O.C. (NSC-102-2314-B-001-001).

Citation:

Ta-Chien Chan, Chien-Min Wang, Yung-Mei Lee, Li-Ting Feng: GeoMasker 1.0. In. Taipei, Taiwan: Academia Sinica 2013.

Table of contents

Table of contents ..... **I**

Table of figures ..... **II**

1. Introduction **1**

1.1 Why we use GeoMasker? 1
1.2 Overview of methods 2

2. Uses of GeoMasker **4**

3. Installation of GeoMasker **5**

4. Introduction of GeoMasker toolbox- non- barrier **8**

4.1 Non-barrier tool box – Affine method 8
4.2 Non-barrier tool box – Shifting method 10
4.3 Non-barrier tool box – Donut method 11
4.4 Neighbor method 13

5. Introduction of GeoMasker toolbox- barrier **16**

5.1 Barrier tool box – Affine method 16
5.2 Barrier tool box – Shifting method 18
5.3 Barrier tool box – Donut method 19
5.4 Fishnet method 20

6. Examine the results GeoMasker **22**

6.1 Report 22
6.2 FastReport 23

7. Training exercise **25**

7.1 Non-Barrier tool box- Affine method, shifting method and Donut method 25
7.2 Barrier tool box- Affine method, shifting method and Donut method 28
7.3 Non-Barrier tool box- Neighbor method 31
7.4 Fishnet method 33
7.5 Report 34
7.6 Fast Report 37

Table of figures

Figure 1 The principle of Affine method. 2
Figure 2 The principle of shifting method. 3
Figure 3 the principle of donut method. 3
Figure 4 Steps of opening AcrToolbox and "Add toolbox". 5
Figure 5 Pathway of GeoMask_10_1.tbx file. 6
Figure 6 Successful installation of GeoMasker in ArcGIS 10. 7
Figure 7 User interface of Affine method in no-barrier tool box. 8
Figure 8 Example of data points. 9
Figure 9 Example of boundary. [base polygon]. 10
Figure 10 User interface of Shifting method in no-barrier tool box. 10
Figure 11 User interface of Donut method in no-barrier tool box. 12
Figure 12 User interface of Neighbor method in barrier tool box. 13
Figure 13 Example of original point data. 14
Figure 14 Example of boundary map of Neighbor map. Data column is the density information of different areas. 15
Figure 15 User interface of Affine method in barrier tool box. 16
Figure 16 Example of barrier. 18
Figure 17 User interface of Shifting method in barrier tool box. 18
Figure 18 User interface of Donut method in barrier tool box. 19
Figure 19 User interface of Fishnet method. 20
Figure 20 user interface of Report 22
Figure 21 user interface of FastReport. 24
Figure 22 Attribute table of random_samples.shp. 25
Figure 23 Attribute table of kaocity.shp. 25
Figure 24 Result of Affine method of non-barrier tool box. 26
Figure 25 Result of Shifting method of non-barrier tool box. 27
Figure 26 Result of Donut method of non-barrier tool box. 27
Figure 27 Example of barrier -Parks_lakes_dissolved.shp and corresponding attribute table. 28
Figure 28 Result of Affine method of barrier tool box. 29
Figure 29 Result of Shifting method of barrier tool box. 30
Figure 30 Result of Donut method of barrier tool box. 30
Figure 31 Attribute table of KH_town.shp. 31
Figure 32 Example of setting parameters of Neighbor method. 32
Figure 33 Example of result of Neighbor. 32
Figure 34 Example of setting parameters of Fishnet method. 33

Figure 35 Example of result of Fishnet method.....34
Figure 36 Attribute table of random_samples.shp. 35
Figure 37 Attribute table of kaocity.shp.....35
Figure 38 Attribute table of random_samples.shp. 36
Figure 39 Attribute table of random_samples.shp. 37
Figure 40 Attribute table of kaocity.shp.....37
Figure 41 Table of Contents report1.shp.....38
Figure 42 Attribute table of report1.shp.....39

1. Introduction

1.1 Why we use GeoMasker?

Since 1996, the United Stated Congress passed the important privacy protection law, Health Insurance Portability and Accountability Act (HIPAA) and the law would be taken effect in 2003 (U.S. Government Printing Office 1996). U.S. Department of Health and Human Services also declared the corresponding guidelines for protecting the usage of public health information in 2003 (Centers for Disease Control and Prevention 2003) which listed the types of public health information and the requirements before using the information. To respect the importance of health data privacy is an inevitable trend around the world (Lawlor and Stone 2001; Verschuuren et al. 2008). Personal privacy information including those can recognize a person such as name, ID, gender, medical and disease records and the locations are taken seriously in recent years. However, the human-related researches need the spatial and temporal information such as the location of the studied subjects or the patients for further epidemiological studies. Some of the personal information such as disease records can be protected by removing the ID and name, but the spatial information is difficult to remove because the geo-coordinates can help identify the diseases' clusters and the distribution of specific subjects in the real world. Removing the location information will cause the inference difficulties and unable to control the individual's risk factors. However, the retention of spatial information runs the risk to disclose the personal information because that the location information in studies mostly are the residents' locations. By cross-comparing different information, it is possible to know who the people are. Therefore, how to maintain the spatial characteristics such as distributions and clusters and avoid revealing the actual locations has become the most important issue for space-related studies. As a result, the GeoMasker has been developed to provide five geo-masking methods for the users to protect their spatial privacy.

1.2 Overview of methods

GeoMasker provides methods avoiding revealing the information of actual location by masking the data points under the condition that the shifted data points still maintain the spatial characteristics. Five methods of GeoMasker are described as following:

Affine method: User gives the specific length of radius (r) and the angle (θ). Shifting distance can be calculated according to the following equation and the principle of Affine method is presented in Figure 1:

$$(X', Y') = (X + \Delta X, Y + \Delta Y)$$

$$= (X + r \times \cos\theta\ , Y + r \times \sin\theta)$$

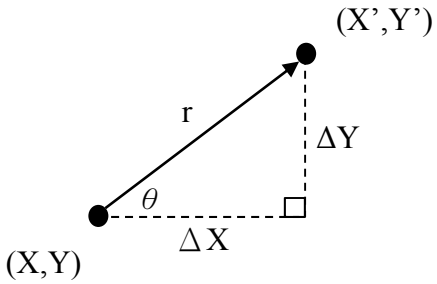


Figure 1 The principle of Affine method.

Shifting method: User gives the specific shifted magnitude of X-coordinate and Y-coordinate, (ΔX, ΔY). The new coordinate (X′,Y′) can be described as following equation:

$$(X', Y') = (X + \Delta X, Y + \Delta Y)$$

(X,Y) means the original location of data points. All the data point shifts the same magnitude of X-coordinate and Y-coordinate (ΔX, ΔY) in shifting method. Figure 2 presents the principle of shifting method.

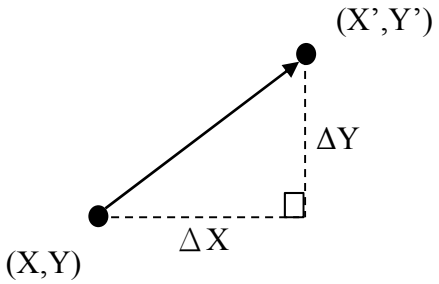


Figure 2 The principle of shifting method.

Donut method: Principle of donut method is similar to affine method and is shown in Figure 3. But the length of radius (r′) of donut method is randomly selected between the maximum and minimum given by user. And the angle (θ′) is randomly selected in the range of 0^o to 360^o. The new coordinate (X′,Y') can be described as following equation:

$$(X', Y') = (X + \Delta X, Y + \Delta Y) = (X + r' \times \cos\theta' , Y + r' \times \sin\theta')$$

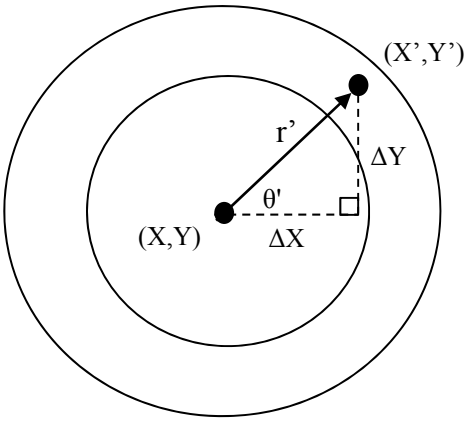


Figure 3 the principle of donut method.

Neighbors: The idea of Neighbors method is to find the closet and similar neighboring polygon. The spatial relationship among polygons was defined as the boundary adjoining. The definition of the similarity based on the smallest attribute's difference such as population or population density between the original polygon and the neighboring polygons. Then, each point was randomly projected on the corresponding selected polygon.

Aggregation: The aggregation method proposed here unlike traditional aggregation into administrative units such as census tracts, townships or cities. The users can create the user-defined fishnet like polygons and aggregated the cases into these polygons. The users will need to define the number of squared cells in x-axis of the layer's extent. Then, the algorithm will partition the x-axis width into the defined number of N_x cells. After the width of each cell is determined, the algorithm used it to partition the y-axis into N_y cells. After all, the total number of polygons is $N_x * N_y$. Finally, the spatial join function was used to aggregate the number of cases into the created small polygons.

2. Uses of GeoMasker

GeoMasker can be used in the human-related spatial data.

For data providers, in order not to reveal the actual location of cases, data providers can utilize GeoMasker to hide the information of cases but still maintain the spatial characteristics.

For researchers, it is important to hide the information about cases before publishing in the paper because that some studies may involve some sensitive information such as specific disease infections which might cause stigma or disturbance on the cases.

In order to protect the personal privacies, the spatial information can be shifted by GeoMasker before or after advanced spatial analysis. User can also examine the results shifted by GeoMasker reporting function and decide whether the method is suitable or not.

3. Installation of GeoMasker

GeoMasker can be installed in ESRI ArcGIS 10 as a tool in toolbox. The installation steps are as following:

Step1. Download the GeoMasker package to PC.

Step2. Run ArcGIS 10 and open ArcGIS toolbox. Right-click on blank area of the ArcGIS toolbox, then choose "Add Toolbox" shown in Figure 4.

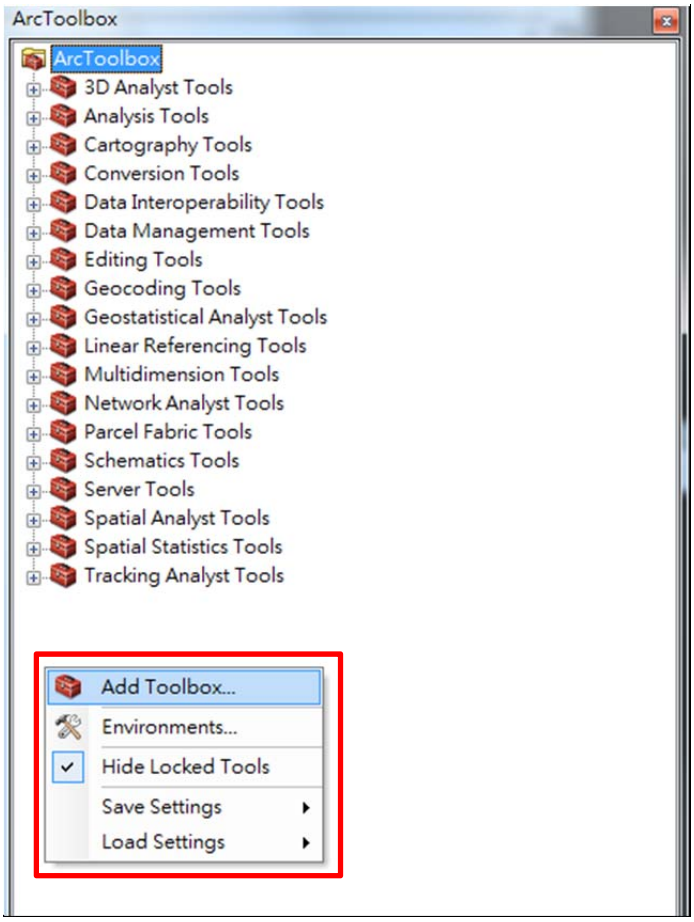


Figure 4 Steps of opening AcrToolbox and "Add toolbox".

Step3. Choose the pathway of GeoMasker file downloaded in Step1. and then select Geomask_10_1.tbx of "…\python\Geomask\GeoMask_10_1.tbx" shown in Figure 5.

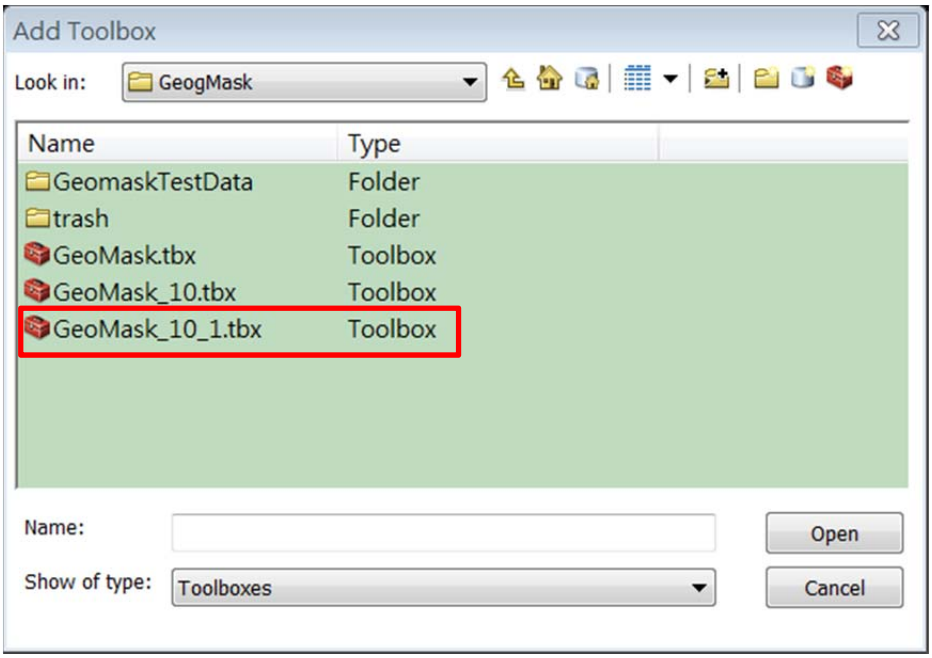


Figure 5 Pathway of GeoMask_10_1.tbx file.

