A toolbox for analyzing the effect of infra-structural facilities on the distribution of activity points

Tohru YOSHIKAWA
Department of Architecture
Tokyo Metropolitan University

Atsuyuki OKABE
Center for Spatial Information Science
University of Tokyo

Thank you chairman.

The title of our presentation is:

A toolbox for analyzing the effect of infra-structural facilities on the distribution of activity points.

This study was carried out by Yoshikawa and Okabe,

and I, Yoshikawa, will talk.

The objective of this study is to develop a toolbox for examining the effect of infrastructural facilities on the distribution of point-like facilities.

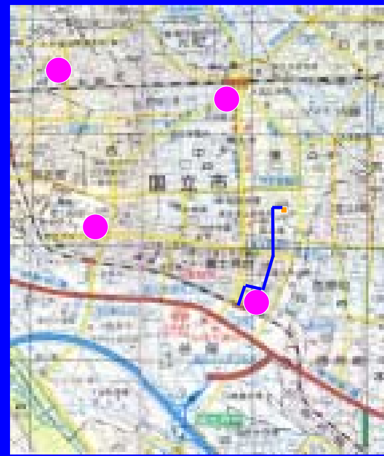
In the real world, we observe various kinds of point-like facilities whose distribution is affected by infrastructural facilities.

For example, the distribution of fast-food stores is likely to be affected by the location of stations.

Our toolbox is useful for examining this kind of effect statistically.

Railway Station :
Infra-structural Facilities

Fast-food Shop:
Point-like Facilities



Let me show you an actual example.

This map shows the distribution of “McDonald’s” Hamburger Shops in My hometown. Three of the four shops are located near the railway stations.

Probably, the owners of these shops determined their locations so as to catch passengers getting on and off at these stations.

As a result Mac shops tend to gather around the stations.

Our toolbox can examine this tendency statistically.

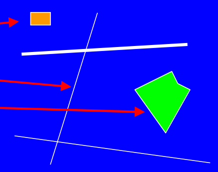
Infra-structural facilities:

- **Classified by shape**

  Point-like: Railway Stations;

  Line-like: Arterial Streets;

  Polygon-like: Large Parks.



- **Last over long time**

We classify infra-structural facilities into three types by their shape:

Point-like facilities, such as railway stations;

Line-like facilities, such as arterial streets;

and Polygon-like facilities, such as large parks.

The infra-structural facilities last over long time, say, fifty years.

Point-like facilities
Called Activity Points

- **Less stable:**

  e.g. One McDonald’s shop frequently appears and disappears.

- **Smaller size:**

  A McDonald’s shop is smaller than a station.

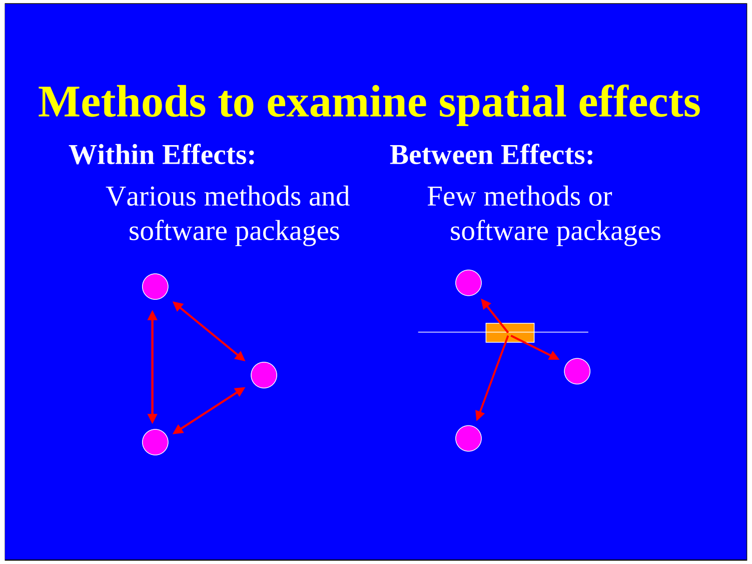
Compared with infra-structural facilities, stores, such as fast-food shops tend to appear and disappear in a fairly short time period,

and their sizes are much smaller than those of infra-structural facilities.

