Discussion Paper No. 75R

A PDF-based analysis of the retail structure in a metropolis

Yukio Sadahiro *

SEPTEMBER, 1998

*Center for Spatial Information Science and Department of Urban Engineering
University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

September 3, 1998

A PDF-based method for analyzing the retail structure in a metropolis

Yukio Sadahiro
Department of Urban Engineering, University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Phone: +81-3-3812-2111 (ext. 6273)
Fax: +81-3-5800-6965
E-mail: sada@okabe.t.u-tokyo.ac.jp

A PDF-based method for analyzing the retail structure in a metropolis

Abstract

The present paper proposed a new method for analyzing the retail structure in a metropolis, using microscale location data of individual retail establishments. The method is based on the probability distribution function (PDF) of establishments estimated from their locational data. The application of the method to the retail structure in Yokohama showed some empirical findings: (1) retail activities that tend to gather with the same type of establishments also gather with other types of establishments, (2) a close relationship exists between the degree of agglomeration and retail function, (3) there are six major patterns of retail distributions, (4) there are five levels of retail agglomerations including isolated stores.

1 Introduction

When analyzing the spatial structure of retailing in a metropolis, we have usually used the locational data of establishments aggregated by spatial units, instead of directly investigating individual location data. There have been a lot of studies based on the data aggregated by retail agglomerations or administrative units (Berry, 1967; Scott, 1970; Davies and Rogers, 1984; Morrill, 1987; Jones and Simmons, 1990; Brown, 1992). Though there are some exceptions such as microscopic studies of store arrangement pattern (Scott, 1959; Carter and Rowley, 1966; Getis and Getis, 1968; Scott, 1970; Brown, 1992), the analysis of urban retail structure has been mainly based on the spatially aggregated locational data.

The use of aggregated data, however, is often problematic in retailing analysis. This is, in short, because there do not exist spatial units used for data aggregation that are defined in a reasonable, consistent and objective manner and suitable for analyzing the spatial structure of retailing. There have been used a wide variety of spatial units in existing studies: shopping areas (Davies 1959; Leeming, 1959; Clark, 1967; Logan, 1968; Ashworth and de Vries, 1985), square lattices (Lee and Koutsopoulos, 1976; Miki, 1980), administrative units (Applebaum, 1966), Voronoi regions (Miki, 1983; Okabe *et al.*, 1992), and so forth. Though numerous definitions of shopping area have been proposed in the literature, they are subjective to some extent (Scott, 1970; Brown, 1992) and no definition can be applied commonly to diverse situations. Inconsistency in the definition of shopping areas makes it difficult to compare shopping areas between different regions and times. One might think that data aggregation on lattices or administrative units are reasonable because they are objectively defined. These methods, however, impose unnatural boundaries to the distribution of retail stores and thus sometimes yield strange results. They also suffer from the problem known as the modifiable areal unit problem (Openshaw, 1984), that is, the results of analysis heavily depend on the spatial units.

The above discussion suggests that it is desirable to use the locational data of establishments without aggregation. It is now practically feasible because the diffusion of geographical information systems (GIS) has been accompanied by the wide spread of microscale spatial database, say, locational data of individual establishments. However, there are only a few methods for analyzing a retail structure using such data as mentioned earlier. A new method needs to be developed which makes full use of the locational data of establishments.

To answer this demand, the present paper develops a method for analyzing the retail structure in a metropolis using microscale locational data of establishments. We take an approach often called the *surface-based method* which treats the distribution of retail stores as a continuous rather than discrete distribution, without aggregating data across spatial units such as shopping areas and administrative units. Hence the method is free from the problems mentioned above: unnatural or subjectively chosen spatial units, difficulties in comparative study, and the modifiable areal unit problem. In addition to this, as we will see later, the surface-based approach is suitable for analyzing the global structure of retailing in relation to its underlying urban structure.

