Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography
Vol. 34, No. 4, 403-411, 2016
http://dx.doi.org/10.7848/ksgpc.2016.34.4.403

ISSN 1598-4850(Print)
ISSN 2288-260X(Online)
Original article

Verifing Formation of Area of Influence of Subway Station through Land Value Distribution Analysis – Case Study on Seoul

Lee, Byoungkil[1] · Lee, Sangkyeong[2]

Abstract

This research has the purpose to develop a method to evaluate whether station's area of influence has been formed, and verify formation of the area of influence through empirical analysis of all subway stations in Seoul. First, we created buffers of 100m intervals from 100m to 1000m, based on subway station exits, calculated the average land price of each buffer, and divided station areas of influence into 10 clusters using K-means clustering with the average land prices as values of observation. Subsequently, we have assumed a decreasing price curve from increasing distance from a nearby subway station, estimated a price curve and evaluated whether the area of influence actually exists using regression analysis of each cluster. The 10 area of influence clusters were largely divided into strong, weak, and no area of influence of subway station. The stations where the strong areas of influence are formed are mainly located in center, sub-centers, and local centers; stations where weak and no areas of influence are formed are mostly located in the adjacent areas of center or sub-centers or suburbs.

Keywords : Subway Station, Area of Influence, Land Price, Land Price Gradient

1. Introduction

The areas of influence of subway stations generally signify a geographical range where a subway station exerts influence, and as these areas of influence become central places in terms of the spatial structure of metropolitan area, the interests of urban planners are large (Koo *et al*., 2016). In particular, as the 'Law on Development and Use of Station's Area of Influence' has been enacted and the development project of the areas of influence of subway stations have become an important means of urban land developments, overall interest has increased.

There has been a diverse amount of research conducted to recognize actual station areas of influence in a large city such as Seoul. There has been research that suggests method to define the range of station's area of influence (Kim *et al*., 2001, Kim *et al*., 2002; Kim *et al*., 2008; Kim and Lee, 2010; Lee *et al*., 2011; Kim *et al*., 2013), analyze spatial structures and land use characteristics inside the areas of influence (Choi *et al*., 2008; Park *et al*., 2009, Lee and Sohn, 2012), or classify the areas of influence according to spatial characteristics (Sung and Kim, 2005, Choi *et al*., 2008; Oh *et al*., 2009; Kim *et al*., 2013). In addition, there was also research conducted on what role the areas of influence play in terms of urban spatial structure (Koo *et al*., 2016).

These efforts helped define the areas of influence, but they also revealed a fundamental limit in the perspective of viewing the areas of influence. The preceding research efforts mostly assume that the areas of influence have already been formed, and this overlooked the fact that there may be stations without any areas of influence. Even though the areas nearby the station do not correspond to the concept of the area of influence, no preceding research handled this aspect. As a result, whether the areas of influence have actually been

Received 2016. 07. 23, Revised 2016. 08. 01, Accepted 2016. 08. 11
1) Member, Dept. of Civil Engineering, Kyonggi University (E-mail: basil@kgu.ac.kr)
2) Corresponding Author, Member, Dept. of Urban Planning, Gachon University (E-mail: skylee@gachon.ac.kr)

formed has not been an interesting subject as a research topic.

Recently, Seoul has been developing station areas of influence for urban revitalization, therefore verifying whether these areas of influence have actually been formed has increasingly become necessary. If the areas of influence did not form, one can suggest a development plan to strengthen the centrality of adjacent area of station in terms of transit oriented development.

Under these recognitions, this research develops the method to evaluate whether a station's area of influence has been formed, and verify formation of the area of influence through empirical analysis of all subway stations in Seoul. Specifically, we divide subway stations into several clusters using land price data nearby the stations, estimate the land price curve for each cluster, and evaluate whether the area of influence has been formed. As land price can be one of good proxy of land use, be acquired in parcel units, and have advantages in terms of utility, we use land price data rather than other data such as development density.

