Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography, Vol. 31, No. 6-2, 601-609, 2013
http://dx.doi.org/**10.7848/ksgpc.2013.31.6-2.601**

Assessing the Appropriateness of the Spatial Distribution of Standard Lots Using the L-index

Sang-Kyeong Lee[1)] · Byoungkil Lee[2)]

Abstract

Standard lots, which are used to assess values of individual lots in Korea, have been criticized for their improper distribution. However, there has been very little evaluation for the spatial distribution of standard lots, and an evaluation method has never been developed. In order to overcome this situation, we attempt to assess the appropriateness of the spatial distribution of standard lots using the L-index and Monte Carlo simulation. The L-index is a well-known indicator of the complete spatial randomness (CSR) of points in spatial statistics. If the L-index of standard lots is similar to that of individual lots, the former is considered to be randomly distributed according to the latter. By analyzing L-indices of two study areas, Gangnam and Seongdong, we find a statistically significant difference in Gangnam area and a relatively small difference in Seongdong area. We confirm that the spatial distribution of standard lots is not CSR and that the L-index is useful as an evaluation method. These results suggest that the standard lot selection and management guidelines need to be modified to apply the spatial distribution of individual lots to the standard lot selection process.

Keywords : Standard Lot, Spatial Distribution, Complete Spatial Randomness, L-index

1. Introduction

Standard lots are selected parcels that are used as criteria assessing values of individual lots based on the official land value system, and 500,000 such lots have been allocated nationwide as of 2013. Values of them are assessed every January 1 by appraisers and used as data for land value index tables. They are legally managed under the standard lot selection and management guidelines, which set the general criteria for the allocation and selection of standard lots by region, zoning, and land use. In spite of several revisions, these guidelines have been criticized for vague criteria (Lee, 2012). Jeong and Hwang (2004), Park (2006) and Lee (2012) suggested specific criteria based on decision-making model and GIS, but they didn't cover the criteria of the spatial distribution of standard lots and assess appropriateness.

Standard lot selection should consider the individual lot distribution because official values of individual lots are determined from compared standard lots. From a statistical point of view, a proper distribution of standard lots suggests a high probability of concurrence with the spatial distribution of individual lots. This can be assessed by the L-index, which was invented by Ripley (1976) as a method for globally analyzing the distribution patterns of point entities using distances between a specific point and all other points. Diggle and Chetwynd (1991) and Gatrell et al. (1996) developed the L-index method by verifying the relationship between the outbreak of disease and environmentally harmful facilities. These previous studies suggest that the L-index can be used as the indicator of the complete spatial randomness (CSR) of points. If the L-index of standard lots is similar to that of individual lots, the former is considered

1) Department of Urban Planning, Gachon University, Republic of Korea (E-mail: skylee@gachon.ac.kr)
2) Corresponding Author, Department of Civil Engineering, Kyonggi University, Republic of Korea (E-mail: basil@kgu.ac.kr)

to be randomly distributed according to the latter. In this perception, we aim to assess the appropriateness of the spatial distribution of standard lots using the L-index and Monte Carlo simulation. We analyze L-indices of two study areas, Gangnam and Seongdong and assess the CSR of standard lots.

2. Research trends

2.1 Standard lot selection and management guidelines

Allocating standard lots to administrative units follows the standard lot allocation criteria. These criteria were completed in 1996 and the sample size of standard lots was defined differently by city size. Under the administrative classification, regions were classified as large cities, mid to small cities, and military areas. However, because standard lot density was uniformly applied with no consideration of local conditions, the first allocation criteria were criticized and required to be changed. In the 2003 revised criteria, regional and industrial factors were considered to allocate standard lots. In the 2008 revised criteria, land value level was considered by city, county, and borough. There were 150,000 standard lots in 1989 when the first official land value system was implemented, 300,000 in 1990, 450,000 in 1996, and 500,000 in 2008.