Step4. Finish the installation steps. ArcGIS toolbox will add GeoMask_10_1 tool in user's ArcGIS 10. Figure 6 shows the successful installation of GeoMasker.

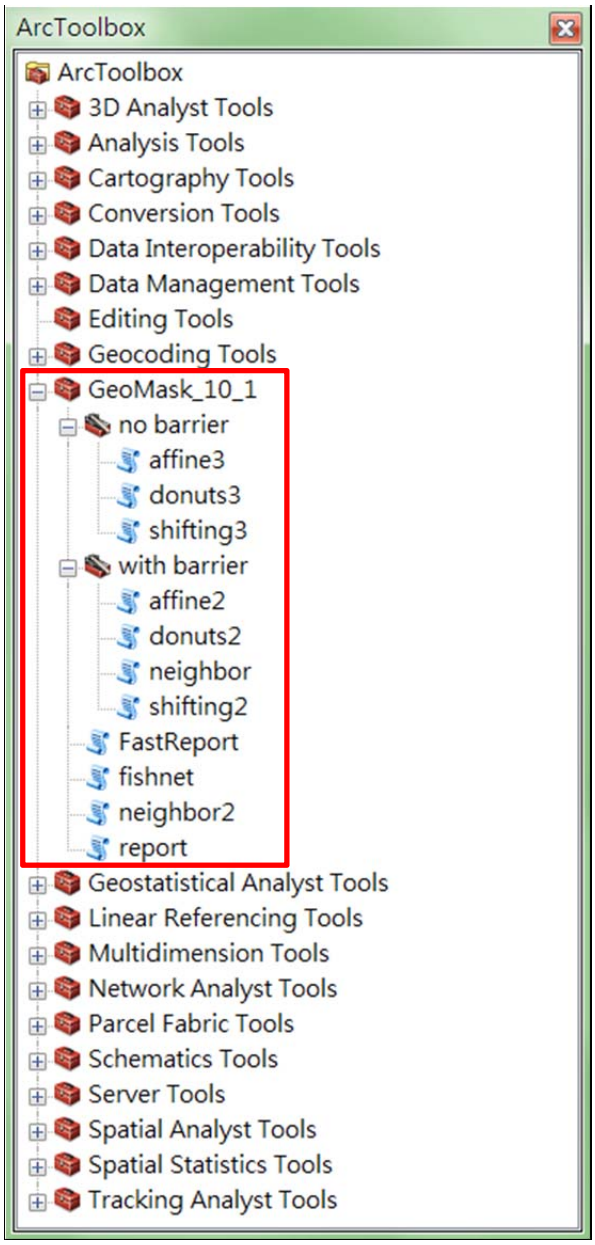


Figure 6 Successful installation of GeoMasker in ArcGIS 10.

Different kinds of GeoMasker tool box will be introduced in user guide subsequently.

4. Introduction of GeoMasker toolbox- non- barrier

Barrier in GeoMasker means areas that the data points should not be shifted to. Barriers can be sea, leak, park which normally people do not live in or stay in. It is unreasonable that the original locations of data points are shifted to those areas. In GeoMasker, user can select whether to use the barrier or non-barrier tool box based on their demands and data format.

4.1 Non-barrier tool box – Affine method

Affine method rotates and shifts the original location of data points to new location to achieve the effect of hiding the original data points. All the data points in Affine method will be rotated the same angle and shifted the same length of radius. If new location of data point is out of boundary, the new data point will automatically be rotated randomly- selected angle from 0^o to 360^o and shifted to the user defined radius when the out of boundary events occurred. The user interface of Affine method is shown in Figure 7.

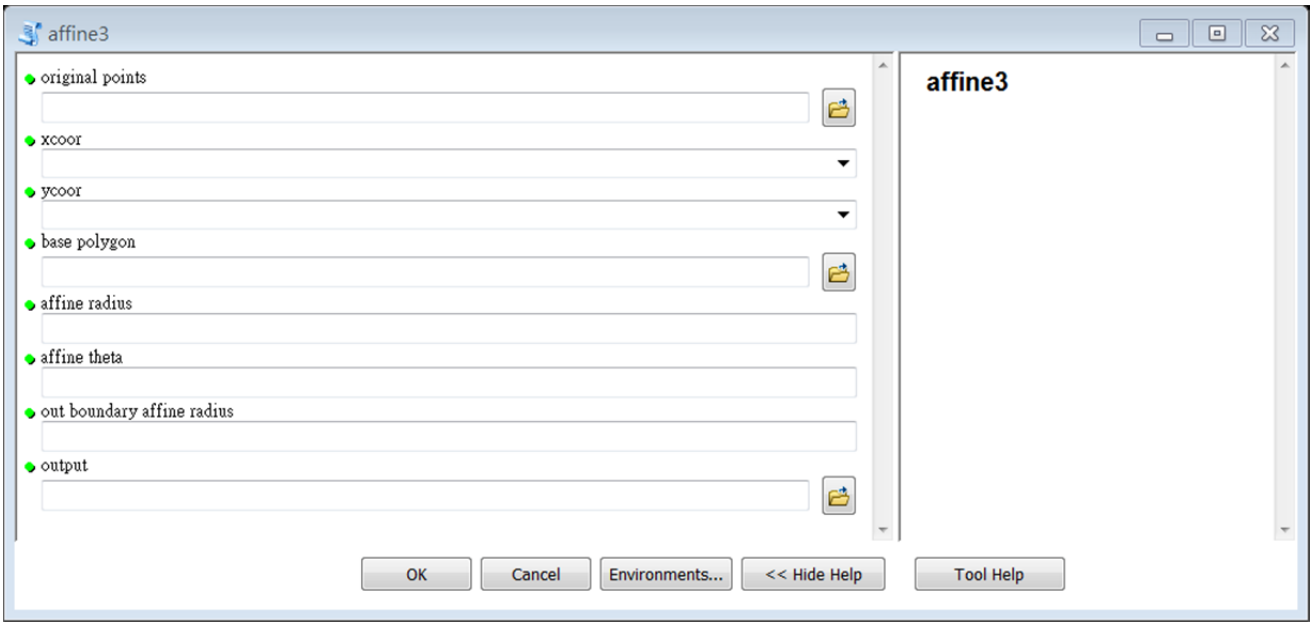


Figure 7 User interface of Affine method in no-barrier tool box.

original points: Data point with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as

y-coordinate in attribute table.

base polygon: The boundary that data points should be inside. If there is no specific boundary, user will have to create a wide extent of boundary which contains all the data points. The boundary must be a polygon and saved as a .shp file.

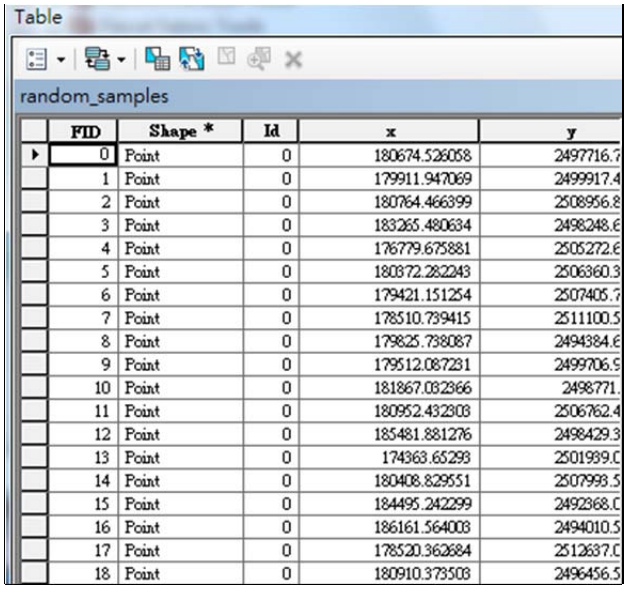
affine radius: Setting the length of radius used in Affine method. Too large length of radius may cause GeoMasker stop functioning.

affine theta: Inputting integers from 0 to 360 represents the angle from 0^o to 360^o.

out boundary affine radius: Setting the radius for those data points which are out of boundary after first rotated by GeoMasker to be rotated again.

output: Setting the pathway of user's output data points.

Figure 8 shows the example of data points with corresponding columns of x-coordinate and y-coordinate in attribute table. Example of boundary is shown in Figure 9. Both the original data point and boundary information must be saved as .shp file.



Table

random_samples

FID	Shape *	Id	x	y
0	Point	0	180674.526058	2497716.7
1	Point	0	179911.947069	2499917.4
2	Point	0	180764.466399	2508956.8
3	Point	0	183265.480634	2498248.6
4	Point	0	176779.675881	2505272.6
5	Point	0	180372.282243	2506360.3
6	Point	0	179421.151254	2507405.7
7	Point	0	178510.739415	2511100.5
8	Point	0	179825.738087	2494384.6
9	Point	0	179512.087231	2499706.5
10	Point	0	181867.082366	2498771.
11	Point	0	180952.432303	2506762.4
12	Point	0	185481.881276	2498429.3
13	Point	0	174363.65293	2501939.0
14	Point	0	180408.829551	2507998.5
15	Point	0	184495.242299	2492368.0
16	Point	0	186161.564003	2494010.5
17	Point	0	178520.362684	2512637.0
18	Point	0	180910.373503	2496456.5

Figure 8 Example of data points.

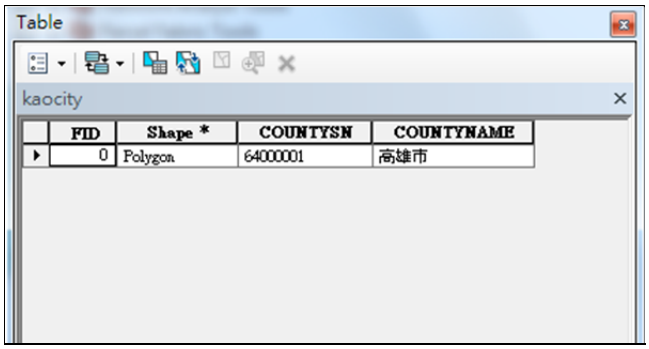


Figure 9 Example of boundary. [base polygon].

4.2 Non-barrier tool box – Shifting method

Shifting method shifts the original location of data points to new locations to achieve the effect of hiding the original data points. All the data points in Shifting method will be shifted the same distance which given by user in x-axis and y-axis. If new location of data point is out of boundary, the new data point will automatically be rotated randomly- selected angle from 0^{o} to 360^{o} and shifted to the user defined radius when the out of boundary events occurred. The user interface of Shifting method is shown in Figure 10.

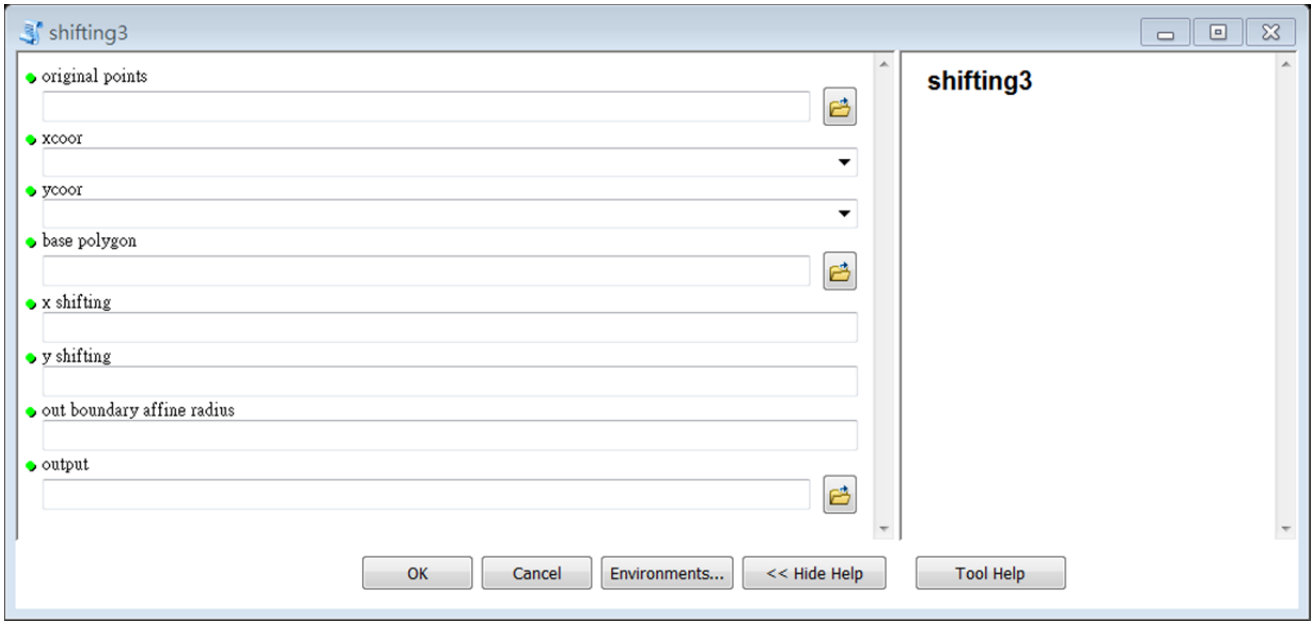


Figure 10 User interface of Shifting method in no-barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as

x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

base polygon: The boundary that all the data points should be inside. If there is no specific boundary, user will have to create a wide extent of boundary which contains all the data points. The boundary must be polygon and saved as a .shp file.

x shifting: Setting the value of shifting distance in x-axis.

y shifting: Setting the value of shifting distance in y-axis.

out boundary affine radius: Setting the radius for those data points which are out of boundary after first rotated by GeoMasker to be rotated again.

output: Setting the pathway of user's output data points.