2. Literature Review

To evaluate whether an area of influence has been formed, firstly, we must select the range of area of influence. 'Law on Development of Station Areas of Influence' defines the area of influence as the station built according to 'Railroad Construction Act', 'Framework Act on Railroad Industry Development', and 'Urban Railroad Act' and its vicinity, without suggesting a specific range. Most of the research which selects the area of influence based on walking range suggests 400-600m (Kim *et al*., 2001; Park *et al*., 2009), whereas research analyzing land use characteristics of the area of influence usually suggest 500m as its range (Sung and Kim, 2005; Lee *et al*., 2011). However, Lee *et al*. (2011) has suggested 550m-700m through spatial characteristics analysis according to types, and Bae (1982) suggested 1km as the area of influence's range, considering 0.8-1.1km distances between stations, 0.8km walking distances, and administrative districts. From these research efforts, we can observe that the range will differ according to the station's characteristics, and setting the range to be as large as possible is necessary to minimize the research limitations.

Research which analyzed land use characteristics and classified the types of the areas of influence can provide implications on the evaluation method. Sung and Kim (2005), through a cluster analysis based on boarding and alighting amounts, categorized areas of influence and analyzed its relation to land use characteristics. Lee *et al*. (2011) categorized the areas of influence by land use characteristics and boarding and alighting characteristics, selected 50m interval bands from a station, and set the range of area of influence using 40 variables including land price and t-test for each band. Lee *et al*. (2011) categorized the areas of influence through cluster analysis and analyzed land use characteristics according to category. These research efforts show that analyzing the areas of influence through categorization is more effective than analyzing individually.

Research that used land prices as a means to analyze the spatial structures in an area of influence includes Kim *et al*. (2008) and Choi *et al*. (2008). Kim *et al*. (2008) and Choi *et al*. (2008) found the range of areas of influence by applying second-order polynomial and linear regression to land price data. These researches, however, show limits in theoretical approaches in land price equation and delineating the station's area of influence. Therefore, it is necessary to approach theoretically for estimating the land price curve from land price data.

3. Analysis Method

To evaluate whether the subway station's area of influence has been formed, this research conducts statistical analysis in two steps; cluster analysis in step 1, and regression analysis in step 2. Cluster analysis is done to classify the areas of influence using land price distribution nearby the subway station; we use K-means clustering method in this research. K-means clustering is a non-hierarchical cluster analysis method which is most commonly used in data mining. K-means clustering is a method that divides n observations into k clusters. Using the distances between k cluster centers and n observations, we assign the observations to nearby cluster centers, repeat calculations of cluster centers using assigned observations, and minimize the distance between all observations and allocated cluster centers. This method caters to most of data forms and is easy to apply as no special transformations are necessary, but it needs to preset the number of clusters, k (Bae and Roh, 2005).

Regression is used to estimate land price curve by each cluster. The land price curve which this research attempts to estimate is deduced through the following theoretical approach. Forming the area of influence is that facilities nearby a subway station, through interactions between facilities and the station, are becoming more active. If the subway station exerts a positive influence to the nearby facilities, the land nearer the station will has much higher price than the further land. As a result, the land price will decrease as the distance from the subway station increases.

The variation in land price with distance from a subway station can be summarized through the estimation of a land price gradient. A land price gradient is estimated by a simple regression model with land price as the dependent variable and distance as the independent variable. It is expressed as negative exponential, which takes the form:

$$P(d) = P_0 e^{-\alpha d} \quad (1)$$

Here, the land price at distance *d* meter from a subway station, *P(d)*, has two components. P_0 is the model's estimate of the level of land price at a subway station and α is the estimate of the coefficient on distance, which means the percentage reduction in land price with each unit increase in distance (i.e., each meter) from a subway station. In order to estimate Eq. (1), we transform it into a linear expression by taking the natural logarithms of both sides of the equation, which yields:

$$\log(P(d)) = \log(P_0) - \alpha d \quad (2)$$

Eq. (2) can be estimated by ordinary least squares regression, in which dependent variable is the log of land price, log(*P(d)*) and the independent variable is the land's distance, *d*.

4. Data Preparation

In this research, we have used the subway stations in the city of Seoul as of 2010, to evaluate whether station areas of influence have been formed by station. To analyze decreasing land prices according to increasing distance from a nearby subway station, we have used official land prices as of 2010; we have excluded parcels which are not the subject of trade such as road, river, ditch, park and railway. Fig. 1 shows the distribution of the parcels used for analysis.