Selecting standard lots in a small area follows the standard lot selection criteria. These criteria are defined as follows: the representation of land values, the unbiasedness of land characteristics, the stability of land use, and the determination of land distinction. However, because these selection criteria are focused on declaring the principle of standard lot selection, there has been no attempt to pinpoint the location of standard lots or to specify spatial distribution.

2.2 Researches related to standard lot selection

Ko (1997) shows that when calculating the optimal number of standard lots based on statistical sampling, allowable error is exceedingly high and reliability is low. Jung (2002) reveals problems in standard lot selection criteria, including the lack of statistical sampling and the abstractness of principles. Jeong and Hwang (2004) suggest that selection of standard lots should represent land value level, and show a selection method based on GIS and spatial statistics. Related to the redesign of the standard lot distribution, Ju, Ahn and Kwon (2010) suggest easing the imbalance of standard lot density by decreasing the number of standard lots in large city areas, residential areas, and commercial areas, while increasing the number of standard lots in natural environment conservation areas. Problems and improvement plans brought up in this previous research are generally applied and used in many of the revision processes for the standard lot selection and management guidelines. However, despite these revisions, the criteria for standard lot selection and spatial distribution remain vague. There is much research on standard lot selection method using GIS to overcome possible limits, including the following: Kim (1995), Lee (1999), Jeong and Hwang (2004), Park (2006), Yun (2007), and Lee (2012). However, because such research using GIS are all focused on improving the standard lot selection criteria, there has been no attempt to assess the spatial distribution of standard lots statistically, as this study attempts to do.

2.3 Point pattern analysis method

Point patterns are defined as a series of locations ($s_1, s_2, \cdots$) where for a specific space, *R*, s_i is a vector coordinate of i^{th} event in *R*. Event is standard term used in the point process to distinguish the observed location from a random location in *R* (Diggle, 1983). *R* can be a rectangle or polygon, and because a border effect occurs in any case, a buffer zone should be used. Otherwise, the border must be considered at the time of analysis, to supplement the effect.

The simplest statistical model for point patterns in space is Complete Spatial Randomness (CSR). CSR means the event is independently distributed according to the equal statistical distribution in the target area *R*. If point patterns clearly show the features of aggregation or regular distribution, they can be visually identified. If it is not visually possible to identify either random distribution or aggregation and regular distribution, point patterns should be statistically tested against a null hypothesis of CSR. However, independent events with concentration may be a case of no concentration, when they are compared with their background or environment (Lee, 2008). For example, regarding the standard lots and individual lots in this study, if standard lots are extracted from individual lots, this means the standard lots are distributed in the CSR according to the individual lots. Here, if individual lots are aggregated, the standard lots selected from those individual lots are also aggregated. Therefore, whether or not the distribution of standard lots is in statistically significant CSR should be considered and assessed in comparison to the distribution of individual lots.

The index most used to verify the CSR distribution of point patterns is the K-index derived from Ripley's K-function. The K-index compares the number of actual points existing within a certain distance and those theoretically expected to determine whether a point distribution is random (Eq. (1)).

$$\widehat{K}(r) = \frac{R}{n^2}\sum\sum\frac{I_h(d_{ij})}{w_{ij}} = \frac{1}{\lambda^2 R}\sum\sum\frac{I_h(d_{ij})}{w_{ij}} \tag{1}$$

Here, r is search radius, R is the target area, n is the number of events, $\lambda\ (=\frac{n}{R})$ is the density of the event, and d_{ij} is the distance between the event s_i and s_j. $I_h()$ is the indicator function (if $d_{ij} < r$, 1, otherwise, 0), and w_{ij} is a weight that can eliminate border effects.

If the point distribution is CSR, a point's probability of occurring in every location is identical and independent from other points. Therefore, the average number of points expected to be found within a distance r from a certain location is $\lambda\pi r^2$. In an isotropic situation without spatial interaction, it is $K(r) = \pi r^2$, in aggregated distribution $K(r) > \pi r^2$, and in regular distribution $K(r) < \pi r^2$.

