Brand Value of Area-Images Extracted from Spatial Distribution of Building Names

Toshihiro Osaragi and Kenichi Ogawa
Tokyo Institute of Technology

Key words: Brand-value, Area-image, Building name, Spatial distribution, Logit model

Abstract: Our area-images are composed of evaluations derived from the various kinds of elements or activities within different regions. In quantitative terms, the extraction of area-images is difficult. In this paper, we focus on a phenomenon in which a part of people's area-image can be observed in the names of buildings. In the first instance, a model based on the random utility theory is constructed to describe the spatial distribution of building names. Secondly, the proposed model is calibrated using actual data from the city of Tokyo (Setagaya Ward), and effects of such area-elements or activities on area-images are then estimated. Finally, values for the area-images are quantitatively estimated and their spatial distribution is represented on a map.

1. INTRODUCTION

The various images or impressions that people have in a certain area have an influence to no small extent on their daily activities in that area. For example, residents who will select a particular residential area as a domicile are affected by the image of that area. Also, people decide to shop by the image of a shopping area. For this reason, city planners have sought to research and understand area-images from a variety of viewpoints. Among these are a series of research projects on area-images that were carried out by Doi et al. (1995) and Nishii et al. (1996, 1997), as well as research by Saito et al. (1997).

Up until now, area-image research has been conducted primarily by detailed examination of the survey results concerning the qualitative aspects of different area-images. However, it has not been easy to perform relative value appraisals for multiple areas, since area-images are of complex structures derived from a wide variety of factors and the research requires covering extensive areas.

Given this background, this research evaluates a method for extracting quantitative values from area-images based on the spatial distribution of building names that represent people's impressions of the area, i.e., the area-image constructed based not only on people's daily activities and elements of lives, but also on a variety of historical elements and organizational factors.

While area-names indicate the area as actual spaces, at the same time they convey the area-image concerning such spaces' actual conditions and characteristics. Therefore, by attaching the area-name as one part of a building name, the impression of the building is enhanced by the area-image. In other words, the probability is considered to be high that the names of buildings will include such areas' area-names as these representing favorable area-images. Taking this phenomenon into account, this research attempts to extract quantitative *brand values of area-image* (i.e., impressions concerning the relative preferences for an area) by examining the naming behavior for building names (figure 1).

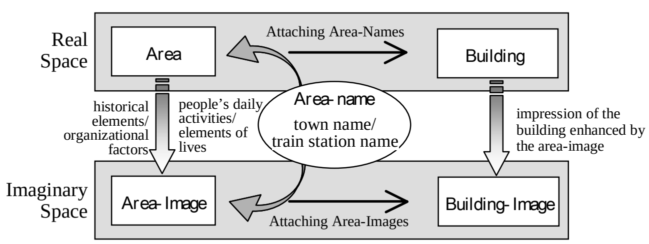

*Figure 1.*Concept of naming behavior for building names

Nakama (1994) has conducted prior research and reported that various geographical factors are associated with the distribution of certain words and expressions used in different area-names, and clarified the distribution of the area-images and the propagation factors for the naming process.

From the same perspective, Yamazaki et al. (1997) used a potential map to visualize the region, which has a particular prevalent area-image, thereby it became possible to visually understand a particular area's strength in its area-image. In addition, Taniguchi and Araki (1996) introduced the concept of area-name choice probability. They demonstrated that hierarchical characteristics exist in area recognition by showing that the range of locations where building names contain the area-name “Tsukuba” has been expanding over year.

The model for this research is from Masuda and Osaragi (1996) and pays attention to the fact that the area-names attached to building names do not necessarily coincide with the nearest train station nor the location of the town in which the building is located, but that the distribution characteristics of building names are to be explained by such factors as busy or lively activities in urban area surrounding the train station. However, this model does not quantify values for area-images. For this reason, the authors employed a geographic information system and constructed data of larger scale in an attempt to extract quantitative information, but they were not able to arrive at a solution for how to measure the quantified constituent factors (Osaragi and Ogawa, 2001). Asami and Kondo (2001) also tried to analyze regional brand value by paying attention to building names, but the constituent factors of regional brand values were left to be elucidated.