Forms of data points and boundary information are the same as examples mentioned in Affine method which are shown in Figure 8 and Figure 9.

4.3 Non-barrier tool box – Donut method

Donut method rotates and shifts the original location of data points to new location to achieve the effect of hiding the original data points. Different from Affine method, Donut method rotates angle which is randomly selected in the range of 0^o to 360^o, so user doesn't have to set the angle value in Donut method. The radius in Donut method is different from Affine method, too. User has to provide maximum and minimum of radius, then the radius will be randomly selected a value between the maximum and minimum as the radius value. The user interface of Donut method is shown in Figure 11.

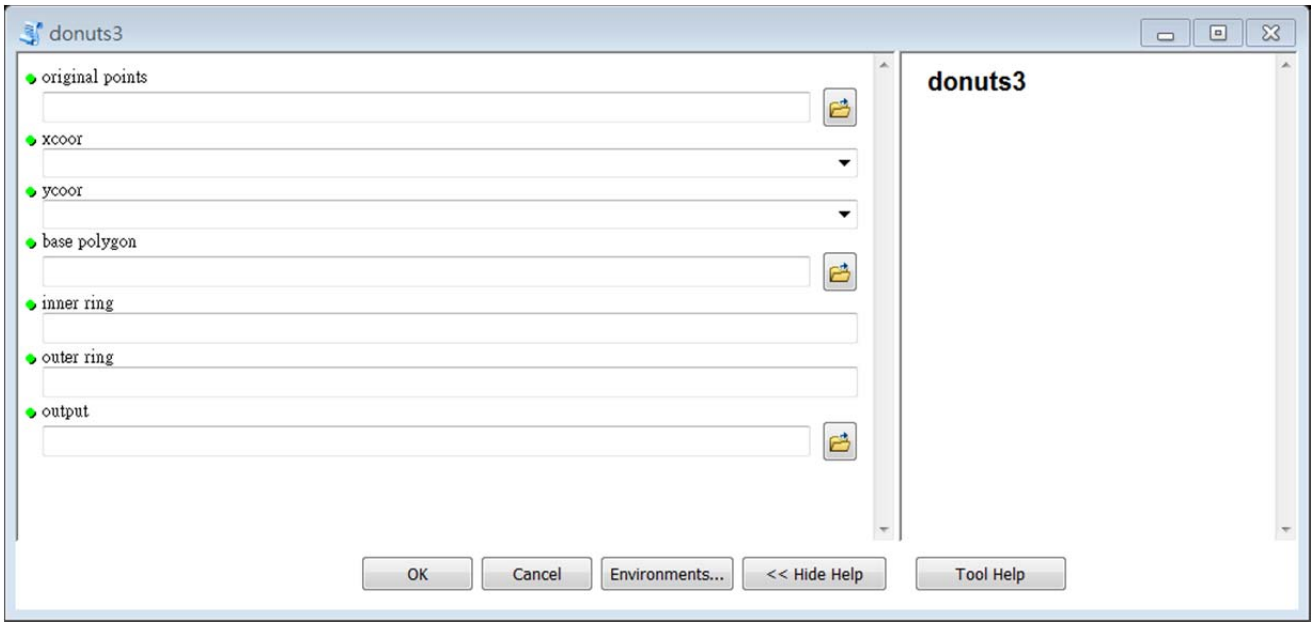


Figure 11 User interface of Donut method in no-barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

base polygon: The boundary that all the data points should be inside. If there is no specific boundary, user will have to create a wide range of boundary which contains all the data points. The boundary must be a polygon and saved as a .shp file.

inner ring: the minimum of radius that a data point should be shifted.

outer ring: the maximum of radius that a data point should be shifted.

output: Setting the pathway of user's output data points.

Data points and boundary information are the same as examples mentioned in Affine method which are shown in Figure 8 and Figure 9.

4.4 Neighbor method

Neighbor method considers the characteristics such as the population density of the boundary map of data point as barrier. The main consideration is that population density is different between areas. If data points are just shifted without any restriction, it will probably change the characteristics of boundary area such as the density of population in different area and it will be easy to guess that the data points are dealt with some processes. In such case, Neighbor method calculates the neighbor matrix of original data points and finds out the most similar neighbor area. Original data points are randomly moved to the most similar neighbor area to achieve the effect of hiding the original data points. The user interface of Neighbor method is shown in Figure 12.

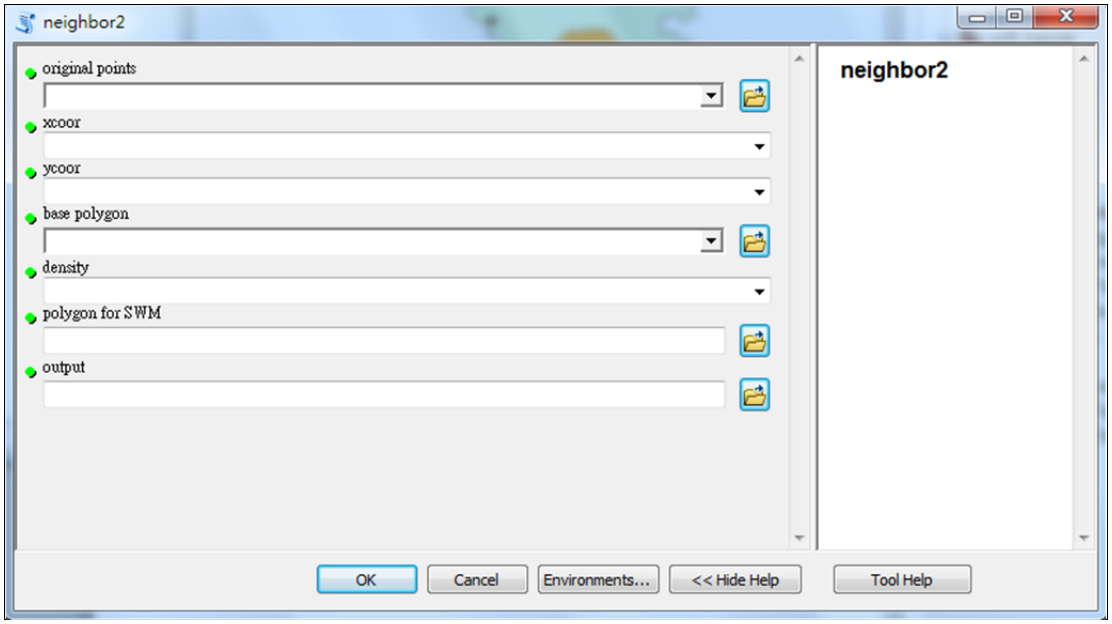


Figure 12 User interface of Neighbor method in barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

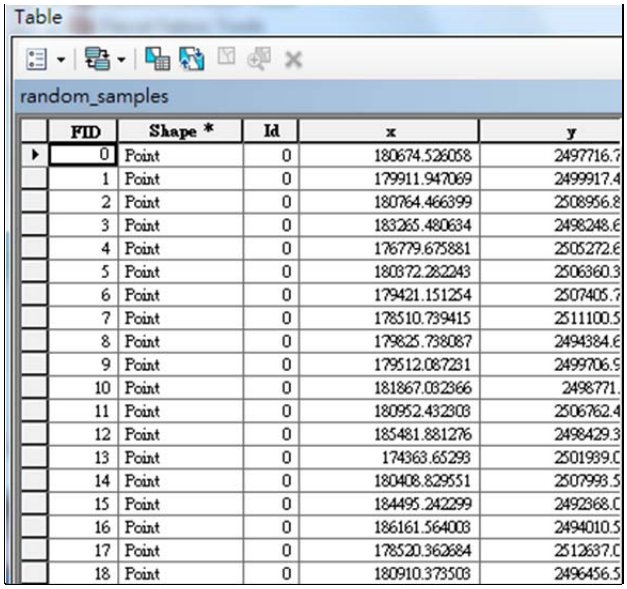
base polygon: The boundary that all the data points should be inside. The boundary must be a polygon and saved as a shp file. Different from other methods

mentioned above, ID code of different area and the density information of the boundary map which is used in Neighbor method should be built firstly by user. The example of boundary map of Neighbor method is shown as Figure 14.

density：The densities of different areas in boundary map. User can select the column of density information in the attribute table of base polygon .shp.

polygon for SWM：Polygon .shp file used to create spatial weight matrix. Usually the same as boundary map.

output: Setting the pathway of user's output data points.

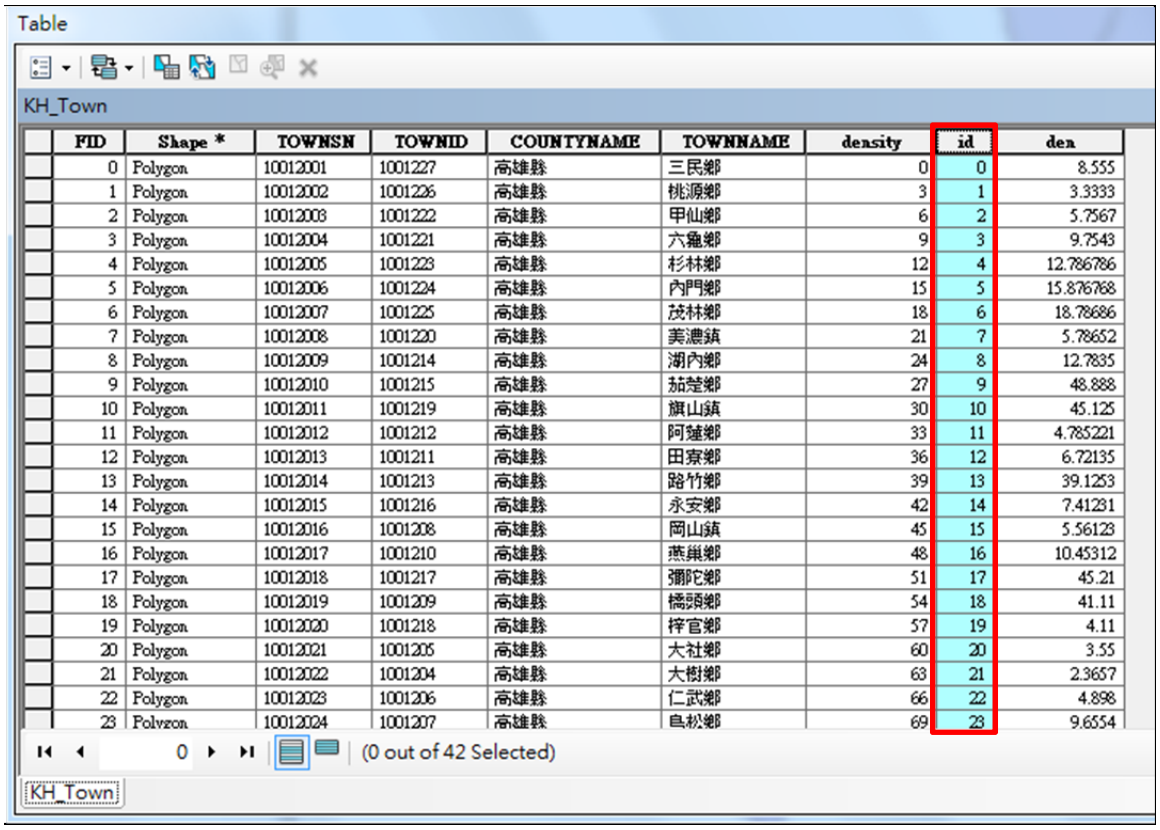


Table

random_samples

FID	Shape *	Id	x	y
0	Point	0	180674.526058	2497716.7
1	Point	0	179911.947069	2499917.4
2	Point	0	180764.466399	2508956.8
3	Point	0	183265.480634	2498248.6
4	Point	0	176779.675881	2505272.6
5	Point	0	180372.282243	2506360.3
6	Point	0	179421.151254	2507405.7
7	Point	0	178510.739415	2511100.5
8	Point	0	179825.738087	2494384.6
9	Point	0	179512.087231	2499706.9
10	Point	0	181867.082366	2498771.
11	Point	0	180952.432303	2506762.4
12	Point	0	185481.881276	2498429.3
13	Point	0	174363.65293	2501939.0
14	Point	0	180408.829551	2507993.5
15	Point	0	184495.242299	2492368.0
16	Point	0	186161.564008	2494010.5
17	Point	0	178520.362684	2512637.0
18	Point	0	180910.373503	2496456.5

Figure 13 Example of original point data.