Therefore, in the K-index graph for distance, it is possible to determine at which distance the aggregation occurs, but because the general graph increases by r, the L-index, shown in Eq. (2), is used in practical applications instead of the K-index.

$$L(r) = \sqrt{\frac{K(r)}{\pi}} - r \tag{2}$$

Eq. (2) gives the L-index, originally suggested by Ripley (1976), which was supplemented by Cressie (1991) for advantages stemming from the concurrence between the L-index for CSR distribution and the X-axis on the graph through distance deduction. If two different point patterns are separately random, both L-indices become zero.

Significance verification of the L-index's difference from zero, indicating whether points are concentrated or regularly distributed compared to CSR distribution, principally uses the Monte Carlo simulation (Diggle, 1983). This method arbitrarily generates the same number of point objects as the target event and calculates the L-index many times over to find the minimum and maximum bounds and the statistical confidence interval.

3. Testing and analysis

3.1 Data and software

The data used in this study are the coordinates of individual lots and standard lots in Seoul in 2010. This study selects the Gangnam-gu as study area #1, as it has regularity in individual lot distribution due to new town development, and the Seongdong-gu as study area #2, which includes varied lots including streams, old neighborhoods, Eungbong Mountain, and the Seoul Forest.

As shown in Fig. 2, study area #1 shows a regular distribution of individual lots and standard lots, and study area #2 shows very irregularly distributed and concentrated individual lots and standard lots.

Fig. 1. Location of study areas

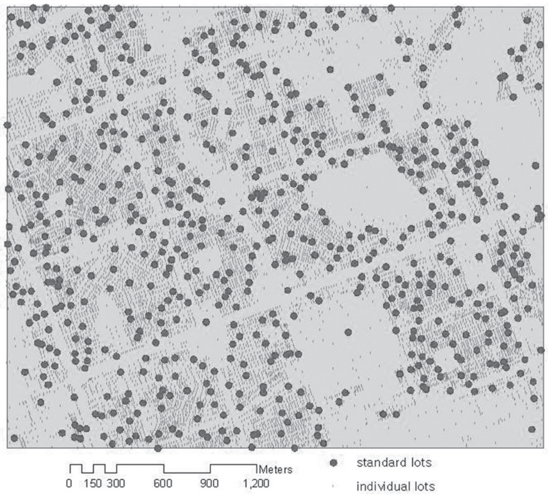

Fig. 2. Study area #1 (Gangnam area)

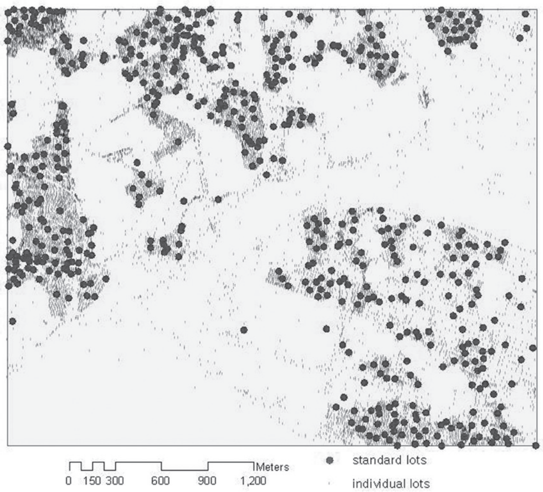

Fig. 3. Study area #2 (Seongdong area)

The number of individual lots and standard lots by zone within study area #1 are given in table 1, the occupied area in table 2, and the average area in table 3. Here, study area #1 has a small deviation in land size, and the extracted standard lot ratio is a little high in the semi-residential zone and in all zones combined. This results from following the standard lot selection guidelines. Study area #2 instead shows large deviations in size by use, and the extracted standard lot ratio is much higher than general commercial areas. This results from following standard lot selection criteria, but with relatively larger ratios than other zones.