This research develops the area-name choice probability concept in order to quantify area-image brand values, and attempts to compare different areas on a relative basis. In addition, it evaluates the constituent factors of area-images, i.e., the kinds of area activities and area characteristics that create brand values of area-images.

2. BUILDING-NAME MODEL AND AREA-IMAGE BRAND VALUE

2.1 Area-names (Train Station Name/Town Name) Added to Multiple Dwelling Residences

In this research, the term *area-name* is used to refer to the town name and/or train station name. To be exact, the station name refers only to the facilities of the train station itself, but in everyday life and conversation it means an expansive area around the train station. Whenever pointing out the shopping areas, etc., surrounding in front of the station, the train station name is used even more frequently than the town name. Therefore, the train station name can also be considered as being included in the collection of area-names.

Examining by building area, the proportion of buildings that have area-names attached to the building name increases as the building area increases (figure 2). This research makes it clear that the area-name is included as part of the building name nearly 80% of the time whenever the building area is 400 m^2 or more. This suggests that greater importance is attached to locality-related information, as the size of the buildings becomes larger and the buildings are located closer to the nearest train station.

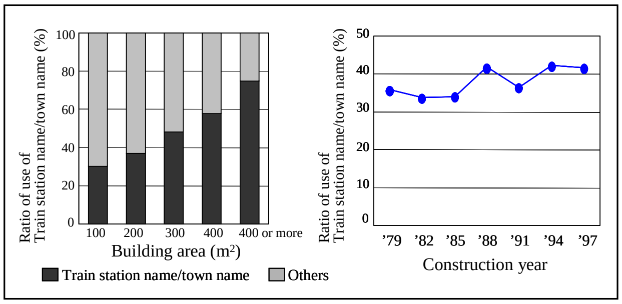

Figure 2. Train station name/town name used for building name

Next, by examining the year of buildings' constructions, it has become clear in this research that the proportion of buildings that were built in 1980 or later that have area-names included in the building names is relatively consistent at about 40% in all cases. In addition, although there was considerable variety in the building names used over a period of years, the proportion of building names that included the area-name remains consistently high. Therefore, there is not a great impact from change of tastes, fashion, or other factors over time.

2.2 Construction of Building-Name Model

There are two considerations for the evaluation of building names: 1) whether the area-name is used, or whether a different name is used instead of the area-name; and 2) which area-name is used if an area-name is used. By paying attention to the process in 2), the description below attempts to establish a model for the choice behavior in naming the buildings.

Looking at the actual spatial distribution of building names, the area-name included within the name of the building does not necessarily correspond to the name of the nearest train station or the town in which the building is located. This phenomenon can be considered to manifest that there are differences for each train station and town name in the brand value of the area-image (figure 3).

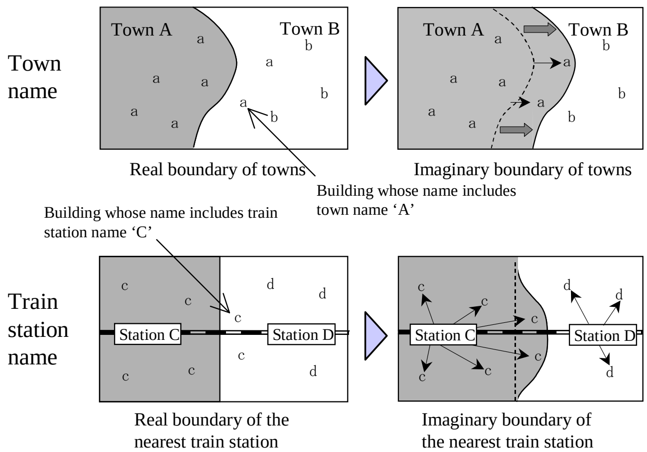

Figure 3. Distortion of spatial distribution of building names

However, an important point that should be noted here is that when attaching a town name that is outside the actual location of the building, or when using the name of a train station that is not the nearest station, psychological resistance is recognized for this unnatural naming behavior. The degree of psychological resistance can be considered to increase as the distance from the corresponding train station or town border increase.