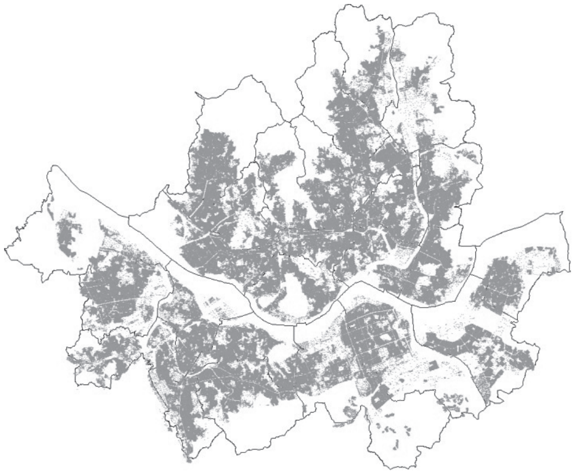

Fig. 1. Distribution of parcels used for analysis

We set the limit of the range of area of influence to effectively analyze land price distributions from subway stations. To minimize negative effects of the range selection, we want to set the limit of the range of area of influence as widely as possible, and set 1000m from the subway station exits as the limit, following Bae(1982)'s standards. We have created 100m interval buffers from 100m to 1000m from the subway station and calculated the average land price of each buffer. We have used ArcMap's spatial join function for calculating the average land prices. Fig. 2 shows parcels and buffer areas in Gangnam, Seoul. The reason why buffer shapes are not uniform is because we have used subway station exits as the starting points. We have chosen the exits as starting points because a station center cannot be set exactly.

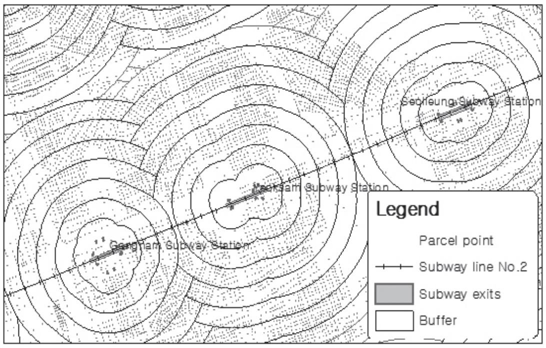

Fig. 2. Buffers and parcel points in the Gangnam area

5. Analysis Results

5.1 Results of K-means clustering

For step 1, we have conducted a cluster analysis for 241 station areas of influence, using a K-means clustering as the analysis method. We have used 10 buffers with 100m interval from the subway station as variables and used each buffer's average land price as values of observations. We have used 10 as the number of clusters, k, considering spatial structure of the city of Seoul: center, sub-center, local center, district center, and suburb. Table 1 shows the results classified subway stations of Seoul to 10 clusters.

Table 1. Classification of subway station areas of influence using K-means clustering