Table

KH_Town

FID	Shape *	TOWNSN	TOWNID	COUNTYNAME	TOWNNAME	density	id	den
0	Polygon	10012001	1001227	高雄縣	三民鄉	0	0	8.555
1	Polygon	10012002	1001226	高雄縣	桃源鄉	3	1	3.3333
2	Polygon	10012003	1001222	高雄縣	甲仙鄉	6	2	5.7567
3	Polygon	10012004	1001221	高雄縣	六龜鄉	9	3	9.7543
4	Polygon	10012005	1001223	高雄縣	杉林鄉	12	4	12.786786
5	Polygon	10012006	1001224	高雄縣	內門鄉	15	5	15.876768
6	Polygon	10012007	1001225	高雄縣	茂林鄉	18	6	18.78686
7	Polygon	10012008	1001220	高雄縣	美濃鎮	21	7	5.78652
8	Polygon	10012009	1001214	高雄縣	湖內鄉	24	8	12.7835
9	Polygon	10012010	1001215	高雄縣	茄萣鄉	27	9	48.888
10	Polygon	10012011	1001219	高雄縣	旗山鎮	30	10	45.125
11	Polygon	10012012	1001212	高雄縣	阿蓮鄉	33	11	4.785221
12	Polygon	10012013	1001211	高雄縣	田寮鄉	36	12	6.72135
13	Polygon	10012014	1001213	高雄縣	路竹鄉	39	13	39.1253
14	Polygon	10012015	1001216	高雄縣	永安鄉	42	14	7.41231
15	Polygon	10012016	1001208	高雄縣	岡山鎮	45	15	5.56123
16	Polygon	10012017	1001210	高雄縣	燕巢鄉	48	16	10.45312
17	Polygon	10012018	1001217	高雄縣	彌陀鄉	51	17	45.21
18	Polygon	10012019	1001209	高雄縣	橋頭鄉	54	18	41.11
19	Polygon	10012020	1001218	高雄縣	梓官鄉	57	19	4.11
20	Polygon	10012021	1001205	高雄縣	大社鄉	60	20	3.55
21	Polygon	10012022	1001204	高雄縣	大樹鄉	63	21	2.3657
22	Polygon	10012023	1001206	高雄縣	仁武鄉	66	22	4.898
23	Polygon	10012024	1001207	高雄縣	鳥松鄉	69	23	9.6554

0 (0 out of 42 Selected)

KH_Town

Figure 14 Example of boundary map of Neighbor map. Data column is the density information of different areas.

5. Introduction of GeoMasker toolbox- barrier

In barrier tool box, the principle of rotating and shifting method are the same as non-barrier tool box. Only one thing is different is that additional barrier should be considered in barrier tool box. Not only should the boundary of all data points but also other barrier such as sea, leak and park be considered and set by user.

5.1 Barrier tool box – Affine method

Principle of Affine method in barrier tool box is the same as in non-barrier tool box. But in barrier tool box, user has to provide the barrier condition. The user interface of Affine method is shown in Figure 15.

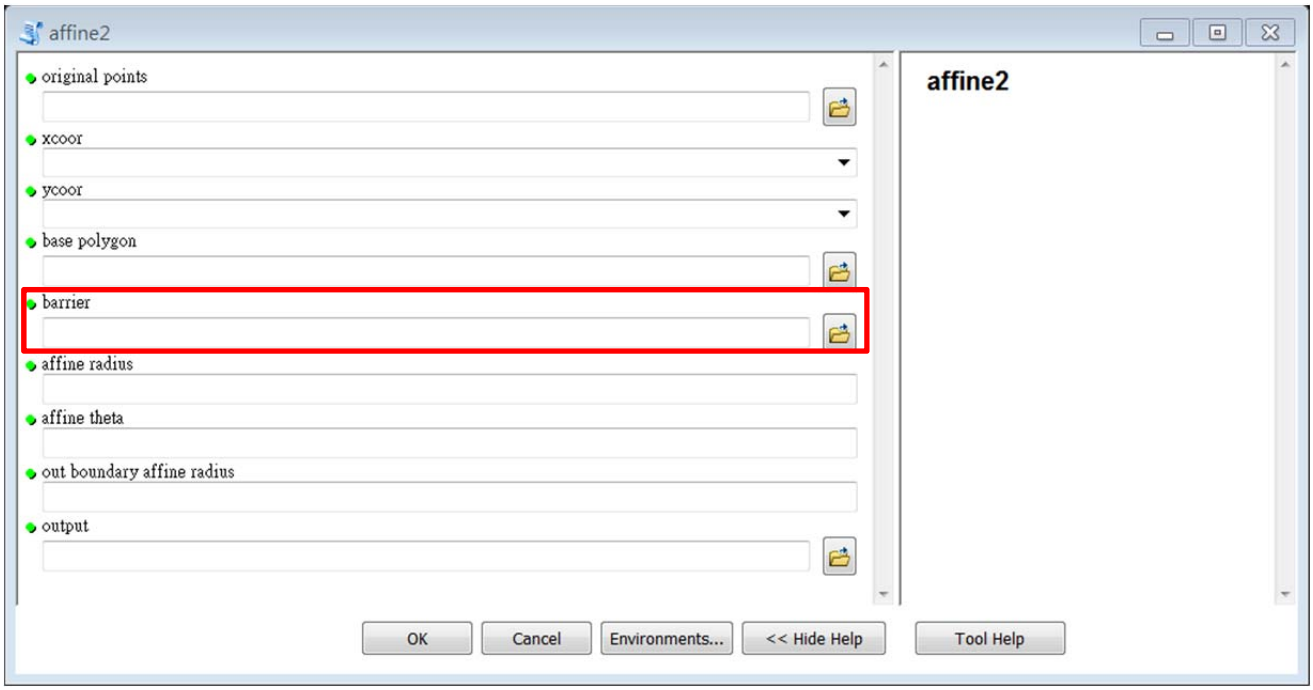


Figure 15 User interface of Affine method in barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

base polygon: The boundary that all the data points should be inside. If there is no specific boundary, user will have to create a wide extent of boundary which contains all the data points. The boundary must be a polygon and saved as a .shp

file.

barrier: Setting the area that new points won't be inside. The barrier must be a polygon and saved as a .shp file.

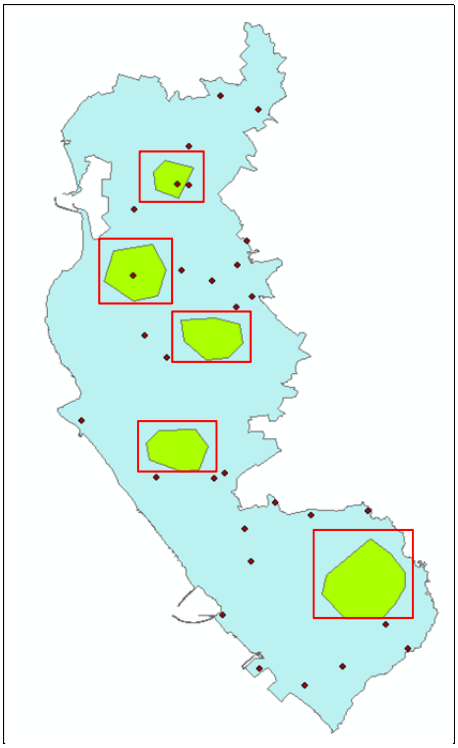
affine radius: Setting the radius which used in Affine method. Too large length of radius may cause GeoMasker stop functioning.

affine theta: Inputting integers from 0 to 360 which represent the angle from 0^{o} to 360^{o}.

out boundary affine radius: Setting the radius for those data points which are out of boundary after first rotate by GeoMasker to rotate again.

output: Setting the pathway of user's output data points.

Data points and boundary information are the same as examples mentioned in Affine method in non-barrier tool box which are shown as Figure 8 and Figure 9. The example of barrier is shown in Figure 16, green area are the barriers in the map.



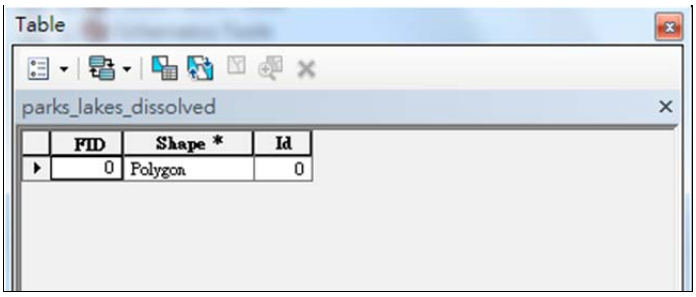


Figure 16 Example of barrier.

5.2 Barrier tool box – Shifting method

Principle of Shifting method in barrier tool box is the same as in non-barrier tool box. But in barrier tool box, user has to provide the barrier condition. The user interface of Shifting method is shown in Figure 17.

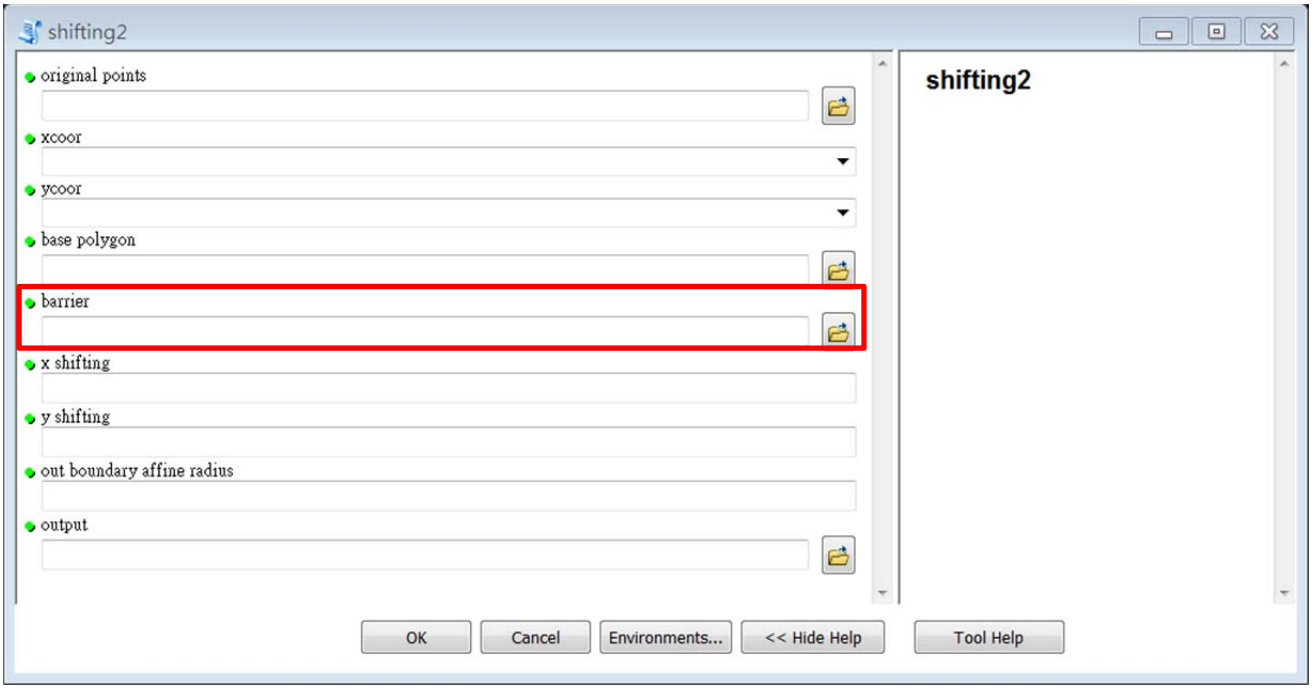


Figure 17 User interface of Shifting method in barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

base polygon: The boundary that all the data points should be inside. If there is no

specific boundary, user will have to create a wide extent of boundary which contains all the data points. The boundary must be polygon and saved as a .shp file.

barrier: Setting the area that new points won't be inside. The barrier must be a polygon and saved as a .shp file. Example of barrier is the same as Figure 16.

x shifting: Setting the value of shifting distance in x direction.

y shifting: Setting the value of shifting distance in y direction.

Out boundary affine radius: Setting the radius for those data points which are out of boundary after first rotate by GeoMasker to rotate again.

output: Setting the pathway of user's output data points.

Data points and boundary information are the same as examples mentioned in Affine method which are shown as Figure 8 and Figure 9.

5.3 Barrier tool box – Donut method

Principle of Donut method in barrier tool box is the same as in non-barrier tool box. But in barrier tool box, user has to provide the barrier condition. The user interface of Donut method is shown in Figure 18.

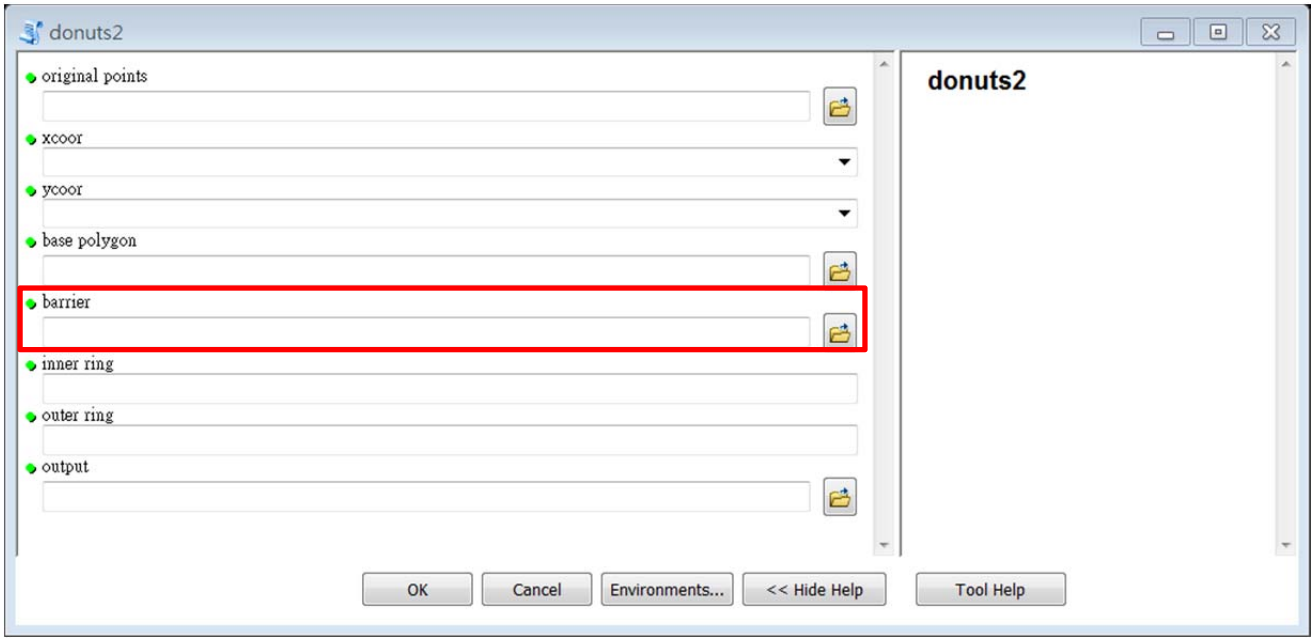


Figure 18 User interface of Donut method in barrier tool box.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

base polygon: The boundary that all the data points should be inside. If there is no specific boundary, user will have to create a wide extent of boundary which contains all the data points. The boundary must be a polygon and saved as a .shp file.

barrier: Setting the area that new points won't be inside. The barrier must be a polygon and saved as a .shp file. Example of barrier us the same as Figure 16.

inner ring: the minimum of radius that a data point should be shifted.

outer ring: the maximum of radius that a data point should be shifted.

output: Setting the pathway of user's output data points.