Table 1. Numbers of lots by zones (Study area #1)

	Individual lots		Standard lots		b/a
	Numbers (a)	Ratio	Numbers (b)	Ratio	
1st exclusive residential zone	736	4.6%	18	2.7%	2.4%
1st general residential zone	1357	8.5%	39	6.0%	2.9%
2nd general residential zone	6778	42.5%	215	32.8%	3.2%
3rd general residential zone	4601	28.9%	225	34.4%	4.9%
Semi-residential zone	226	1.4%	14	2.1%	6.2%
General commercial zone	2236	14.0%	144	22.0%	6.4%
Total	15934	100.0%	655	100.0%	4.1%

Table 2. Sum of lot areas by zones (Study area #1)

	Individual lots		Standard lots		b/a
	Sum of area (a)	Ratio	Sum of area (b)	Ratio	
1st exclusive residential zone	282313.0	3.3%	6919.1	2.0%	2.5%
1st general residential zone	1023148.0	12.0%	13509.7	3.9%	1.3%
2nd general residential zone	2690857.6	31.5%	90943.9	26.1%	3.4%
3rd general residential zone	2759842.9	32.3%	135536.5	38.9%	4.9%
Semi-residential zone	112567.6	1.3%	6175.6	1.8%	5.5%
General commercial zone	1673369.7	19.6%	95622.1	27.4%	5.7%
Total	8542098.8	100.0%	348706.9	100.0%	4.1%

Table 3. Average area of lots by zones (Study area #1)

	Individual lots	Standard lots
	Mean area of lot (㎡)	Mean area of lot (㎡)
1st exclusive residential zone	383.6	384.4
1st general residential zone	754.0	346.4
2nd general residential zone	397.0	423.0
3rd general residential zone	599.8	602.4
Semi-residential zone	498.1	441.1
General commercial zone	748.4	664.0
Total	536.1	532.4

Table 4. Numbers of lots by zones (Study area #2)

	Individual lots		Standard lots		b/a
	Numbers (a)	Ratio	Numbers (b)	Ratio	
1st general residential zone	990	5.2%	19	3.5%	1.9%
2nd general residential zone	10499	54.6%	272	50.5%	2.6%
3rd general residential zone	3212	16.7%	80	14.8%	2.5%
Semi-residential zone	715	3.7%	31	5.8%	4.3%
General commercial zone	230	1.2%	16	3.0%	7.0%
Neighboring commercial zone	76	0.4%	2	0.4%	2.6%
Semi-industrial zone	3045	15.8%	117	21.7%	3.8%
Natural green zone	448	2.3%	2	0.4%	0.4%
Total	19215	100.0%	539	100.0%	2.8%

Table 5. Sum of lot areas by zones (Study area #2)

	Individual lots		Standard lots		b/a
	Sum of area (a)	Ratio	Sum of area (b)	Ratio	
1st general residential zone	807053.3	11.0%	24584.1	8.6%	3.0%
2nd general residential zone	2461344.2	33.6%	59361.8	20.7%	2.4%
3rd general residential zone	1271198.2	17.4%	107170.8	37.4%	8.4%
Semi-residential zone	139180.8	1.9%	7328.2	2.6%	5.3%
General commercial zone	54399.4	0.7%	20468.5	7.1%	37.6%
Neighboring commercial zone	2862.1	0.0%	60.5	0.0%	2.1%
Semi-industrial zone	1366074.8	18.6%	58846.1	20.5%	4.3%
Natural green zone	1224243.5	16.7%	8587.0	3.0%	0.7%
Total	7326356.3	100.0%	286407.1	100.0%	3.9%

Table 6. Average area of lots by zones (Study area #2)

	Individual lots	Standard lots
	Mean area of lot(㎡)	Mean area of lot(㎡)
1st general residential zone	815.2	1293.9
2nd general residential zone	234.4	218.2
3rd general residential zone	395.8	1339.6
Semi-residential zone	194.7	236.4
General commercial zone	236.5	1279.3
Neighboring commercial zone	37.7	30.3
Semi-industrial zone	448.6	503.0
Natural green zone	2732.7	4293.5
Total	381.3	531.4

In comparing and analyzing spatial distributions for standard lots and individual lots, standard lots are considered incidents, and individual lots are considered environment, and then compared by the calculation of the L-index of two spatial features. For the subset, data from the study area ArcMap was used, and to calculate the L-index, CrimeStat Ⅲ, specialized and used for spatial statistics analysis, and R, a statistics program were used.