c1 N=55	Ahyeon, Balsan, Bangbae, Cheonggu, Cheongnyangni, Daebang, Daeheung, Dangsan, Dongguk Univ., Dongmyo, Dunchon-dong, Ewha Womans Univ., Gaerong, Gangbyeon, Gangdong-gu Office, Gangdong, Geoyeo, Gil-dong, Gongdeok, Gubeundari, Gwangheungchang, Hannam, Hapjeong, Hongik Univ., Hyochang Park, Isu, Itaewon, Jegi-dong, Konkuk Univ., Macheon, Mapo, Mok-dong, Mullae, Munjeong, Myeongil-dong, Naebang, Noksapyeong, Noryangjin, Ogeum, Omokgyo, Sadang, Sangsu, Seobinggo, Seongsu, Sinchon, Sindang, Singil, Sinseol-dong, Sookmyung Women's Univ., Ttukseom Resort, Ttukseom, Yangcheon-gu Office, Yeongdeungpo-gu Office, Yeongdeungpo Market, Yeongdeungpo
c2 N=9	Apgujeong, Daechi, Dogok, Euljiro 3(sam)-ga, Gangnam, Jongno 3(sam)-ga, Samseong, Seolleung, Yeoksam
c3 N=50	Aeogae, Amsa, Banghwa, Bokjeong, Bomun, Boramae, Chang-dong, Changsin, Children's Grand Park, Daerim, Dapsimni, Dongnimmun, Dorimcheon, Eungbong, Gaehwasan, Geumho, Gireum, Guil, Gunja, Guro, Gurogongdan, Gusan, Guui, Gwangnaru, Haengdang, Hangangjin, Hansung Univ., Hyehwa, Jangji, Jangseungbaegi, Junggye, Majang, Mangwon, Namguro, Nowon, Oksu, Sangil-dong, Sangwangsimni, Seoul Nat'l Univ., Sindaebang, Sindap, Sindorim, Singeumho, Sinjeong, Sinjeongnegeori, Sinpung, Ujangsan, Wangsimni, Yaksu, Yangpyeong
c4 N=22	Anguk, Banpo, Cheongdam, Euljiro 4(sa)-ga, Express Bus Terminal, Gangnam-gu Office, Hak-dong, Hangnyeoul, Ichon, Jamwon, Maebong, Nambu Bus Terminal, Nonhyeon, Samgakji, Seodaemun, Seokchon, Seongnae, Seoul Nat'l Univ. Of Education, Seoul, Sinsa, Songpa, Yeouknaru
c5 N=4	City Hall, Euljiro 1(il)-ga, Gwanghwamun, Jonggak
c6 N=5	Jamsil, Seocho, Sinyongsan, Sport Complex, Yongsan
c7 N=1	Gyeongbokgung
c8 N=3	Hoehyeon, Myeong-dong, Sincheon
c9 N=18	Bangi, Cheonho, Chungjeongno, Chungmuro, Daecheong, Dongdaemun Stadium, Dongdaemun, Dongjak, Garak Market, Godeok, Irwon, Jongno 5(o)-ga, Mongchiontoseong, Namyeong, Olympic Park, Suseo, Yangjae(Seocho-gu Office), Yeouido
c10 N=74	Achasan, Anam, Banghak, Beotigogae, Bongcheon, Bonghwasan, Bulgwang, Cheonwang, Danggogae, Dobong, Dobongsan, Dokbawi, Doksan, Dolgoji, Eungam, Gaebong, Garibong, Gimpo Int'l Airport, Gongneung, Gupabal, Hagye, Hankuk Univ. of Foreign Studies, Hanyang Univ., Hoegi, Hongje, Hwagok, Hwarangdae, Janghanpyeong, Jeungsan, Junggok, Junghwa, Kkachisan, Korea Univ., Madeul, Mapo-gu Office, Meokgol, Mia, Miasamgeori, Muakjae, Myeonmok, Nakseongdae, Namseong, Namyaeryeong, Nokbeon, Nokcheon, Onsu, Oryu-dong, Saejeol, Sagajeong, Sangbong, Sangdo, Sanggye, Sangwolgok, Seokgye, Seongbuk, Seongsan, Siheung, Sillim, Sindaebangsamgeori, Sinimun, Songjeong, Soongsil Univ., Ssangmun, Sungshin Women's Univ., Suraksan, Susaek, Suyu, Taereung, Wolgok, Wolgye, Yeokchon, Yeonsinnae, Yongdap, Yongmasan

5.2 Results of land price curve estimation

Following step 1 cluster analysis, we have conducted a simple regression for each of 10 clusters as step 2. We have used the distance from the subway station as the independent variable, measured in 100m interval, and the average land price of each buffer as the dependent variable. In Table 2, we present estimates of Eq. (2) for subway station areas of influence by cluster. For total data, the coefficient on distance α is 0.016 and is statistically significant at a 1% significance level. As R-squared value is only 0.7%, however, goodness of fit is very weak. This can also been observed through Fig. 3. As land price data are significantly biased to low value in close distance from subway stations, it worsens data fitting conditions. This result tells us that land price curve need to be estimated by cluster for better fittings.

Table 2 shows that the α values of c1, c2, c5, c6, c7, c9, and c10 match the assumptions of Eq. (1) with a statistical significance at 1% significance level. These provide estimates of slopes of land price gradients and signify that with each 100 meter increase in distance from a subway station, land price decreases by 1.0%, 4.3%, 10.6%, 9.7%, 11.6%, 5.0%, and 1.6%, respectively. The constant in this model is the estimate of log (P_0) of Eq. (2). To derive land prices at a subway station, we transformed the constant, log (P_0) into P_0 by taking exponential and thus, those are estimated at 4.1, 12.7, 28.1, 14.2, 19.5, 7.3, and 2.4 million won per square meter. C3 and c8 show α values that were not statistically significant, whereas the sign of c4's α value shows an opposite sign of that of Eq. (1), and α value is statistically significant at a 1% significance level.