Data points and boundary information are the same as examples mentioned in Affine method which are shown as Figure 8 and Figure 9.

5.4 Fishnet method

Different from methods mentioned above, Fishnet utilizes the grids to hide the actual location of original points. The user interface of Fishnet method is shown in Figure 19.

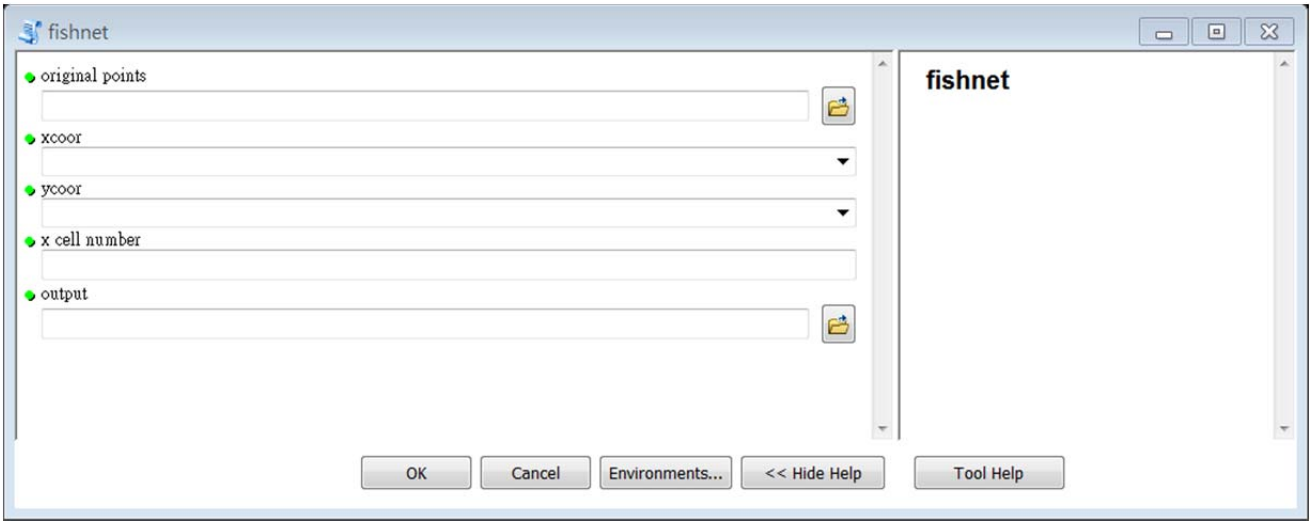


Figure 19 User interface of Fishnet method.

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: x coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

x cell number: Setting the number of cell in x-axis.

output: Setting the pathway of user output data points.

6. Examine the results GeoMasker

For different types of geomasking methods, it is important to examine the results of those methods because that the results will be used to conduct advanced analysis or may be published. If the results are significantly different from the original data points and the new data points lose the original geographical feature, results dealt with GeoMasker will be not suitable to conduct advanced analysis and will be easy to guess that the data points are dealt with some processes. Therefore, GeoMasker provides some methods to examine the shifts of data points whether to meet the demand of users or not.

6.1 Report

Report utilizes the characteristics such as the population density of the boundary map to compare the difference between after shifting and before shifting. The user interface of Report is shown as Figure 20.

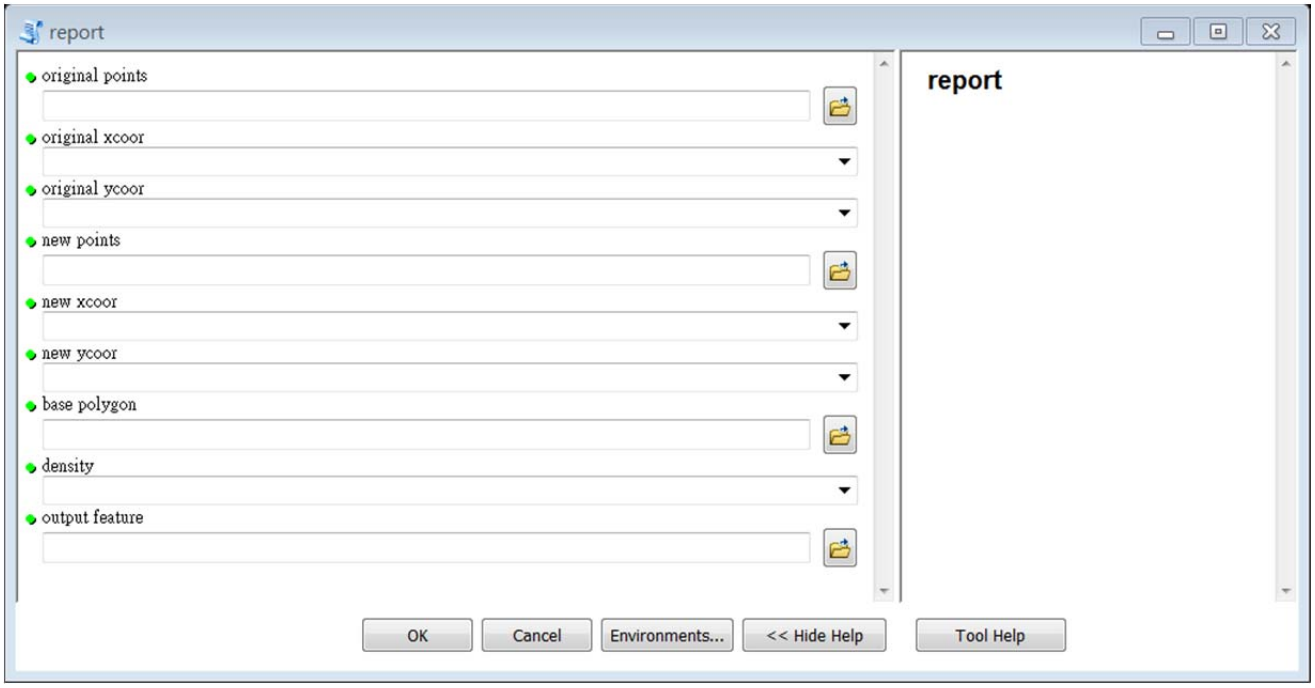


Figure 20 user interface of Report

original points: Data points with x-coordinate and y-coordinate in attribute table.

xcoor: X coordinate of each data point. In GeoMasker, user can select column as x-coordinate in attribute table.

ycoor: Y coordinate of each data point. In GeoMasker, user can select column as y-coordinate in attribute table.

new points: The shp file which contains the moved new data points dealt by Affine method, Shifting method, Donut method, and Neighbor method.

new xcoor: x coordinate of new points in the new point shp file. User can select the column which represents the x-coordinate of new data points.

new ycoor: y coordinate of new points in the new point shp file. User can select the column which represents the y-coordinate of new data points.

base polygon: The boundary that data point should be inside. The boundary must be a polygon and saved as a shp file. ID code of different area and the density information of the boundary map which is used in Report should be built first by user. The example of boundary map of Report is shown the same as Figure 14.

density：The densities of different areas in boundary map. User can select the column of density information in the attribute table of base polygon shp. Density can help user to test the GeoMasker result.

output feature: Setting the pathway of user output data points.

6.2 FastReport

FastReport considers that user may want to compare the difference between after shifting and before shifting by not only one characteristic. The user interface of Report is shown as Figure 21.

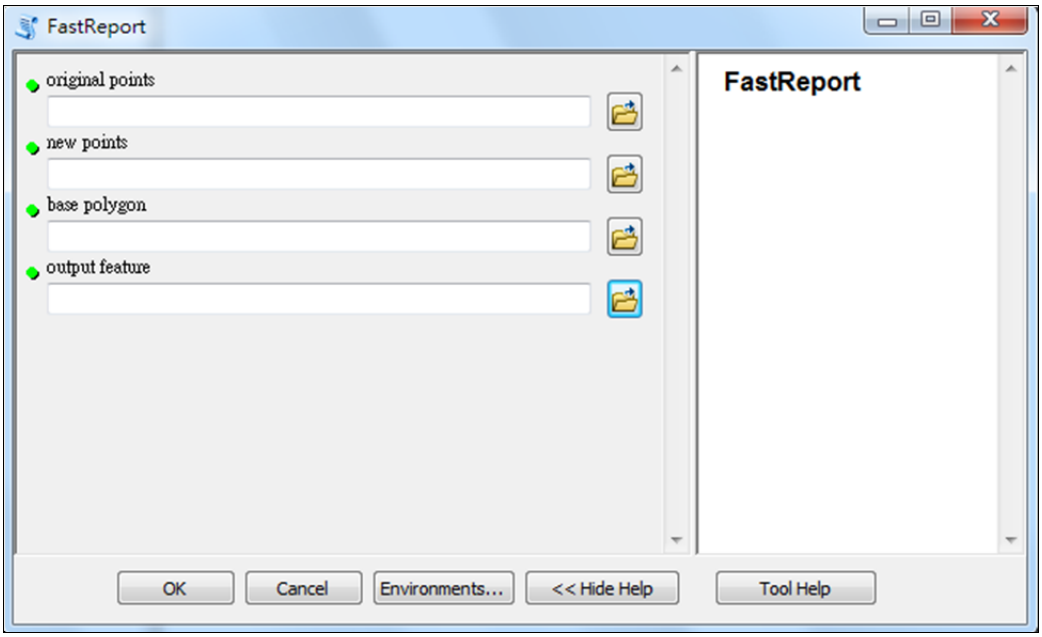


Figure 21 user interface of FastReport.

original points: Data points with x-coordinate and y-coordinate in attribute table.

new points: The .shp file which contains the moved new data points dealt by Affine method, Shifting method, Donut method, and Neighbor method.

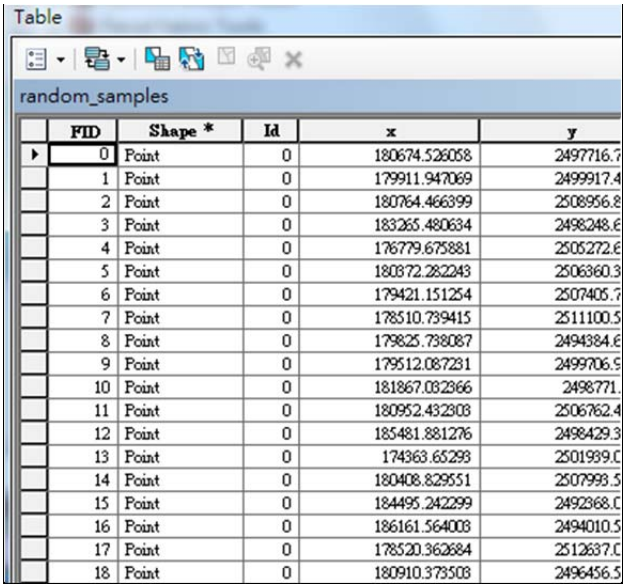
base polygon: The boundary that data point should be inside. The boundary must be a polygon and saved as a .shp file. ID code of different area and the density information of the boundary map which is used in Report should be built first by user. The example of boundary map of Report is shown the same as Figure 14.

output feature: Setting the pathway of user output data points.

7. Training exercise

7.1 Non-Barrier tool box- Affine method, shifting method and Donut method

The following section offer example of Affine method, Shifting method, Donut method and Neighbor of non-barrier tool box of GeoMasker. Users can, on their own, follow the steps outlined in example to better understand the GeoMasker process and the interaction of the components described in the User's Guide. All of three methods are demonstrated with the same example original point data, "random_samples.shp", and the same base polygon, "kaocity.shp". Two kinds of example data are shown as following:

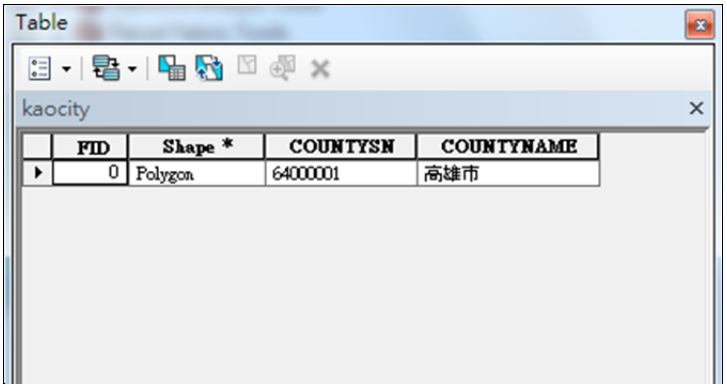


Table

random_samples

FID	Shape *	Id	x	y
0	Point	0	180674.526058	2497716.7
1	Point	0	179911.947069	2499917.4
2	Point	0	180764.466399	2508956.8
3	Point	0	183265.480634	2498248.6
4	Point	0	176779.675881	2505272.6
5	Point	0	180372.282243	2506360.3
6	Point	0	179421.151254	2507405.7
7	Point	0	178510.739415	2511100.5
8	Point	0	179825.738087	2494384.6
9	Point	0	179512.087231	2499706.9
10	Point	0	181867.032366	2498771.
11	Point	0	180952.432303	2506762.4
12	Point	0	185481.881276	2498429.3
13	Point	0	174363.65293	2501939.0
14	Point	0	180408.829551	2507993.5
15	Point	0	184495.242299	2492368.0
16	Point	0	186161.564003	2494010.5
17	Point	0	178520.362684	2512637.0
18	Point	0	180910.373503	2496456.5

Figure 22 Attribute table of random_samples.shp.