For instance, a McDonald’s shop appears or disappears every day in Japan,

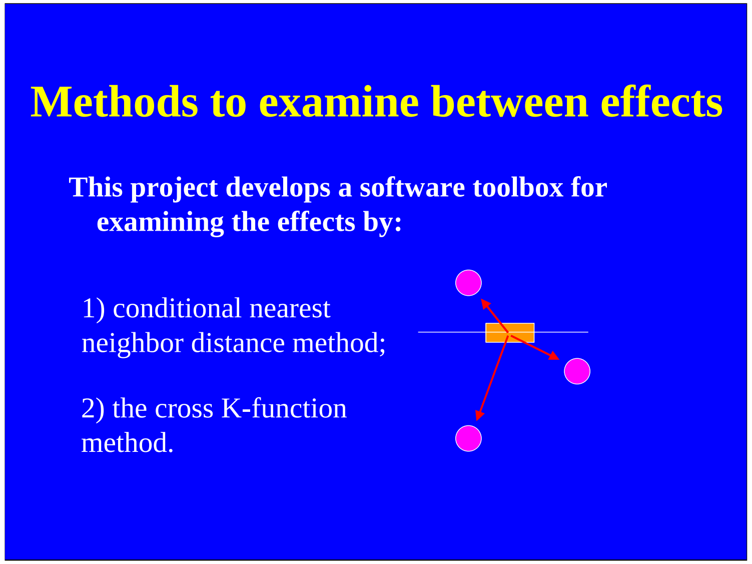
and the store is much smaller than a station in size.

Thus we may regard them as points and we call the points activity points.



In the literature, we can find many methods and software packages that examine spatial effects within the same kind of facilities.

However, we can find few methods and software packages that examine spatial effects between different kinds of facilities, say, infra-structural facilities and activity points.



We have developed a toolbox to deal with the between effect of infra-structural facilities on the distribution of activity points.

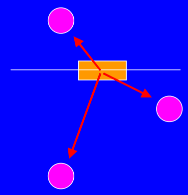
The toolbox consists of two tools, namely:

a tool for the conditional nearest neighbor distance method proposed by Okabe and Miki;

and a tool for the cross K-function method proposed by Ripley.

Toolbox

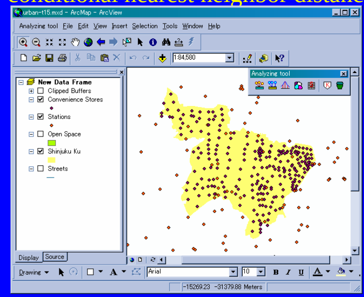
- Written in Macros with VB
- Based on ArcView 8.2
- “Analyzing Tool” on ArcView menu bar



This toolbox is written in Macros with VB and runs on ArcView 8.2. We add an additional menu item to the menu bar on ArcView window, called the “Analyzing Tool” menu.

Distances from activity points to the nearest infra-structural facilities

conditional nearest neighbor distance method -1



Let us demonstrate how to use this tool.

First we show the procedure for using the conditional nearest neighbor distance method.

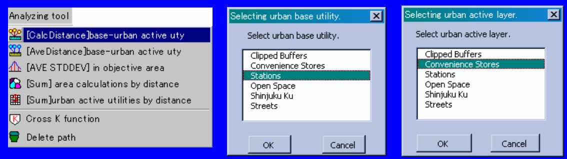
The orange points of this map show the location of railway stations, and the purple points show the location of convenience stores in Shinjuku-Ku, Tokyo Metropolis.

You might expect that these stores tend to locate near the stations because many pedestrians gather around the stations.

If this expectation is true, the distances from these stores to their nearest stations are smaller than the distances from randomly distributed stores to their nearest stations.

Distances from activity points to the nearest infra-structural facilities

conditional nearest neighbor distance method -1



The distribution of the distances is obtained using the [Calc Distance] submenu of the “Analyzing Tool” menu.

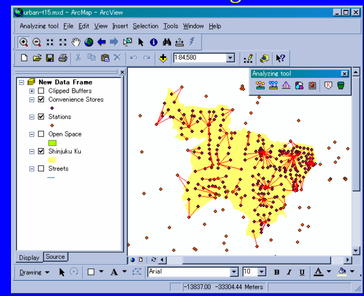
We select this submenu, then two dialog boxes appear.

As you see in these boxes, we can select the layer of infra-structural facilities and that of activity points.

Suppose we choose the layers of the stations and the convenience stores and click this OK button.

Distances from activity points to the nearest infra-structural facilities

conditional nearest neighbor distance method -1



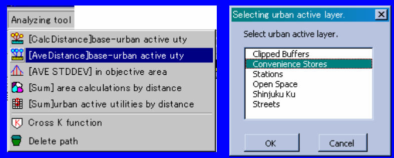
Then, as you see in this picture, we obtain the distances.

This map shows the nearest station from each convenience store.

The values of the distances are attached to the table of the activity point layer.