3.2 L-index analysis and considerations

To verify the statistical significance of distributed standard lots in the study area, the Monte Carlo simulation is necessary. The CrimeStat Ⅲ program includes a simulation function for significance verification, but this function assumes a homogeneous space to create points randomly, while distributions of standard lots and individual lots are imbalanced because of geographical and land use features. Therefore, they would appear to be concentrated distributions compared to randomly created points.

Thus, we execute a Monte Carlo simulation through 10 random samplings from individual lots, which are the population of standard lots, to verify the statistical significance of standard lot distribution. In the sampling, R, which is a statistical program, is used frequently, and the L-index is calculated from sampled individual lots for upper and lower bounds to calculate a 95% confidence interval.

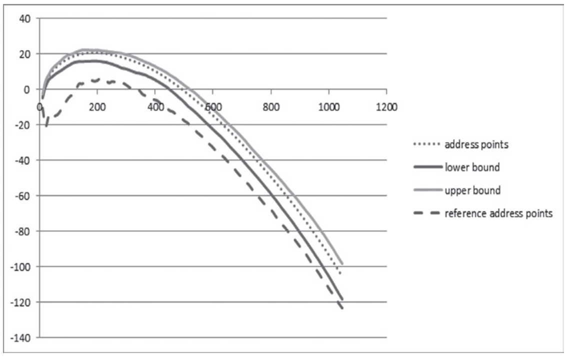

Fig. 4. L-indices of study area #1

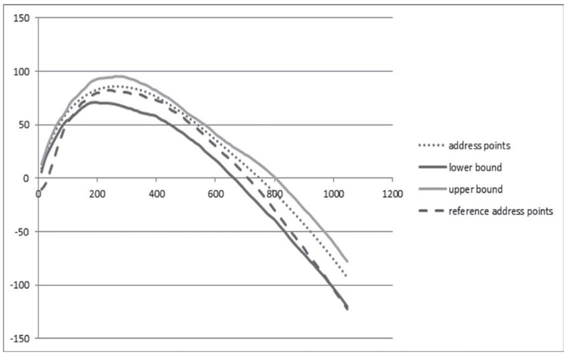

Fig. 5. L-indices of study area #2

Table 7. L-indices and confidence interval (Study area #1)

r(m)	Individual lots	Standard lots	Lower bound	Upper bound	Average	Standard deviation	95% confidence interval	
10	-3.84	-10.42	-4.80	-2.60	-3.84	0.81	-4.34	-3.33
105	17.9	-3.96	13.91	19.7	17.5	1.7	16.45	18.55
209	20.61	6.2	15.87	21.94	19.35	1.9	18.18	20.53
419	8.58	-7.97	3.62	11.41	6.83	3.06	4.93	8.72
523	-3.87	-20.88	-10.73	-0.29	-6.08	3.85	-8.46	-3.69
628	-18.89	-36.94	-26.72	-15.32	-21.29	4.45	-24.05	-18.53
733	-36.62	-55.68	-45.63	-33.13	-39.01	4.83	-42.01	-36.02
837	-56.53	-75.93	-66.79	-51.69	-58.91	5.33	-62.21	-55.61
942	-79.2	-98.29	-90.47	-72.52	-82.05	6.02	-85.77	-78.32
1047	-105.48	-123.26	-118.22	-98.18	-108.85	6.93	-113.14	-104.55