Table

kaocity

FID	Shape *	COUNTYSN	COUNTYNAME
0	Polygon	64000001	高雄市

Figure 23 Attribute table of kaocity.shp.

Different settings of three methods are shown as following table, after setting the parameters user can push the bottom "OK" to run the GeoMasker.

Affine method	Shifting method	Donut method
original points: random_samples.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **affine radius:** 1000 **affine theta:** 45 **out boundary** **affine radius:** 1000	**original points:** random_samples.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **X shifting:** 2000 **Y shifting:** 0 **out boundary** **affine radius:** 2000	**original points:** random_sample.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **Inner ring:** 1000 **Outer ring:** 2000

Different methods can cause different shift of original data point shown as Figure 24, Figure 25 and Figure 26. The bigger points are the original data points and the smaller points are the shifted new data points.

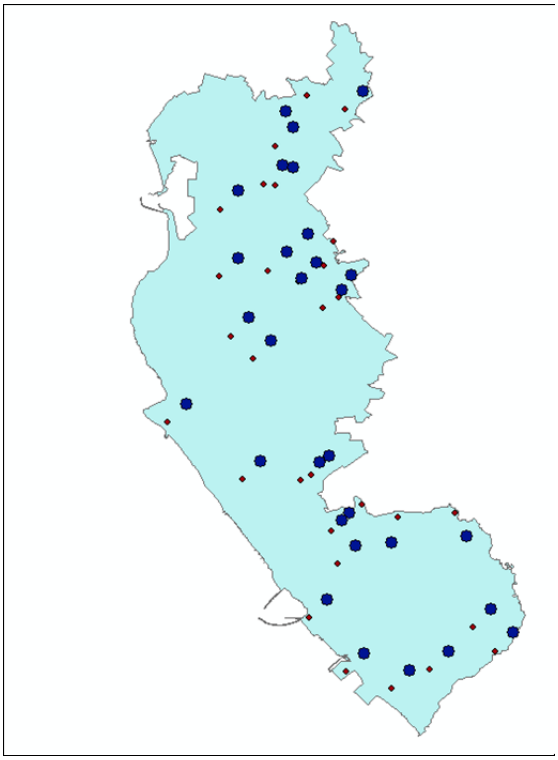


Figure 24 Result of Affine method of non-barrier tool box.

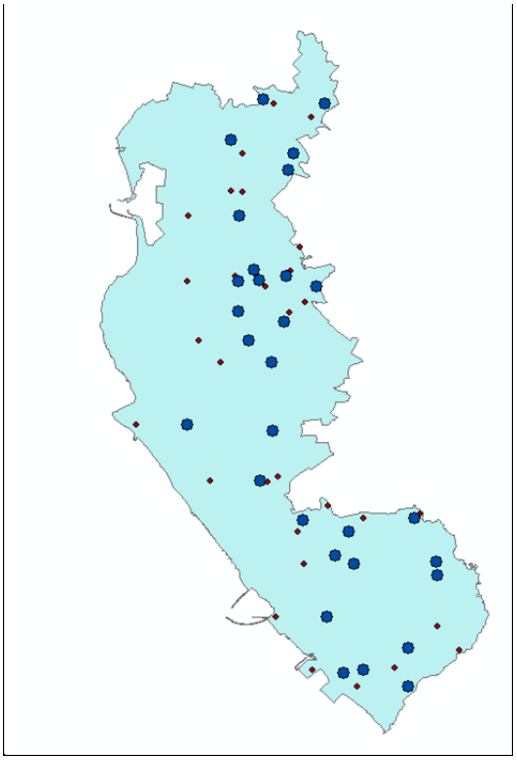


Figure 25 Result of Shifting method of non-barrier tool box.

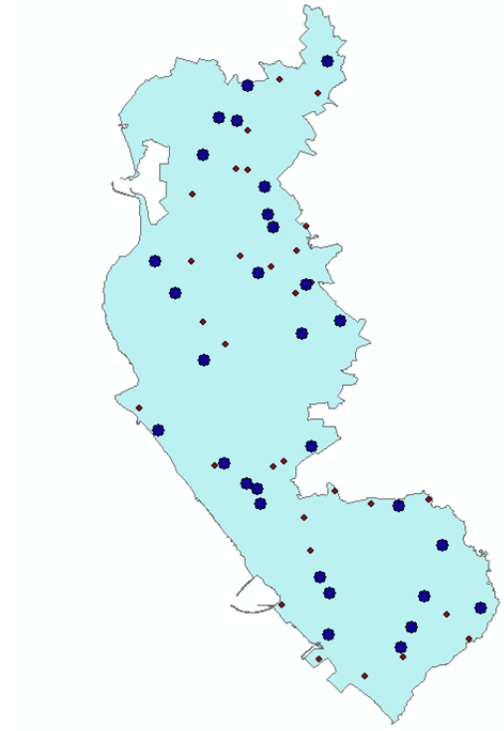


Figure 26 Result of Donut method of non-barrier tool box.

7.2 Barrier tool box- Affine method, shifting method and Donut method

This section offers examples of Affine method, shifting method and donut method of barrier tool box of GeoMasker. All of three methods are demonstrated with the same example original point data, "random_samples.shp", and the same base polygon, "kaocity.shp". Two example data are shown as Figure 22 and Figure 23. Example of barrier, "parks_lakes_dissolved.shp" is shown in Figure 27:

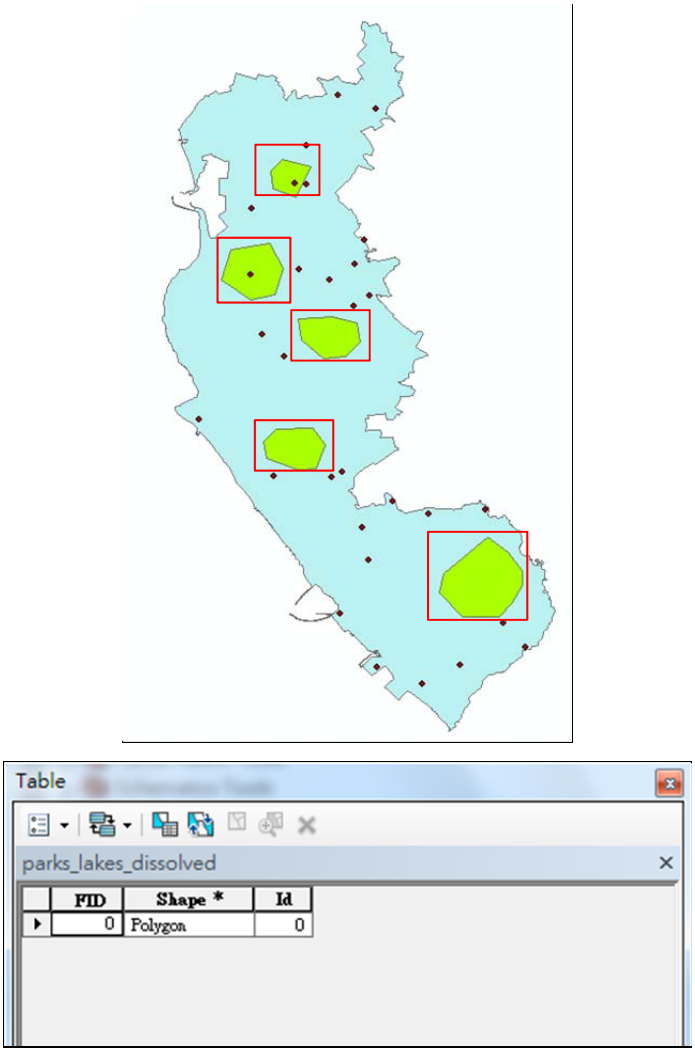


Figure 27 Example of barrier -Parks_lakes_dissolved.shp and corresponding attribute table.

Different settings of three methods are shown as following table, after setting the parameters user can push the bottom "OK" to run the GeoMasker.

Affine method	Shifting method	Donut method
original points: random_samples.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **barrier**: parks_lakes_dissolved.shp **affine radius:** 1000 **affine theta:** 45 **out boundary** **affine radius:** 1000	**original points:** random_samples.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **barrier**: parks_lakes_dissolved.shp **X shifting:** 2000 **Y shifting:** 0 **out boundary** **affine radius:** 2000	**original points:** random_samples.shp **xcoor:** x **ycoor:** y **base polygon:** kaocity.shp **barrier**: parks_lakes_dissolved.shp **Inner ring:** 1000 **Outer ring:** 2000

Different methods can cause different shifts of original data point shown as Figure 28, Figure 29 and Figure 30. The bigger points are the original data points and the smaller points are the shifted new data points.

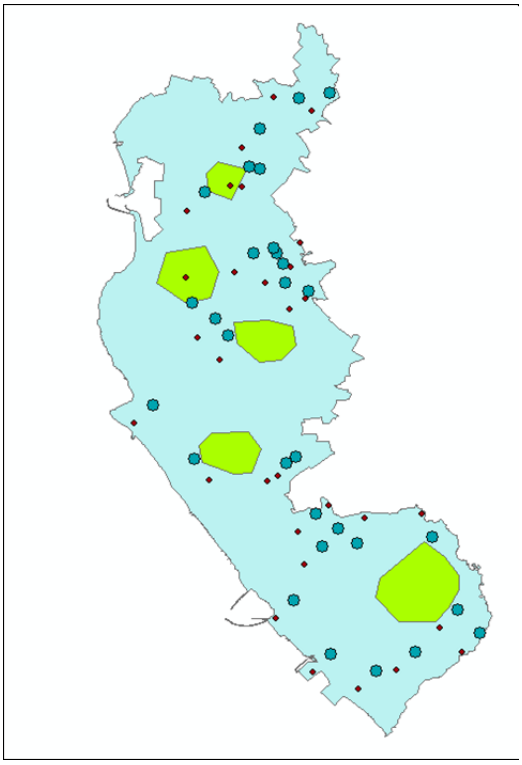


Figure 28 Result of Affine method of barrier tool box.

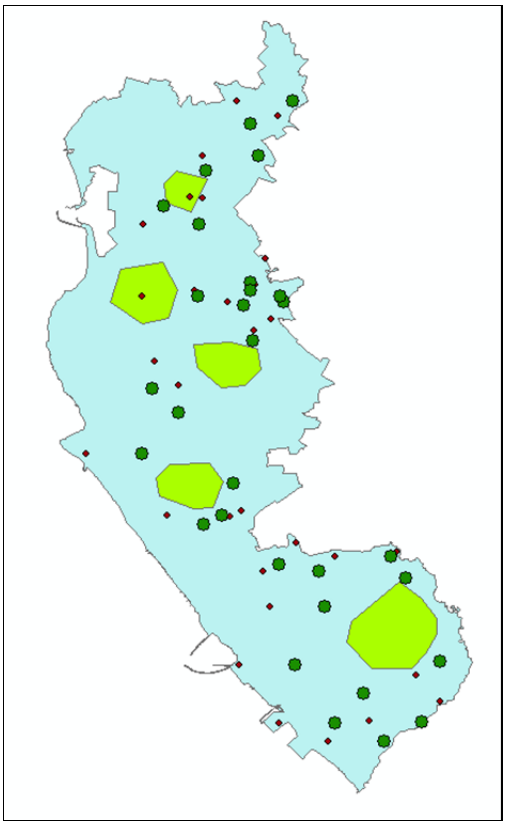


Figure 29 Result of Shifting method of barrier tool box.

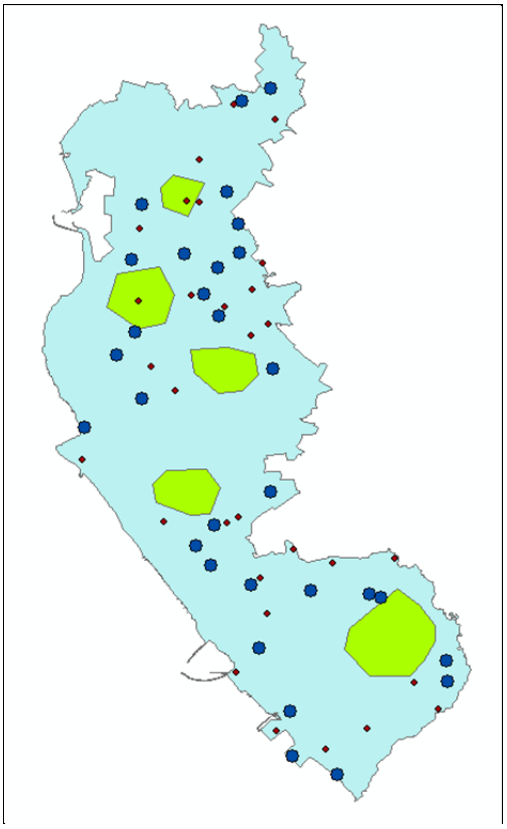


Figure 30 Result of Donut method of barrier tool box.