Therefore, a model was constructed that describes the spatial distribution of building names. We anticipate that, within the area where a building is located, a building name that gives the most favorable impression in terms of area-image is selected from a collection of area-names (train station names/town names).

The probability for a building located in *Location i* to use a name from *Area j* is described by the following Logit model according to the random utility theory.

$$P_{ij} = \frac{\exp[U_{ij}]}{\sum_k \exp[U_{ik}]} \qquad (1)$$

The fixed term *Utility* U_{ij} is considered to be made up of the *Area-Image Brand Value* U_j of *Area j* and the *Resistance* C_{ij} for using a name from *Area j*. This is expressed by the following formula.

$$U_{ij} = U_j + C_{ij} \qquad (2)$$

When the area-name assigned to a building is the same as the actual location of the building, resistance is not expressed, and the value of *Resistance* C_{ij} is 0. When a train station name or town name is used that is not within the vicinity of the building location, then the resistance function increases according to distance. Hence, *Resistance* C_{ij} is based on *Distance From Town Border* X_{ij1}, *Distance From Station* X_{ij2}, and *Unknown Parameters* β_1, β_2. Refer to figure 4 for details concerning this model.

$$C_{ij} = \beta_1 X_{ij1} + \beta_2 X_{ij2} \tag{3}$$

The above model is called the *Basic Model*, and model estimation was attempted using the Setagaya District data. From the sample number of 8620 residential and commercial buildings, there were 94 area-names that became the set of choices. The estimated results are shown in figure 5(1).

(1) Supposition

Building at location i will employ a name, which includes the highest utility, from the set of available names.

(2) Supposition for the set of available names

Names of towns located at 1500 m or less from the location of building i.
Names of train stations located at 2000 m or less from the location of building i.
When building i employs name j, the value of j will be as follows:

$$j = \{j_1, j_2\}$$ where j_1 is town name, j_2 is train station name

The following 3 cases can be possible:
(a)In case the town name is exactly the same as the train station name, town name and train station name are written as l and l' respectively.
(b)In case there is only a town name, it is written as m.
(c) In case there is only a train station name, it is written as n.
Hence, the values of j_1 and j_2 can be as follows:

$$j_1 = \{l, m\}, \quad j_2 = \{l', n\}$$

(3) Construction of Logit Model

Probability in which the building i will employ the name j can be expressed as follows, where non-stochastic part of utility which building i will receive is described as U_{ij}.

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_k \exp(U_{ik})}$$

(4) Modeling of utility U_{ij}

Consider utility U_{ij} can be described using the area-image brand value, U_j, and resistance C_{ij}.

$$U_{ij} = U_j + C_{ij}$$

Consider the resistance C_{ij} increases according to the distance X from location i to a town/ a train station j. The resistance can be described using unknown parameters β_1, β_2.

$$C_{ij} = \beta_1 X_{ij_1} + \beta_2 X_{ij_2}$$

X_{ij1}: distance from the boundary of a town
X_{ij2}: distances from a train station

The distance X_{ij_1} is defined to be 0 for the case that buildings employ a name of town where they are located.

Figure 4. Basics of Building-Name Model

2.3 Improving the Building-Name Model

Although the fitness of the *Basic Model* is high, by evaluating the spatial distribution of the estimated error, it was confirmed that improvements for evaluating the distance from the town boundaries and the station should be possible by modeling the distance resistance. Therefore, the function type for distance resistance was reconsidered, and it was discovered that the best evaluation is possible by exponential function type for town boundaries and the linear function type for the station (figure 5(2)). A town has a clear borderline, and the feeling of resistance is considered to rapidly increase after crossing the border. This explains the appropriateness of applying an exponential function type for distance from the town boundaries.