Average distance from activity points to the nearest infra-structural facilities

conditional nearest neighbor distance method -2



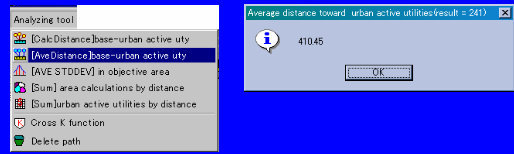
We can also obtain the average distance from all activity points to their nearest infra-structural facilities.

Let me show how to do it.

We select the [AveDistance] submenu, then this dialog box appears. As is seen in the box, we can choose the activity point layer that contains the distances. Let us select the convenience stores layer and click the OK button.

Average distance from activity points to the nearest infra-structural facilities

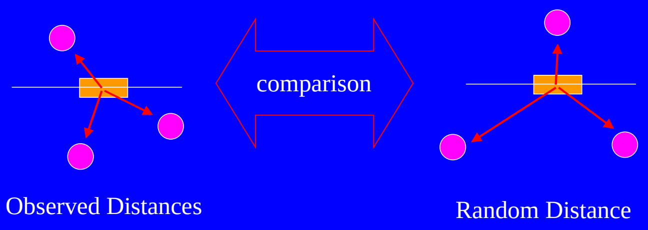
conditional nearest neighbor distance method -2



Then we can also obtain the average distance from all activity points to their nearest infra-structural facilities.

Distribution of the distance from a random point to its nearest infra-structural facilities

conditional nearest neighbor distance method -3

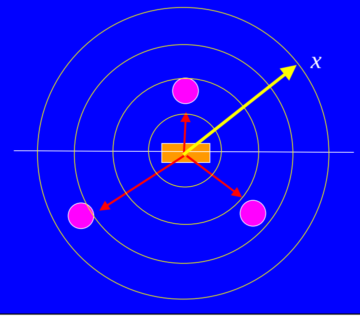


In order to test whether the observed distances are significantly smaller or larger than the distances from randomly distributed stores,

we have to obtain the probability distribution function of the random distances.

Distribution of the distance from a random point to its nearest infra-structural facilities

conditional nearest neighbor distance method -3



This figure shows how to obtain this function.

When a point is randomly placed, the probability of the point being placed in a circle centered at the station with the radius x is proportional to the area of the circle.

Distribution of the distance from a random point to its nearest infra-structural facilities

conditional nearest neighbor distance method -3

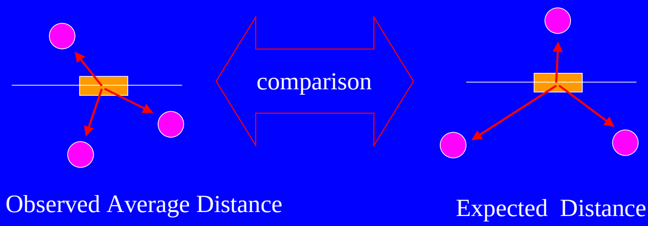
We can calculate this area using the “Buffers” menu of ArcView.

In theory, the distribution function is continuous with respect to the radius, but in practice, we use discrete values.

Obviously if we use small intervals, the computational time increases.

Comparison between the observed average distance and the expected distance

conditional nearest neighbor distance method -4

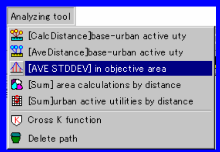


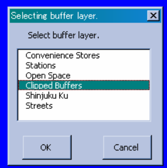
Once we obtain the distribution function of the random distances, we can test the null hypothesis that the stores are randomly located.

One of the methods is to compare the observed average distance and the expected random distance.

Comparison between the observed average distance and the expected distance

conditional nearest neighbor distance method -4



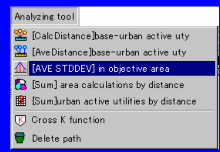


We can obtain the expected random distance from the buffers using the [Ave StDev] submenu of "Analyzing Tool" menu. The tool also reports the standard deviation of the random distances.

When we select the [Ave StDev] submenu, this dialog box appears. We can choose the layer of the buffers around the stations. Then click the OK button.

Comparison between the observed average distance and the expected distance

conditional nearest neighbor distance method -4



The tool reports the expected distance and the standard deviation of the random distances.

Comparison between the observed average distance and the expected distance

conditional nearest neighbor distance method -4

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

The null hypothesis can be tested using z.

By comparing the observed average distance with the expected distance, we can test the null hypothesis that the stations do not affect the location of the convenience stores. We can test the null hypothesis using the value of z given by this equation. If the null hypothesis is true, the distribution of z is the standard normal distribution.

Comparison between the observed average distance and the expected distance

conditional nearest neighbor distance method -4

$$z = \frac{410.45 - 445.67}{\frac{214.03}{\sqrt{241}}} = -2.554$$

In the case of Shinjuku-Ku, z is -2.554 and the null hypothesis is rejected with the significance level of 95%. Thus we can conclude that the convenience stores are located near the stations. This result implies that the stations attract the stores.