Table 8. L-indices and confidence interval (Study area #2)

r(m)	Individual lots	Standard lots	Lower bound	Upper bound	Average	Standard deviation	95% confidence interval	
10	9.15	-10.46	5.75	12.90	9.23	2.52	7.67	10.80
105	63.72	54.2	54.94	68.57	62.08	4.6	59.23	64.93
136	72.44	66.42	62.87	77.96	70.6	5.49	67.19	74
147	**74.76**	**70.32**	**65.52**	**81.46**	**73.11**	**5.52**	**69.69**	**76.53**
209	**83.82**	**80**	**70.28**	**93.46**	**81.36**	**7.58**	**76.66**	**86.06**
314	**84.49**	**79.06**	**65.01**	**92.73**	**79.18**	**9.08**	**73.55**	**84.8**
419	**73.25**	**71.17**	**54.45**	**78.56**	**66.75**	**8.71**	**61.35**	**72.15**
523	**53.66**	**49.45**	**35.19**	**57.43**	**46.86**	**8.69**	**41.48**	**52.25**
628	**29.41**	**22.54**	**9.56**	**35.65**	**22.94**	**9.4**	**17.11**	**28.76**
733	**4.18**	**-6.94**	**-20.34**	**16.36**	**-2.5**	**11.87**	**-9.85**	**4.86**
785	**-9.44**	**-24.19**	**-35.48**	**4.69**	**-16.5**	**13.19**	**-24.68**	**-8.32**
796	-12.31	-28.19	-37.59	2.28	-19.23	13.26	-27.45	-11.02
838	-24.22	-43.5	-50.85	-9.75	-30.91	14.13	-39.67	-22.15
942	-56.38	-81.37	-83.88	-41.43	-63.37	15.19	-72.79	-53.95
1047	-93	-122.37	-120.21	-77.81	-100.49	15.96	-110.38	-90.6

In the study area #1 for Fig. 4 and Table 7, the distribution of individual lots is concentrated within a 200m radius with a very low value; further, if it exceeds a 500m radius the distribution is regular. This explains how, as shown in Fig. 2, since the Gangnam area is planned, lots in the block show a slightly concentrated phenomenon, and the regularly arranged blocks are explained by the L-index. On the other hand, in study area #2 for Fig. 5 and Table 8, the distribution of individual lots is concentrated within a 300m radius, and is regular if it exceeds a 750m radius; further, the distribution of concentrated lot numbers in the old neighborhood shows a segmented form following natural geography.

In both study areas, the distribution of standard lots in all sections shows relatively lower values than the distribution of individual lots. Compared to a CSR distribution in which the overall standard lot selections are randomly extracted from individual lots, this result is considered to ensure collateral regularity in standard lot distribution from deliberated selection under the standard lot selection and management guidelines.

Study area #1 shows a negative L-index up to within a 100m radius, and so it is considered substantially regular for individual lot distribution, but its standard lots do not match the spatial distribution of individual lots, as it deviates from the 95% confidence interval in all sections under Monte Carlo simulation.

As the section between a 145m and 790m radius (the grey part of table 8) is included in the confidence interval in study area #2, we conclude that the standard lot distribution follows the individual lot distribution, in that extent of radius. However, the L-index is negative in the short distance within a 30m radius and the distribution of standard lots within a 200m radius shows a relatively lower density than that of individual lots. Because the overall number of standard lots is 2.8% for Seongdong, as compared to 4.1% for Gangnam, any future standard lot addition should apply the individual lot distribution.

4. Conclusions

In this study, we analyze L-indices to assess the appropriateness of the spatial distribution of standard lots. Study areas comprise Gangnam-gu, where the individual lot distribution shows the regularity of a newly built town, and Seongdong-gu, where there are diverse lot compositions around the Jungrangcheon River. Targeting these two areas, L-indices are calculated, and the analysis results are as follows. First, through the spatial distribution of individual lots, the effectiveness of the L-index was analyzed. For the Gangnam newly planned area, a small radius shows through a low L-index in that the individual lot distribution in the block is assessed as almost a CSR distribution. However, for Seongdong, the L-index value is relatively high, such that the individual lot distribution is concentrated because of natural geographical effects.