7.3 Non-Barrier tool box- Neighbor method

This section offers examples of Neighbor method of non-barrier tool box of GeoMasker demonstrated with the example original point data, "random_samples.shp", and base polygon, "KH_Town.shp". Example original point data is shown as Figure 22. But the base polygon, KH_Town.shp , should add a new column, ID, which is the same as column, FID, which is shown in Figure 31:

Table

KH_Town

FID	Shape *	TOWNSN	TOWNID	COUNTYNAME	TOWNNAME	density	id	den
0	Polygon	10012001	1001227	高雄縣	三民鄉	0	0	8.555
1	Polygon	10012002	1001226	高雄縣	桃源鄉	3	1	3.3333
2	Polygon	10012003	1001222	高雄縣	甲仙鄉	6	2	5.7567
3	Polygon	10012004	1001221	高雄縣	六龜鄉	9	3	9.7543
4	Polygon	10012005	1001223	高雄縣	杉林鄉	12	4	12.786786
5	Polygon	10012006	1001224	高雄縣	內門鄉	15	5	15.876768
6	Polygon	10012007	1001225	高雄縣	茂林鄉	18	6	18.78686
7	Polygon	10012008	1001220	高雄縣	美濃鎮	21	7	5.78652
8	Polygon	10012009	1001214	高雄縣	湖內鄉	24	8	12.7835
9	Polygon	10012010	1001215	高雄縣	茄萣鄉	27	9	48.888
10	Polygon	10012011	1001219	高雄縣	旗山鎮	30	10	45.125
11	Polygon	10012012	1001212	高雄縣	阿蓮鄉	33	11	4.785221
12	Polygon	10012013	1001211	高雄縣	田寮鄉	36	12	6.72135
13	Polygon	10012014	1001213	高雄縣	路竹鄉	39	13	39.1253
14	Polygon	10012015	1001216	高雄縣	永安鄉	42	14	7.41231
15	Polygon	10012016	1001208	高雄縣	岡山鎮	45	15	5.56123
16	Polygon	10012017	1001210	高雄縣	燕巢鄉	48	16	10.45312
17	Polygon	10012018	1001217	高雄縣	彌陀鄉	51	17	45.21
18	Polygon	10012019	1001209	高雄縣	橋頭鄉	54	18	41.11
19	Polygon	10012020	1001218	高雄縣	梓官鄉	57	19	4.11
20	Polygon	10012021	1001205	高雄縣	大社鄉	60	20	3.55
21	Polygon	10012022	1001204	高雄縣	大樹鄉	63	21	2.3657
22	Polygon	10012023	1001206	高雄縣	仁武鄉	66	22	4.898
23	Polygon	10012024	1001207	高雄縣	鳥松鄉	69	23	9.6554

0 (0 out of 42 Selected)

KH_Town

Figure 31 Attribute table of KH_town.shp.

Setting parameters are shown in Figure 32. User should select the data of polygon for SWM in the folder of user's PC. After setting the parameters, user can push the bottom "OK" to run the GeoMasker.

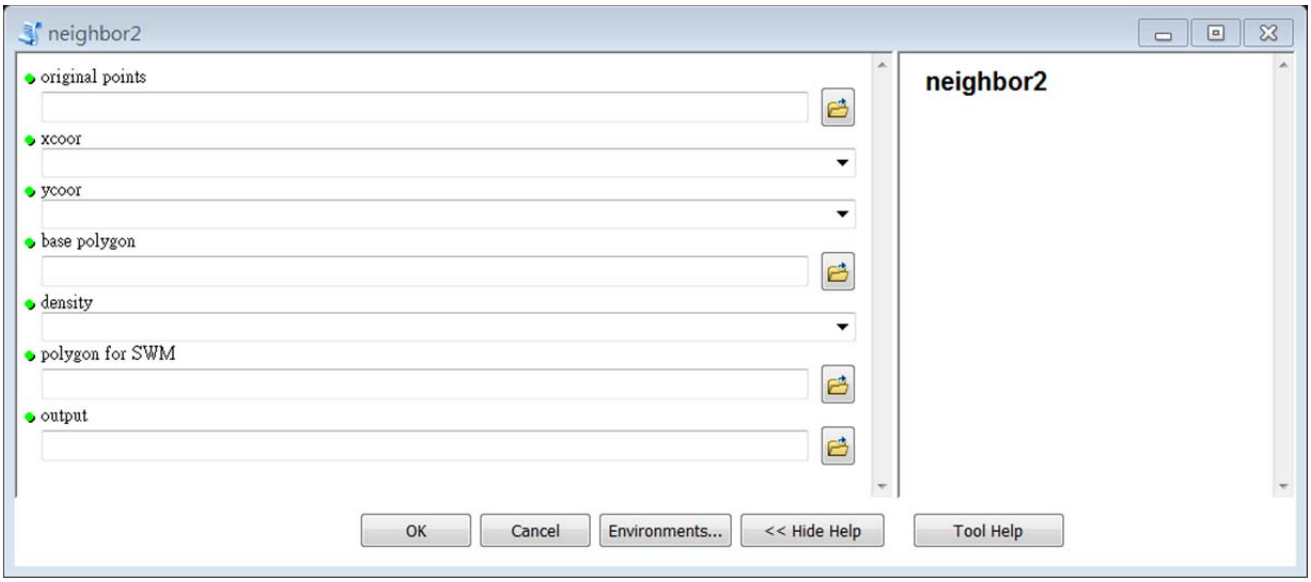


Figure 32 Example of setting parameters of Neighbor method.

The result of Neighbor method is shown in Figure 33. The bigger points are the original data points and the smaller points are the shifted new data points. Original data points are moved to the most similar neighbor.

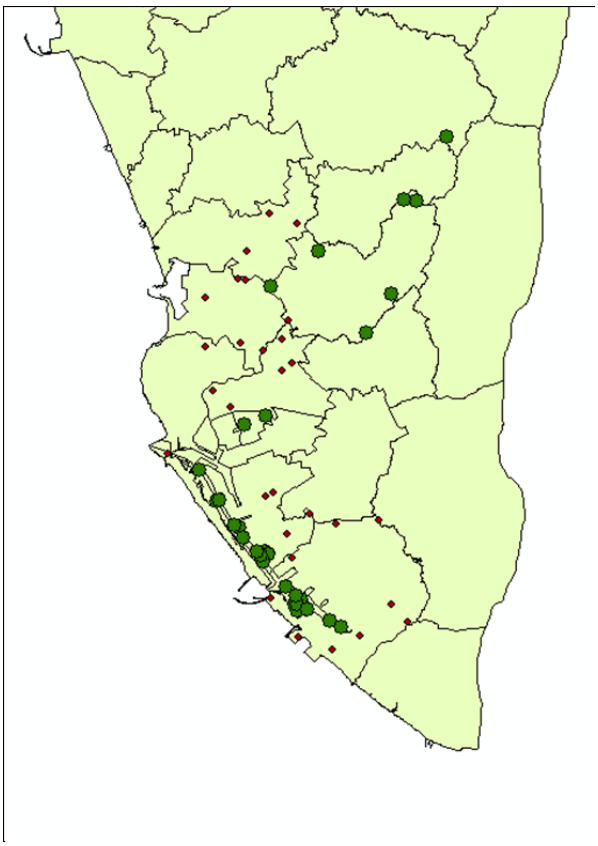


Figure 33 Example of result of Neighbor.

7.4 Fishnet method

This section offers examples of Fishnet method of GeoMasker demonstrated with the example original point data, “random_samples.shp” shown as Figure 22. Different from other methods, Fishnet doesn’t need the base polygon. User can choose the cell number to determine the grid number of X-axis. In the example, the cell number is chosen as 50.

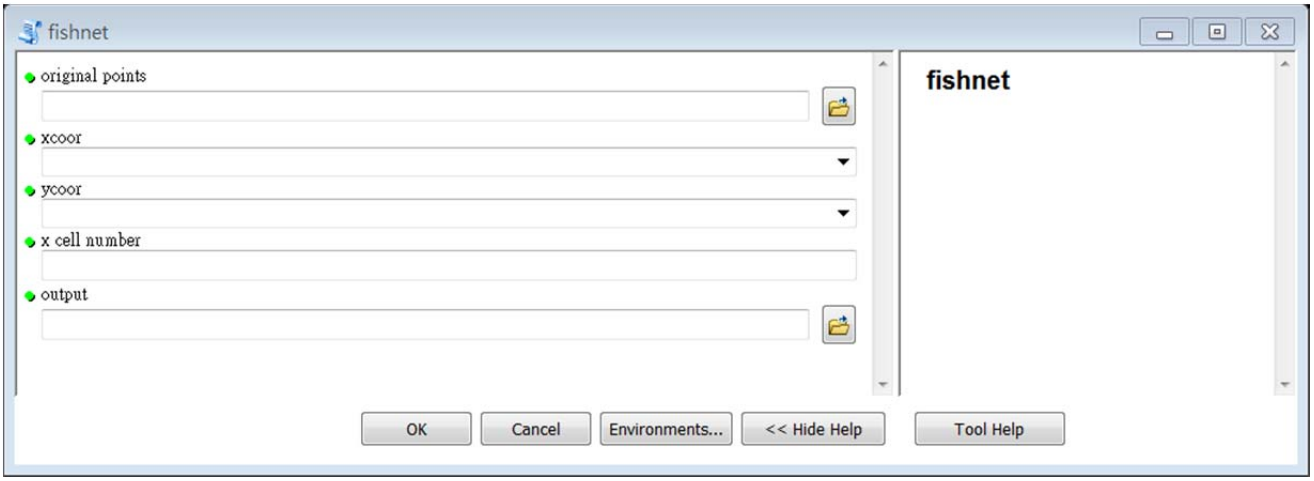


Figure 34 Example of setting parameters of Fishnet method.

The result of Fishnet method is shown in Figure 35. “Join_Count” column in attribute table of this shp file records the number of original data points contained in each grid.

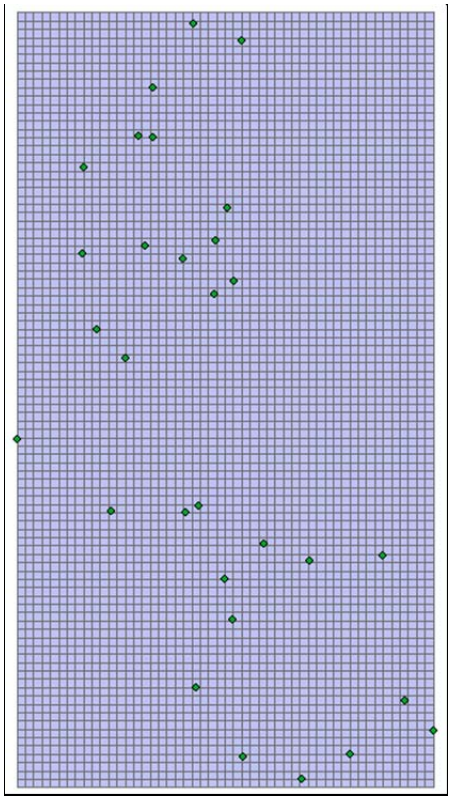
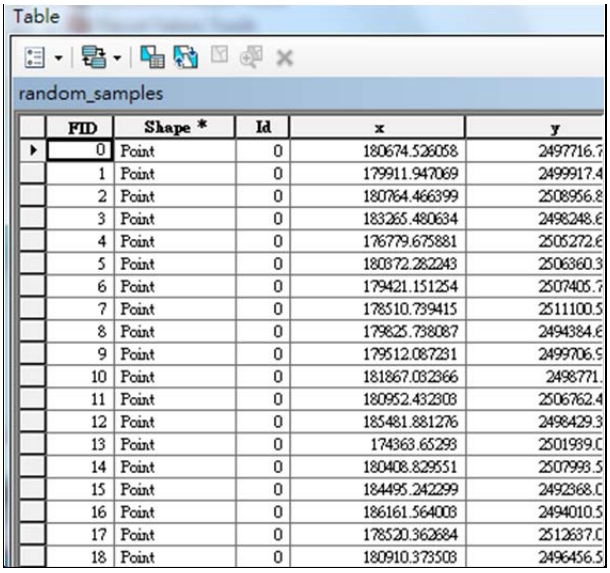


Figure 35 Example of result of Fishnet method.

7.5 Report

The following section offer example of Report method tool box of GeoMasker. Users can, on their own, follow the steps outlined in example to better understand the GeoMasker process and the interaction of the components described in the User's Guide. One method is demonstrated with the same example original point data, "random_samples.shp", and the same base polygon, "kaocity.shp". Two kinds of example data are shown as following:

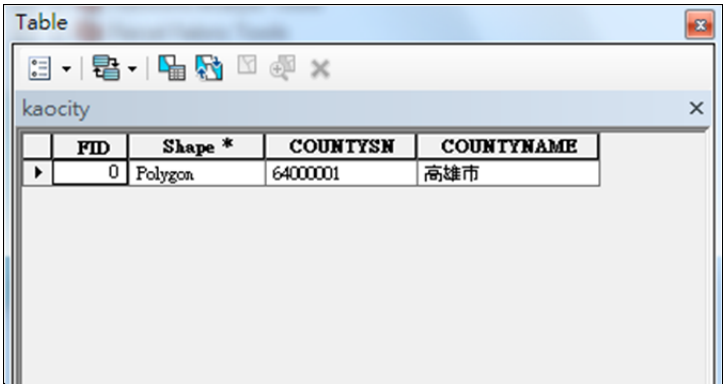


Table

random_samples

FID	Shape *	Id	x	y
0	Point	0	180674.526058	2497716.7
1	Point	0	179911.947069	2499917.4
2	Point	0	180764.466399	2508956.8
3	Point	0	183265.480634	2498248.6
4	Point	0	176779.675881	2505272.6
5	Point	0	180872.282243	2506360.3
6	Point	0	179421.151254	2507405.7
7	Point	0	178510.739415	2511100.5
8	Point	0	179825.738087	2494384.6
9	Point	0	179512.087231	2499706.9
10	Point	0	181867.082366	2498771.
11	Point	0	180952.432303	2506762.4
12	Point	0	185481.881276	2498429.3
13	Point	0	174363.65293	2501989.0
14	Point	0	180408.829551	2507998.5
15	Point	0	184495.242299	2492368.0
16	Point	0	186161.564003	2494010.5
17	Point	0	178520.362684	2512637.0
18	Point	0	180910.373503	2496456.5

Figure 36 Attribute table of random_samples.shp.