From Fig. 4 to 13, we present scatter plots of land prices by distance from a subway station with the land price curve representing estimates of Eq. (2). The results from these figures are as follows;

C1 cluster shows low land prices in subway exits and low gradients, signifying weak an area of influence. In c2, land price at subway exits is remarkably high at KRW 12 million and shows high gradient, signaling a strong area of influence. C3 shows low land prices near exits and almost no gradient, which shows no area of influence. C4 shows high land prices near subway exits, but the prices rise as the distance from the station increases. This is characteristics of the adjacent area of center or sub-centers, which the area approaches to center or sub-centers as the distance from the station increases. C5 shows highest land prices near subway exits and gradient, showing characteristics of stations of center. C6 also shows high land prices and gradient. These stations are located at sub-centers or local centers. C7 is Gyeongbokgung Station. It is classified as a separate cluster from C2 or C5 but shows similar traits in terms of land price distributions with these clusters. C8 stations are located close to the center but do not show any characteristics of station areas of influence. C9 is located in district centers and shows a strong area of influence. C10 shows weak areas of influence and is an example of suburbs.

Table 2. Results of regression analysis by cluster

	Total	c1	c2	c3	c4	c5	c6	c7	c8	c9	c10
Constant	15.185	15.216	16.355	14.865	15.683	17.153	16.466	16.786	16.094	15.802	14.673
	(618.9*)	(762.1*)	(475.4*)	(822.5*)	(477.2*)	(313.7*)	(235.1*)	(103.8*)	(122.6*)	(382.7*)	(775.7*)
Coefficient on Distance	-0.016	-0.010	-0.043	-0.001	0.015	-0.106	-0.097	-0.116	-0.015	-0.050	-0.016
	(-4.1*)	(-3.1*)	(-7.7*)	(-0.4)	(2.8*)	(-12.0*)	(-8.6*)	(-4.4*)	(-0.7)	(-7.5*)	(-5.3*)
R-squared	0.007	0.017	0.403	0.000	0.034	0.792	0.604	0.712	0.017	0.239	0.038
Observations	2387	548	90	494	218	40	50	10	30	180	727
Land price on station exit	3.9	4.1	12.7	2.9	6.5	28.1	14.2	19.5	9.8	7.3	2.4
Gradient (%)	-1.6	-1.0	-4.3	-0.1	1.5	-10.6	-9.7	-11.6	-1.5	-5.0	-1.6

T-statistics are in parentheses and * is statistically significant at a 1% significant level.