Next, because there is a possibility of differences in the function of resistance according to differences in building use, the resistance coefficients were estimated separately for residential and commercial use (figure 5(3)). As a result, compatibility improved, and the model was refined. Observing the resistance coefficients, the value for residential use was less than that for commercial use. This is because, in the case of multiple dwelling residences as compared to shopping areas, etc., the residents have a strong consciousness of addresses, and, accordingly, it can be inferred that, in residential use, it is harder to attach a town name or train station name to a building that is not the nearest station or that is outside the boundaries of that town.

Next, the resistance coefficient values may vary between when both an existing train station name and a town name are the same (hereafter referred to as *duplicate name*), and when an existing train station name and town name are different (a name that is only the town name or only the train station name). Therefore, the resistance coefficients are divided into two more categories and estimated (figure 5(4)). Looking at the estimated values of the resistance coefficients, the absolute value for duplicate names is less than that for names that are only the town name or only the train station name. In other words, it can be understood that in cases when the train station name and town name are the same, that name is more easily attached to building names that are relatively farther away. Favorable results could be obtained by looking at the *Likelihood Ratio*, *AIC*, or *Hitting Ratio*, and the model was further improved. The model below uses eight resistance coefficients for *Resistance* C_{ij}, and it is called the *Building-Name Model*.

(1) Basic model

Evaluation Indices	Likelihood Ratio	0.733
	AIC	19168
	Fitting Ratio	67.0%
Estimated parameters	Distance from boundary of towns $\beta 1$	-6.76
	Distance from train stations $\beta 2$	-3.25

(2) Model 1

Evaluation Indices	Likelihood Ratio		0.742	0.743	0.766	0.768
	AIC		18568	18474	16845	16659
	Fitting Ratio		67.7%	67.8%	71.0%	71.2%
Estimated parameters	Distance from boundary of towns	(function type)	log(X+1)	log(X+1)	exp[X]	exp[X]
	Distance from train stations	(function type)	log(X+1)	X	log(X+1)	X

(3) Model 2

Evaluation Indices	Likelihood Ratio		0.770
	AIC		16586
	Fitting Ratio		71.2%
Estimated parameters	Distance from boundary of towns	residential $\beta 1$	-2.40
		commercial $\beta 3$	-2.01
	Distance from train stations	residential $\beta 2$	-3.56
		commercial $\beta 4$	-3.33

(4) Model 3 [Building Name Model]

Evaluation Indices	Likelihood Ratio			0.779
	AIC			15943
	Fitting Ratio			71.4%
Estimated parameters	Distance from boundary of towns	residential	single $\beta 1$	-2.77
			double $\beta 1'$	-1.84
		commercial	single $\beta 3$	-2.41
			double $\beta 3'$	-1.44
	Distance from train stations	residential	single $\beta 2$	-5.68
			double $\beta 2'$	-2.45
		commercial	single $\beta 4$	-5.28
			double $\beta 4'$	-2.38

single: the name is the same as an existing train station name or a town name
double: the name is the same as both an existing train station name and a town name

Figure 5. Improvement of models

2.4 Area-Image Brand Value

Town names and train station names were considered as two different categories because a town name clearly expresses the area within the borders of the town, whereas the boundary for the area indicated by the train station name is not necessarily clear. Therefore, buildings that contained the train station name were eliminated from the samples (*N*=7586), and the area-image brand value was estimated using the town name only as a selection. These results are displayed in figure 6. From the figure, it is apparent that the area-image brand value is higher in areas running along train lines.