Table

kaocity

FID	Shape *	COUNTYSN	COUNTYNAME
0	Polygon	64000001	高雄市

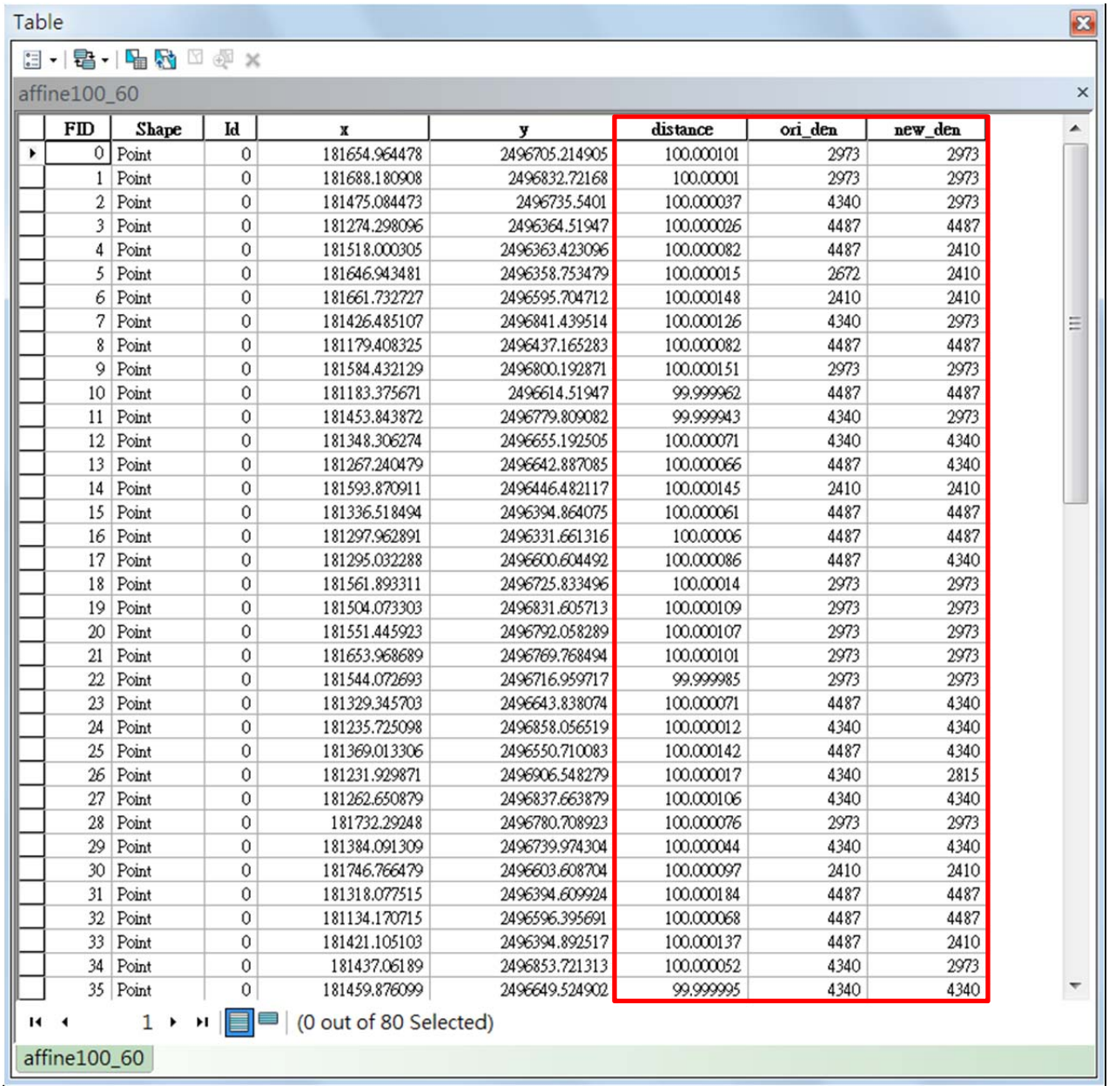
Figure 37 Attribute table of kaocity.shp.

Different settings of one method is shown as following table, after setting the parameters user can push the bottom “OK” to run the GeoMasker.

Report method
original points:
random_samples.shp
xcoor: x
ycoor: y
new points:
affine100_60.shp
new xcoor: x
new ycoor: y
base polygon: kaocity.shp
density: CENSUS

Output is a Attribute table presented in a way,Attribute table Join point layer after processing,ori_den: population density value represents the original point where the layers,new_den: represents a new point where the layers of the population density.

The statements provided to the users for their reference.



Table

affine100_60

FID	Shape	Id	x	y	distance	ori_den	new_den
0	Point	0	181654.964478	2496705.214905	100.000101	2973	2973
1	Point	0	181688.180908	2496832.72168	100.00001	2973	2973
2	Point	0	181475.084473	2496735.5401	100.000037	4340	2973
3	Point	0	181274.298096	2496364.51947	100.000026	4487	4487
4	Point	0	181518.000305	2496363.423096	100.000082	4487	2410
5	Point	0	181646.943481	2496358.753479	100.000015	2672	2410
6	Point	0	181661.732727	2496595.704712	100.000148	2410	2410
7	Point	0	181426.485107	2496841.439514	100.000126	4340	2973
8	Point	0	181179.408325	2496437.165283	100.000082	4487	4487
9	Point	0	181584.432129	2496800.192871	100.000151	2973	2973
10	Point	0	181183.375671	2496614.51947	99.999962	4487	4487
11	Point	0	181453.843872	2496779.809082	99.999943	4340	2973
12	Point	0	181348.306274	2496655.192505	100.000071	4340	4340
13	Point	0	181267.240479	2496642.887085	100.000066	4487	4340
14	Point	0	181593.870911	2496446.482117	100.000145	2410	2410
15	Point	0	181336.518494	2496394.864075	100.000061	4487	4487
16	Point	0	181297.962891	2496331.661316	100.00006	4487	4487
17	Point	0	181295.032288	2496600.604492	100.000086	4487	4340
18	Point	0	181561.893311	2496725.833496	100.00014	2973	2973
19	Point	0	181504.073303	2496831.605713	100.000109	2973	2973
20	Point	0	181551.445923	2496792.058289	100.000107	2973	2973
21	Point	0	181653.968689	2496769.768494	100.000101	2973	2973
22	Point	0	181544.072693	2496716.959717	99.999985	2973	2973
23	Point	0	181329.345703	2496643.838074	100.000071	4487	4340
24	Point	0	181235.725098	2496858.056519	100.000012	4340	4340
25	Point	0	181369.013306	2496550.710083	100.000142	4487	4340
26	Point	0	181231.929871	2496906.548279	100.000017	4340	2815
27	Point	0	181262.650879	2496837.663879	100.000106	4340	4340
28	Point	0	181732.29248	2496780.708923	100.000076	2973	2973
29	Point	0	181384.091309	2496739.974304	100.000044	4340	4340
30	Point	0	181746.766479	2496603.608704	100.000097	2410	2410
31	Point	0	181318.077515	2496394.609924	100.000184	4487	4487
32	Point	0	181134.170715	2496596.395691	100.000068	4487	4487
33	Point	0	181421.105103	2496394.892517	100.000137	4487	2410
34	Point	0	181437.06189	2496853.721313	100.000052	4340	2973
35	Point	0	181459.876099	2496649.524902	99.999995	4340	4340

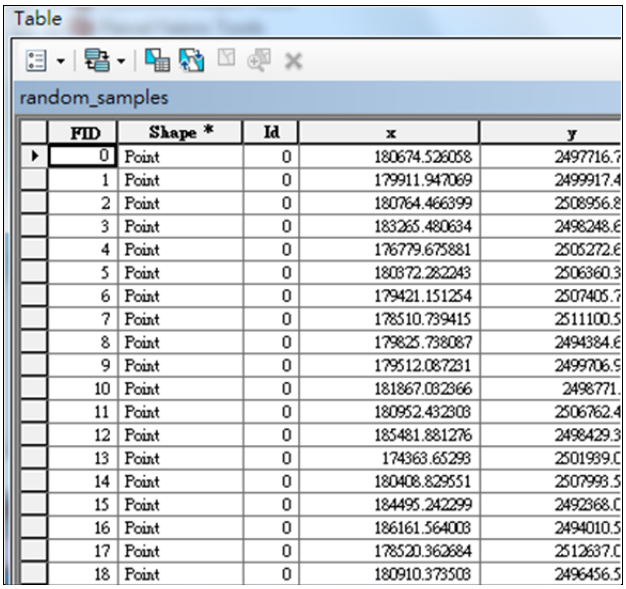
1 (0 out of 80 Selected)

affine100_60

Figure 38 Attribute table of random_samples.shp.

7.6 Fast Report

The following section offer example of Report method tool box of GeoMasker. Users can, on their own, follow the steps outlined in example to better understand the GeoMasker process and the interaction of the components described in the User's Guide. One method is demonstrated with the same example original point data, "random_samples.shp", and the same base polygon, "kaocity.shp". Two kinds of example data are shown as following:



Table

random_samples

FID	Shape *	Id	x	y
0	Point	0	180674.526058	2497716.7
1	Point	0	179911.947069	2499917.4
2	Point	0	180764.466399	2508956.8
3	Point	0	183265.480634	2498248.6
4	Point	0	176779.675881	2505272.6
5	Point	0	180372.282243	2506360.3
6	Point	0	179421.151254	2507405.7
7	Point	0	178510.739415	2511100.5
8	Point	0	179825.738087	2494384.6
9	Point	0	179512.087231	2499706.9
10	Point	0	181867.032366	2498771.
11	Point	0	180952.432303	2506762.4
12	Point	0	185481.881276	2498429.3
13	Point	0	174363.65293	2501939.0
14	Point	0	180408.829551	2507993.5
15	Point	0	184495.242299	2492368.0
16	Point	0	186161.564003	2494010.5
17	Point	0	178520.362684	2512637.0
18	Point	0	180910.373503	2496456.5

Figure 39 Attribute table of random_samples.shp.


Table

kaocity

FID	Shape *	COUNTYSN	COUNTYNAME
0	Polygon	64000001	高雄市

Figure 40 Attribute table of kaocity.shp.

Different settings of one method is shown as following table, after setting the parameters user can push the bottom "OK" to run the GeoMasker.

Fast Report method
original points: random_samples.shp **new points:** **base polygon:** village.shp **output feature:** report1.shp

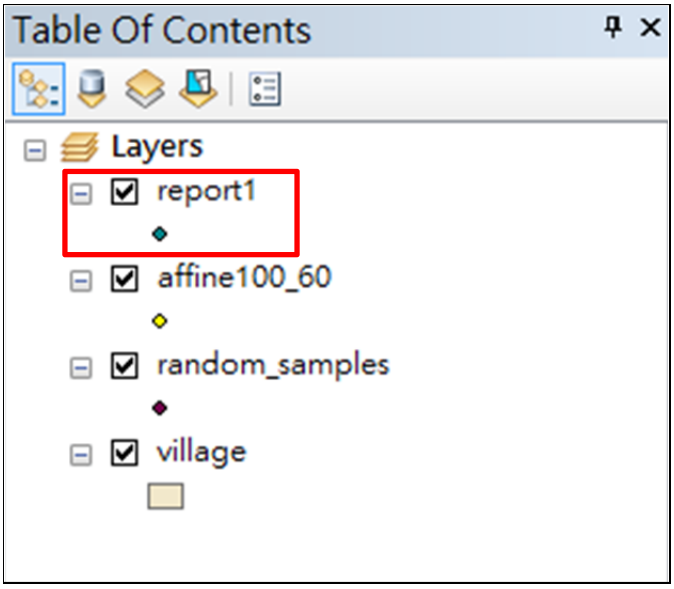


Figure 41 Table of Contents report1.shp.

The output is a new point: report1.shp, focus in the Attribute table, Figure 42, the correspond to to the field data name of the new point multi-an increase of ”_1”. For example, density and density_1, represent the point of origin and the point corresponding to the population density values. There many corresponds to the field data, not just users only from a single field view GeoMask results.

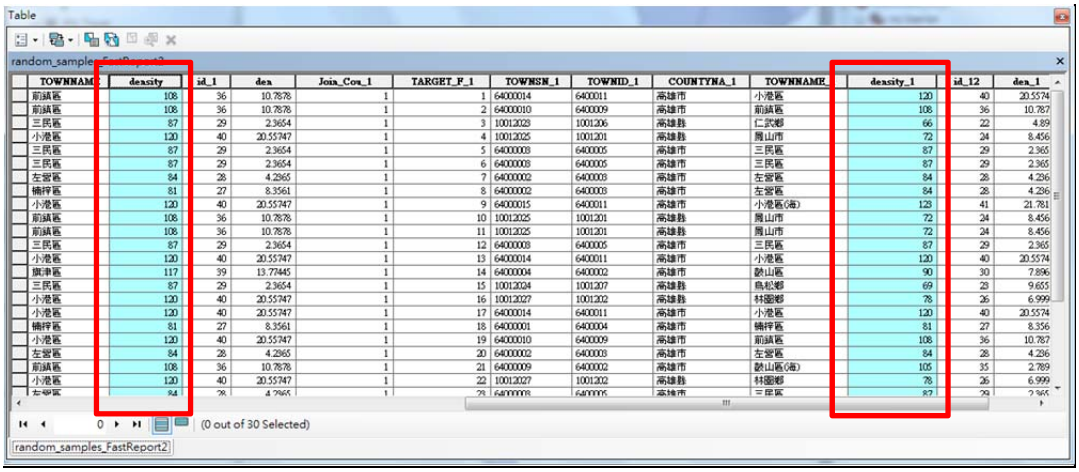


Figure 42 Attribute table of report1.shp.