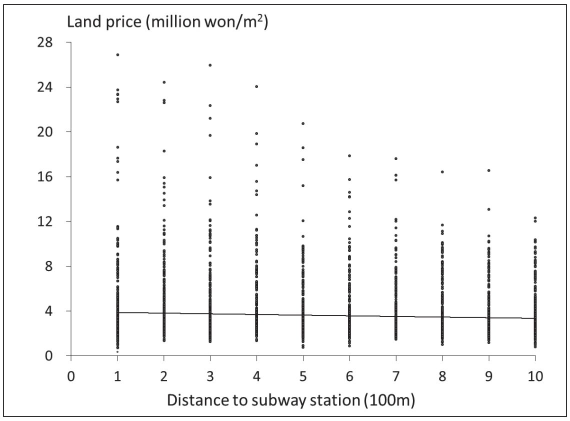

Fig. 3. Land price curve of Total

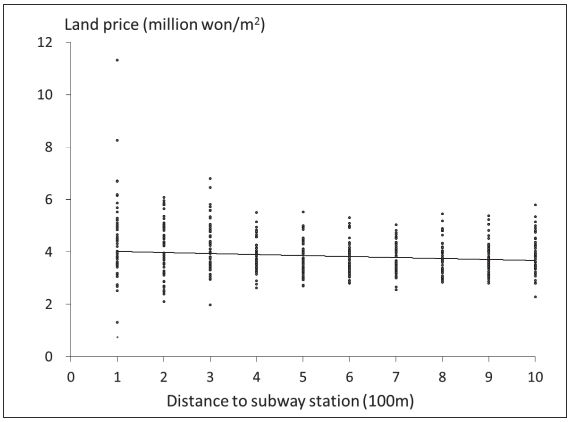

Fig. 4. Land price curve of c1

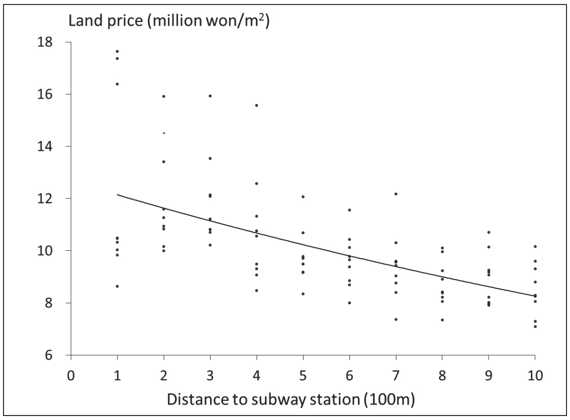

Fig. 5. Land price curve of c2

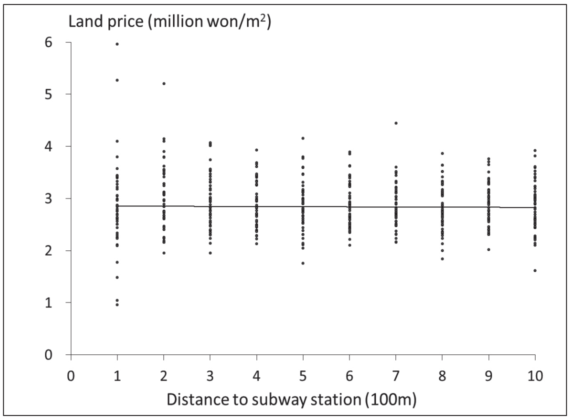

Fig. 6. Land price curve of c3

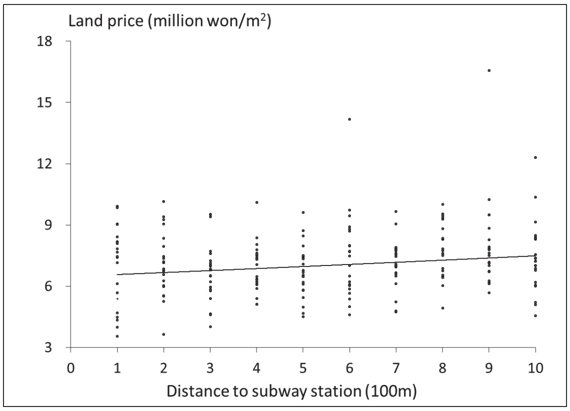

Fig. 7. Land price curve of c4

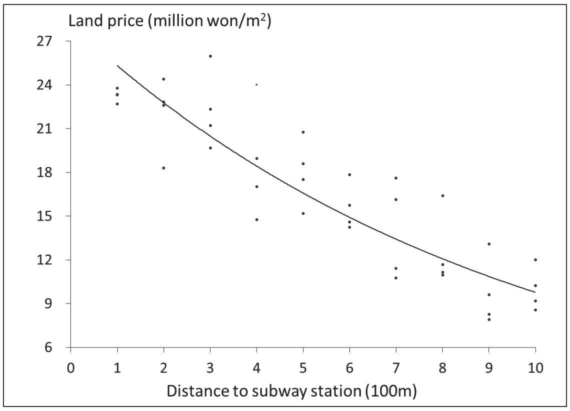

Fig. 8. Land price curve of c5

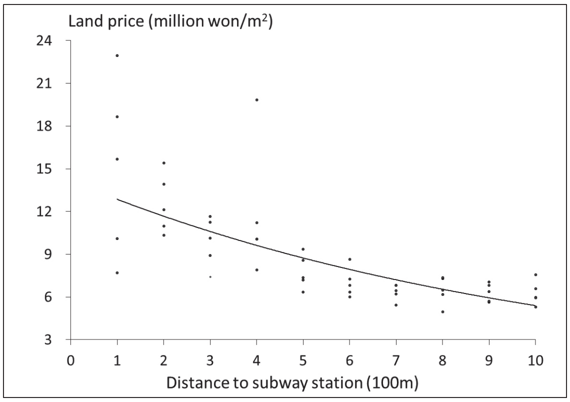

Fig. 9. Land price curve of c6

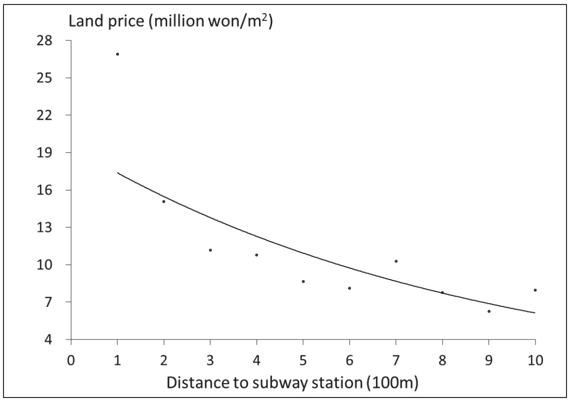

Fig. 10. Land price curve of c7

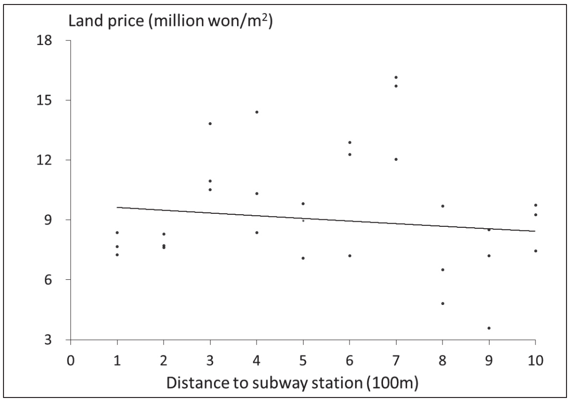

Fig. 11. Land price curve of c8

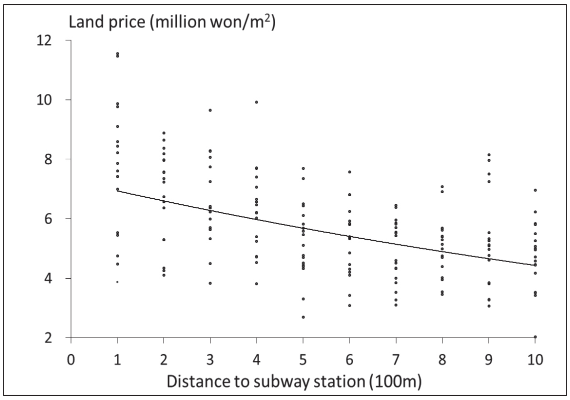

Fig. 12. Land price curve of c9

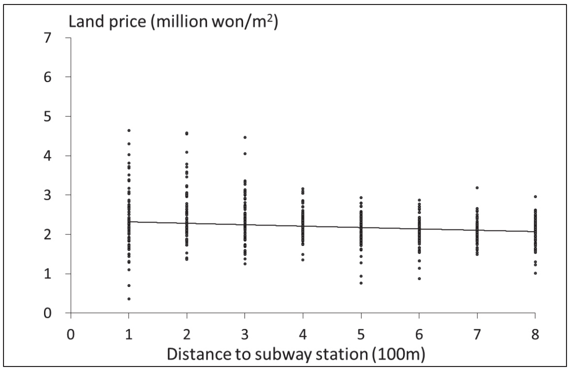

Fig. 13. Land price curve of c10

Table 3. Cluster analysis results

Cluster	Number of stations	Location of station	State of area of influence
C1	55	Adjacent area of center or suburbs	Weak
C2	9	Sub-centers	Strong
C3	50	Suburbs	None
C4	22	Adjacent area of center or sub-centers	None
C5	4	Center	Strong
C6	5	Sub-centers or local centers	Strong
C7	1	Gyeongbokgung	Strong
C8	3	Adjacent area of center	None
C9	18	District centers	Strong
C10	74	Suburbs	Weak

In summary, subway stations in c2, c5, c6, c7, and c9 show strong station areas of influence. Land prices are high near subway stations, which considerably decrease from an increase in distance from the exit. Stations in c1 and c10 show weak areas of influence. Land price is low near the exits, and does not decrease much as the distance from the exit increases. C3, c4, and c8 do not show any formations of the areas of influence.

6. Conclusion

In this research, we have developed a method that analyzes whether subway station areas of influence have been formed, and have verified formation of the area of influence through empirical analysis on subway stations in the city of Seoul. The method is comprised of 2 steps; step 1 applies cluster analysis using K-means clustering, and step 2 applies regression analysis to estimate land price curves. The analysis was conducted as follows. 100m-interval buffers are created from 100m to 1000m distances from the subway station exits. Average land prices are calculated for each buffer and, subsequently, K-means clustering is applied to classify the areas of influence into 10 clusters. We then assumed a decreasing land price as the distance from the station increases, conducted a regression analysis for each cluster, and evaluated whether the areas of influence were actually formed using the results.