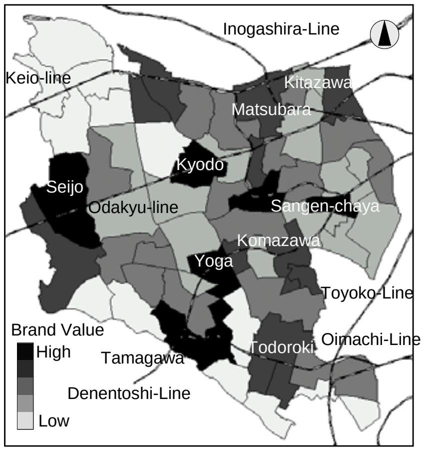

Figure 6. Spatial distribution of area-image brand value

2.5 Visualizing the Region of the Area-Image

Using the estimated *Building-Name Model*, the area-names that have the highest probability of being used for building names are found and shown on the map (figure 7) for each location. It can be derived that the image relating to the expanse of a particular area is influenced by daily recognition of building names within that area. In other words, the map in figure 7 applies to one type of area-image region that implicitly expresses the area that is easiest to be perceived as being a part of that area-image region.

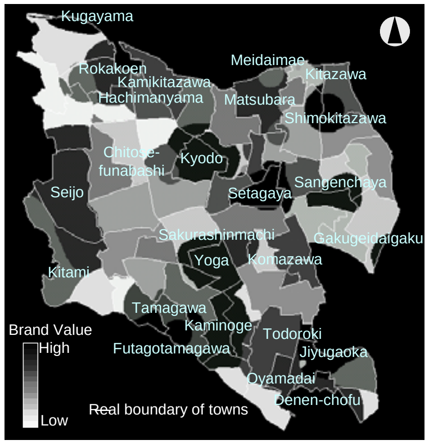

Figure 7. Real/imaginary boundaries of towns

3. FACTORS CONTRIBUTING TO THE AREA-IMAGE BRAND VALUE

3.1 Area-Image Model

Whereas the previous section estimated the area-image brand value using *Brand Value* U_j, this section considers factors that generate differences in brand value. By being rewritten, the model shown at the left of figure 8 now becomes a model describing the area-image brand value using urban activities. The descriptions below differentiate this model, called the *Area-Image Model*, from the *Building-Name Model*.

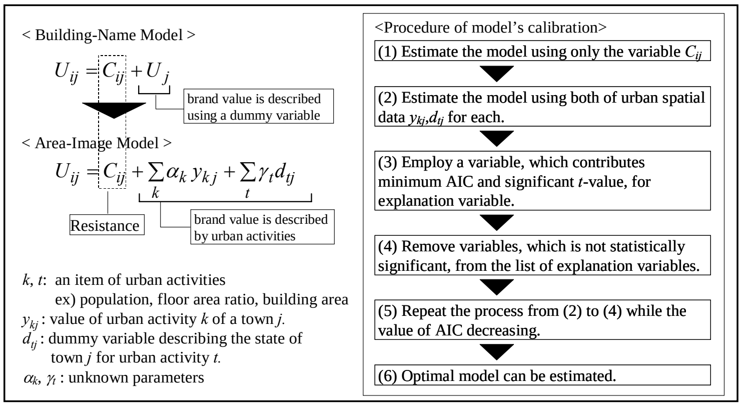

Figure 8. Model for describing area-image and method for estimation

3.2 Area-Image Model Estimation

Because attributive data relating to train stations are scarce, model estimation was performed here covering the data acquired from 70 town names where various urban activities data were available. The estimation was performed by the method shown in figure 8, and AIC was used for model evaluation.

As shown in table 1, there are 37 types of variables for the urban activities data used in the analysis. Table 2 shows a comparison of estimation results of the *Area-Image Model* and the *Building-Name Model*. Favorable compatibility was obtained without using dummy variables in the *Building-Name Model*. The brand value estimated according to the *Building-Name Model* is designated as U_j, and the brand value estimated by the *Area-Image Model* is designated as U_j^*. Both are compared in figure 9.