Goodness-of-fit test

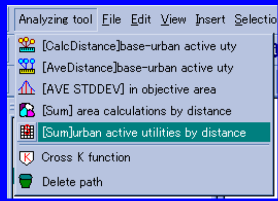
conditional nearest neighbor distance method -4

distance	0-150	-300	-450	-600	-750	-900	900-
stores							
probability							

We can alternatively test the null hypothesis by the goodness-of-fit test. To conduct the test, we divide the continuous distance into several discrete intervals, and count the number of stores in each interval.

Goodness-of-fit test

conditional nearest neighbor distance method -4

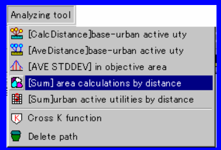


0,150,17
150,300,70
300,450,56
450,600,51
600,750,33
750,900,11
900,1050,3
1050,1200,0
1200,1350,0
1350,1500,0

This count is done by the “[SUM] urban active utilities by distance” submenu.

Goodness-of-fit test

conditional nearest neighbor distance method -4



0,150,S(1),1470535.851
150,300,S(2),3588605.740
300,450,S(3),4568936.024
450,600,S(4),4057118.575
600,750,S(5),2794822.691
750,900,S(6),1403499.654
900,1050,S(7),358484.667
1050,1200,S(8),5399.588
1200,1350,S(9),0
1350,1500,S(10),0

We can also calculate the probabilities from the areas within the distance intervals.

This calculation is done by the “[SUM] area calculations by distance” submenu.

Goodness-of-fit test

conditional nearest neighbor distance method -4

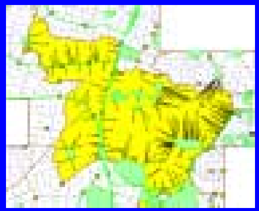
distance	0-150	-300	-450	-600	-750	-900	900-
stores	17	70	56	51	33	11	3
probability	.0806	.1967	.2504	.2223	.1532	.0769	.0199

P =0.01586

In the case of Shinjuku-Ku, the chi square probability calculated from the numbers of the stores is 0.01586. This means that the null hypothesis is rejected with the significance level of 95%.

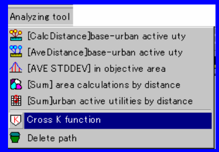
Line-like and polygon-like infra-structural facilities

conditional nearest neighbor distance method -5



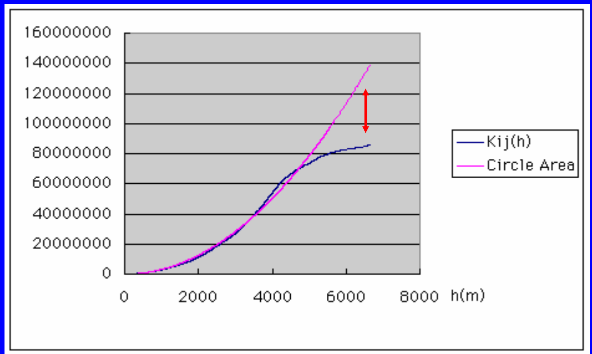
Our toolbox can also deal with line-like infra-structural facilities such as arterial streets and polygon-like ones such as parks in the same way using this toolbox.

Cross *K*-function method



When we expect that stores are affected not only by its nearest station but also the second, third nearest stations and so on, the cross k-function method is suitable for the analysis. This toolbox also provides a submenu for estimating the cross k-function.

Cross *K*-function method



By processing the output file using a database or a spreadsheet software package like Microsoft Access or Excel, we can obtain a cross K-function.

Comparing the observed and expected functions indicated by the red and blue lines, we may say that the convenience stores tend to be fewer in the area far from the stations.

Conclusions

- Our toolbox provides a user-friendly tools for analyzing the effect of infra-structural facilities on activity points.
- Our toolbox can deal with three types of infra-structural facilities:
  - point-like facilities, such as stations;
  - line-like facilities, such as streets;
  - polygon-like facilities, such as parks.

As you might have understood from our presentation, our toolbox provides a user-friendly tool for analyzing the effect of infra-structural facilities on activity points.

The toolbox can deal with three types of infra-structural facilities, namely:

point-like facilities, such as stations;

line-like facilities, such as streets;

and polygon-like facilities, such as parks.

Conclusions

- Our toolbox will be open to non-profit users without charge.
- We are now preparing a web site for this toolbox.

This toolbox will be open to non-profit users without charge.

We are now preparing a web site for this toolbox like SANET.

If you think our tool is useful for your study, please visit our site and use our toolbox.

Thank you for your listening.