According to land price curve estimation results, the 10 clusters are divided into strong, weak, and no areas of influence. Clusters with strong areas of influence are generally located in center, sub-centers, local centers, and district centers; clusters with weak or no areas of influences are mostly located in adjacent area of center or sub-centers or suburbs. This signifies that our method using K-means clustering and land price curve estimation is effective.

In this research, we have used 10 clusters considering the spatial structure of the city of Seoul for K-means clustering. However, considering the diversity of the area of influence characteristics, it is needed to diversify the number of clusters to improve evaluation method. In addition, centers and sub-centers may form the combined area of influence with other nearby areas and thus, it is necessary to improve the analysis method through distinguishing these characteristics. Though we conducted evaluating with land price data, there still are data that show similar to land price distribution, including floating populations and development density. Evaluating with these data is also deemed to be possible.

References

Bae, J. G. (1982), A study on the socio-economic analysis of the rapid rail transit - with reference to the case of subway line No. 1 in Seoul, *Community Development Review*, Vol.

7, pp.43-64. (in Korean with English abstract)

Bae, W. and Roh, S. W. (2005), A study on K-means clustering, *The Korean Communications in Statistics*, Vol. 12, No. 2, pp.497-508. (in Korean with English abstract)

Choi, Y., Kim, T. H., and Park, J. S. (2008), Development of selection model of subway station influence (SIA) in Seoul city using chi-square automatic interaction detection (CHAID), *Journal of the Korean Society for Railway*, Vol. 11, No. 5, pp.504-512. (in Korean with English abstract)

Kim, D. O., Ryu, Y. G., and Choi, H. G. (2002), A study on the setting up method of subway access/egress area by walking and its application, *Journal of the Korea Planning Association*, Vol. 37, No. 5, pp.177-186. (in Korean with English abstract)

Kim, J. and Lee, M. S. (2010), A study for the analysis on relationship between transit riderships and characteristics of transit centers, *Journal of the Architectural Institute of Korea Planning & Design*, Vol. 26, No. 10, pp.305-312. (in Korean with English abstract)

Kim, S. H., Lee, C. M., and Ahn, G. H. (2001), The influence of walking distance to a transit stop on modal choice, *Journal of the Korea Planning Association,* Vol. 36, No. 7, pp.297-307. (in Korean with English abstract)

Kim, S. Y., Eom, S. Y., and Lee, M. H. (2013), A study on spatial range of Seoul subway station area on characteristics of land use, *Journal of the Korea Planning Association*, Vol. 48, No. 1, pp.177-186. (in Korean with English abstract)

Kim, T. H., Lee, Y. T., Hwang, E. P., and Won, J. M. (2008), Development of selection model of subway station influence area (SIA) in new town using categorical and regression tree (CART), *Journal of the Korean Society for Railway*, Vol. 11, No. 3, pp.216-224. (in Korean with English abstract)

Koo, H., Lee, B., and Lee, C. S. (2016), Analysis of changes in spatial structure of Seoul by analyzing the land price changes of station influence areas, *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography,* Vol. 34, No. 1, pp. 63-70. (in Korean with English abstract)

Lee, Y. S., Choo, S. H., and Kang, J. M. (2011), Setting spatial ranges and analyzing characteristics of the adjacent areas of Seoul subway station, *Journal of the Korea Planning Association*, Vol. 46, No. 7, pp.57-72. (in Korean with English abstract)

Lee, Y. S. and Sohn, D. W. (2012), A relationship analysis between subway transit demand and urban spatial characteristics in the subway station area, *Journal of the Urban Design Institute of Korea*, Vol. 13, No. 4, pp.23-32. (in Korean with English abstract)

Oh, Y. T., Kim, T. H., Park, J. J., and Rho, J. H. (2009), An empirical analysis of influencing factors toward public transportation demand considering land use type Seoul subway station area in Seoul, *Journal of the Korean Society of Civil Engineers*, Vol. 29, No. 4D, pp.467-472. (in Korean with English abstract)

Park, S. H., Sohn, D. W., and Lee, J. H. (2009), A spatial analysis of transit centers in Seoul metropolitan region for developing transit oriented urban environment, *Journal of the Korean Society of Civil Engineers*, Vol. 29, No. 1D, pp.111-120. (in Korean with English abstract)

Sung, H. and Kim, T. H. (2005), A study on categorizing subway station areas in Seoul by rail use pattern, *Journal of Korean Society of Transportation*, Vol. 23, No. 8, pp.19-29. (in Korean with English abstract)