Overall, there was variability in the vicinity 45 degrees from the train line, where a large bias was not observed. However, there is a low frequency of using the name Tamagawa-Denen-chofu in building names because the name is too long, and because it is contiguous with the famous high-class residential district Denen-chofu. For this reason, the Brand Value U_j that was estimated using the dummy variable is under-estimated.

Table 1. Variables of urban activities used for models' calibration

Dummy variable		18	Number of wooden houses
1	whether the name is the same as both an existing train station name and a town name	19	Number of wooden apartment houses
Variables defined for each train station		20	Ratio of wooden apartment houses
2	Total number of passengers per day for each train	21	Ratio of wooden structure apartment houses
3	Number of commercial buildings located within 200 m from a train station.	22	Number of buildings
Variables defined for each town		23	Density of buildings
4	Area of town	24	Average area of detached houses
5	Number of families	25	Number of detached houses
6	Population	26	Ratio of detached houses
7	Density of population	27	Ratio of housing
8	Density of families	28	Average number of stories
9	Number of distributors	29	Average building coverage ratio
10	Number of retail stores	30	Average floor area ratio
11	Number of employees of distributors/retail stores	31	Ratio of field area
12	Density of distributors	32	Ratio of housing area
13	Density of retail stores	33	Ratio of commercial area
14	Density of employees of distributors/retail stores	34	Ratio of forest area
15	Number of commercial buildings	35	Ratio of park/ground area
16	Density of commercial buildings	36	Time distance from town to train station
17	Ratio of commercial buildings	37	Time distance from town to Yamanote Line

Table 2. Results of estimation of models

	Likelihood Ratio	AIC	Fitting Ratio
Building-Name Model	0.812	11269	78.70% (5970 / 7586)
Area-Image Model	0.792	12329	77.97% (5915 / 7586)

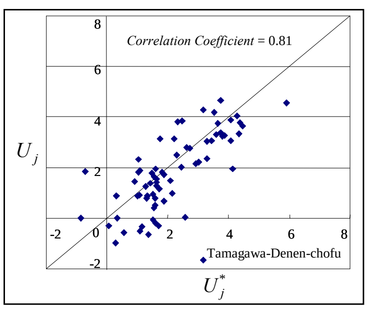

Figure 9. Relationship between values of U_j and U_j^*

3.3 Interpretation of Estimation Parameters

Values for the *Area-Image Model* estimation parameters are shown in table 3. The parameter codes express the positive and negative impacts to the area-image brand value, and the standardized parameter values express the level of influence on the area-image brand value. The largest positive operative force is a dummy variable expressing whether or not the same train station name exists. In other words, the existence of a train station has a strong influence on the formation of an area-image, and it is shown that the train station name plays an important role in attaching generic name to the urban activities around the train station. The average building area is also a positive operative force. In other words, high-class residential streets and their highly related premises greatly affect the area-image. Also, the ratio of the number of commercial buildings to small shops is an important factor toward having a positive operative force. It can be considered that towns that have lively commercial activities and are often featured in the mass media contribute to the formation of a favorable area-image.

Meanwhile, the following functions concern factors that decrease the area-image brand value. First of all, the strongest negative operative force is the proportion of land that is agricultural or vacant because a high ratio is associated with an area still under development, and this can be considered to reduce the brand value. Station accessibility is expressed as the distance in traveling time from a representative point in the town to the nearest train station, and it can be verified that the brand value decreases as distance from the station increases. The ratio of the number of wooden structure apartment houses to the number of apartment houses shows the number of rental apartments of wood construction, and it is apparent that wood fabricated apartments, etc., are a factor that reduces area-image brand values.