Second, the L-indices of standard lots and individual lots are compared and assessed. This assesses whether or not the spatial distribution of standard lots is adequate, or if it is a randomly extracted sample from individual lots, from a spatial viewpoint. In addition, through the Monte Carlo simulation using individual lots, the analysis shows that statistical assessment for standard lot distribution adequacy is possible.

Third, the result that the L-indices of standard lots in both Gangnam and Seongdong are low means that the spatial distribution of standard lots does not correspond to that of individual lots. The reason is that standard lots were not extracted from random sampling, but from standard lot selection criteria, which is an artificial standard. In particular, the L-index of standard lots in Gangnam is out of the lower bounds and the 95% confidence interval in all sections, having low significance as a sample.

In this study, the spatial distribution of standard lots appears not to fully apply to that of individual lots, and this means that standard lots are not representative samples. Considering the roles of standard lots as assessing criteria for official individual land values, the standard lot selection and management guidelines need to be modified to apply the spatial distribution of individual lots to the standard lot selection process. The limits of this study are as follows: bias according to zoning composition and land value distribution are not considered, and the concrete effects of the standard lot selection and management guidelines on standard lot spatial distribution are not considered.

References

Cressie, N., 1991, Statistics for Spatial Data, Wiley & Sons, Inc.

Diggle, P. J., 1983, Statistical Analysis of Spatial Point Patterns, Academic Press, London.

Diggle, P. J. and Chetwynd, A. G., 1991, Second-Order Analysis of Spatial Clustering for Inhomogeneous Population, Biometrics, 47, pp. 1155-1163.

Gatrell, A. C., Bailey, T. C., Diggle, P. J. and Rowlingson, B. S., 1996, Spatial Point Pattern Analysis and Its Application in Geographical Epidemiology, Transactions of the Institute of British Geographers, New Series, 21(1), pp. 256-274.

Jeong, S., and Hwang, C., 2004, The Selection of Standard Lands to Estimate the Public Land Price Based on Spatial Statistical Analysis, Journal of The Korea Planners Association, 39(3), pp. 55-67. (in Korean with English abstract)

Ju, Y., Ahn, J. and Kwon, H., 2010, Research on Improvement of the Publicly Notified Standard Land Price System and Redesign for Distribution, Korea Real Estate Research Institute. (in Korean)

Jung, S., 2002, Research on Improvement of the Standard Land Selection and Distribution Criteria, Real estate policy research, 3(1). (in Korean)

Kim, J., 1995, A Study on the Efficient Selecting of Government Announced Land Using Geographic Information System, MS Thesis of Hanyang University. (in Korean with English abstract)

Ko, S., 1997, Accuracy and Fairness of Publicly Notified Land Price Calculation-in Standard Land Management Aspects, Korean Journal of Policy Analysis and Evaluation 7(2), pp. 131-148. (in Korean)

Lee, B., 2008, Applying the L-index for Analyzing the Density of Point Features, The Journal of GIS Association of Korea, 16(2), pp. 237-247. (in Korean with English abstract)

Lee, H., 2012, Suitability for the Method of Selecting Standard Lot Based on Objective Indexes and GIS Spatial Analysis on the Official Land Price System, MS Thesis of Namseoul University (in Korean with English abstract)

Lee, Y., 1999, A Study on the Alternatives to Raising Propriety of the Public Notification of Land Values, MS Thesis of Chonnam National University. (in Korean with English abstract)

Park, J., 2006, Automation of Land Appraisal Value Using GIS, Ph. D. Dissertation, Kangwon National University. (in Korean with English abstract)

Ripley, B. D., 1976, The Second-Order Analysis of Stationary Point Processes, Journal of Applied Probability, 13, pp. 255-66.

Yun, H., 2007, A Study on the Improvement of the Efficiency of Duties of Selecting Standard Land Using GIS, MS Thesis of Chung-Ang University. (in Korean with English abstract)

(Received 2013. 12. 02, Revised 2013. 12. 05, Accepted 2013. 12. 28)