Table 3. Estimated parameters of Area-Image Model

	Urban activity	Estimated parameter	Standardized parameter
1	Dummy variable describing whether the name is the same as both an existing train station name and a town name	1.7834	0.828
2	Average building area of detached houses (m2)	0.0279	0.472
3	Ratio of commercial buildings (%)	0.0568	0.392
4	Ratio of field area (%)	-0.0845	-0.245
5	Time distance from town to the nearest train station (min.)	-0.0451	-0.197
6	Ratio of wooden structure apartment houses (%)	-0.0134	-0.181
7	Number of retail stores	0.0014	0.144

4. APPLICATION TO EXISTING URBAN MODELS

4.1 Combining Brand Value with Existing Urban Models

The effectiveness of combining the *Area-Image Model* with existing urban models has been confirmed. As for existing urban models, a model developed by Aoki et al. (1994) was used. This model describes the land value according to travelling time and distance using defined travel costs and the condition of urban activities. Figure 10 outlines this and also shows methods for combining models: a model for travel costs only, a model overview for travel costs together with the condition of urban activities, and reconstruction of the models combined with the area-image brand value.

Land value model previously proposed

< Travel Cost >

$$\log P = -C$$

$$C = b_0 + b_1 M + b_2 Y + b_3 M^2 + b_4 M^3$$

P : land value
C : Transportation cost
M : Time distance from home to the nearest station
Y : Time distance from the nearest train station to Yamanote-Line

< Travel Cost + Urban Activities >

$$\log P = -C + \sum_{i=1}^{n} \sum_{j=1}^{m_i} (\log h_{ij}) d_{ij}$$

d_{ij} : Dummy variable describing whether urban activities of place i is j
h_{ij} : Unknown parameters

+

Land value model combined with area-image brand value

< Travel Cost + Urban Activities + Brand Value U_k >

$$\log P = -C + \theta U_k$$

$$\log P = -C + \sum_{i=1}^{n} \sum_{j=1}^{m_i} (\log h_{ij}) d_{ij} + \theta U_k$$

θ : Unknown parameter
U_k : Exponential function of brand value U_j , U_j*

Figure 10. Land value model combined with area-image brand value

4.2 Comparison of Estimation Results

In the applicable area analyzed in this research, the results are shown on table 4 for estimating a model using only travel costs and for estimating a model using a designated floor-area ratio. In the case of the estimation using the floor-area ratio, the compatibility of the real values and the estimated values are shown on the left of figure 11. A relatively favorable result was obtained, but a large residual error is apparent in locations with high real values.

Next, an estimate was performed combining the area-image brand value with the existing land value model above. After converting brand value into a variety of function types, it is shown that the exponential function type is the most preferable. The estimated results are shown on the right of table 4 and on the right of figure 11.

There are few cases where the residual error was large, particularly in locations where the real value was high. Even when observing the AIC value, it is apparent that the model has been improved. For confirmation purposes, the same kind of analysis was attempted using a dummy variable combined with estimated *Brand Value* U_j, but more favorable results were achieved using urban activities and estimated *Brand Value* U_j^*. These results also confirm the usefulness of the *Area-Image Model*.

Table 4. Improvement of land value model by using brand value

Explanation variables	Travel Cost	Travel Cost + Floor-area Ratio	Travel Cost + Bland Value U_j *	Travel Cost + Floor-area Ratio + Bland Value U_j *
Correlation Coefficient	0.731	0.838	0.791	0.881
AIC	-58	-142	-101	-197
D.F.	4	10	5	11

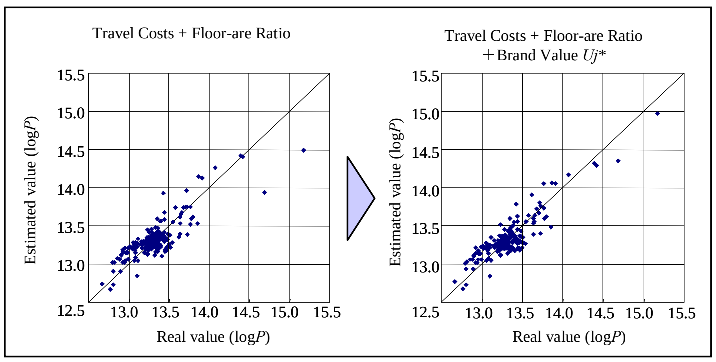

Figure 11. Improvement of land value model by using area-image brand value

5. SUMMARY AND CONCLUSIONS

In this research, a *Building-Name Model* was constructed to describe the choice behavior for assigning building names according to the brand value of the area-image. First, one kind of choice behavior was considered for assigning building names that contain area-names (train station name/town name). Then, quantitative values were extracted for the brand value of the area-image using actual data (data for the spatial distribution of building names). In order to evaluate the constituent factors of this brand value in detail, an *Area-Image Model* was constructed, and in order to perform model estimation, the relationships between various urban activities and the brand value were clarified. Finally, by combining this model with existing urban models in order to create a model that better reflects actual conditions, the research shows the possibility and effectiveness of combining urban models with a goal to quantify area-image brand values.

Using the models proposed above in this research for spatial distributions of building names, quantitative values for area-image brand values can be extracted and analyzed relatively easily where as it has been conventionally difficult to grasp because of its obscurity. In conventional urban models, there has not been great discussion concerning area-image brand values, such as by being included in error terms, etc. However, by using these models, it is possible to proactively combine area-image brand value into different models as important elements to express area characteristics. Compatibility and applicability of these models with conventional urban models is high.

6. REFERENCES

Aoki Y., T. Osaragi, K. Ishizaka, 1994, "Interpolating model for land price data with transportation costs and urban activities", *Environment and Planning B, Planning and Design*, vol.21, p.53-65.

Asami, Y. and Y. Kondo, 2001, "Brand values of area names through an analysis of distribution of buildings containing area names", *Proceedings of Geographical Information Systems Association*, vol.10, p.39-43.

Doi, T., T. Kiuchi, A. Mihoshi, H. Kitagawa, and K. Nishii, 1995, "An image of railroad line area and its structure", *Infrastructure Planning Review*, 12, p.267-374.

Masuda, K. and T. Osaragi, 1996, "Spatial distribution of building names", *Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan*, F-1, p.783-784.

Nakama, K., 1994, "A study on the spread of the district image refered to local area-names", *City planning review. Special issue, Papers on city planning*, 29, p.607-612.

Nishii, K., M. Tanahashi, T. Doi, and T. Kiuchi, 1996, "The LOGMAP modelling for structural analysis of railroad line area image: An empirical study on the attribute – regression model", *Infrastructure Planning Review*, 13, p.49-56.

Nishii, K., M. Tanahashi, T. Doi, and T. Kiuchi, 1997, "The LOGMAP model for image positioning of railroad line area: focusing on difference linguistic image between common and proper nouns", *Infrastructure Planning Review*, 14, p.107-114.

Ogawa, K. and T. Osaragi, 2001, "Regional image analysis based on spatial distribution of building names", *Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan* , F-1, p.829-830.

Osaragi, T. and K. Ogawa, 2001, “Regional image extracted from spatial distribution of names of apartment houses”, *Proceedings of Geographical Information Systems Association*, vol.10, p.295-298.

Saito, K., H. Ishizaki, T. Tamura, and Y. Masuya, 1997, “Analysis on the image structure of city and their relation to physical characteristics”, *Infrastructure Planning Review*, 14, p.467-474.

Taniguchi, M. and S. Araki, 1996, “Regional competition on cognition and their affecting factors: Based on land name & naming”, *Infrastructure Planning Review*, 13, p.225-232.

Yamazaki A., A. Fujii, H. Magaribuchi, and K. Ito, 1997, “A study on urban space refered to spatial distribution of building names”, *Summaries of Technical Papers of Annual Meeting, Architectural Institute of Japan* , F-1, p.783-784.