The main objective of the paper is to propose an analyzing method rather than to investigate an actual retail structure. However, we proceed in parallel with its empirical application since it would help understanding of the analysis procedure. In the following section, we briefly describe the locational data of establishments in Yokohama that are used throughout in the paper. From the data we can estimate the probability density function (PDF) of establishments. We outline the estimation method in Section 3. In Section 4, we propose two indices drawn from PDF to represent agglomerative tendencies of retail establishments. The indices focus on individual retail activities rather than the spatial relationship among them. In contrast to this, we explicitly investigate the spatial relationship between retail activities and the whole structure of retailing in Section 5. Finally, we summarize the conclusions in Section 6.

2 Locational data of retail establishments in Yokohama

Yokohama is a city located in the south of Tokyo, Japan (Figure 1). The 1997 population is some 3, 340, 000 and the covering area is 76 square kilometers. Yokohama is the capital of Kanagawa prefecture, and consequently most of administrative agencies are located around Sakuragicho and Kannai stations (Figure 2). Central business districts are scattered around Yokohama, Sakuragicho, Kannai, and Shin-Yokohama stations.

Figure 1. Location of Yokohama.

Figure 2. Railway network in Yokohama. Only major stations are labelled with their names.

There are around 80, 000 retail and service establishments in Yokohama. As shown in Figure 3, establishments are highly concentrated around railway stations. There seems a close relationship between the number of boarding passengers of stations and the density of their surrounding establishments. Though one might think that is strange, retail agglomerations are usually formed around railway stations in Japan (Okabe and Miki, 1984: Sadahiro, 1994).

Figure 3. Density distribution of establishments in Yokohama.

The source data used in the present paper is the list of establishments appeared in the 1997 NTT (Nippon Telegraph and Telephone Corporation) telephone directory. The directory categorizes establishments into 1, 650 types, each of which contains various numbers of establishments ranging from 2 to 4, 000. From 1, 650 retail types we chose 78 that are most important and influential for the retail structure (Table 1) for plain explanation of the analyzing method. We should note that, however, the method we propose is applicable for any number of retail types or establishments.

Table 1 Retail types analyzed in the present study and the number of establishments

The NTT directory includes not only the phone number but the address of establishments. We converted the addresses into (X, Y) coordinates in a rectangular coordinate system through the address matching technique (Bohman-Carter, 1994; Longley and Clarke, 1995). The obtained spatial data were handled in ARC/INFO ver 7.1 running on a UNIX workstation.

3 Estimating PDF of establishments

Retail establishments seek a location to obtain as much benefit as possible, taking a variety of spatial factors into account. Gasoline stations, for instance, tend to be located along highways and major roads to capture running cars. Women's clothing stores form retail agglomerations in downtown for convenience of comparison shopping. Souvenir shops are clustered around hotels or convention halls hoping that the visitors of the hotels stop at the shops. We can say that the distribution of establishments is a realization of their benefit seeking behavior.

It is almost impossible, however, that all the establishments are located at their best locations because such locations are usually preoccupied by other buildings or landuses. Consequently, establishments are forced to choose relatively better location rather than the best. To consider this more explicitly, we in this paper regard the distribution of establishments as the distribution of samples taken from an underlying probability distribution. We estimate the probability density function (PDF) from a retail distribution, and analyze the PDF in place of the retail distribution itself.

This is a kind of surface-based method which has been often used in geography and GIS. Similar approach has been adopted in the literature to retail distributions (Okabe and Sadahiro, 1994), population distributions (Bracken and Martin, 1989, 1995; Bracken, 1993), and so forth (Scott, 1992; Fotheringham and Rogerson, 1994; Bailey and Gatrell, 1995). Advantages of the surface-based approach can be summarized as follows. First, the surface-based approach is free from the problems caused by determining crisp boundaries for data aggregation as discussed in Section 1. Second, comparison between surfaces can be done straightforward (Bracken and Martin, 1995). Third, since the approach explicitly focuses on the global structure of spatial phenomena rather than its individual outcomings, it helps us to detect the relationship between a distribution and its underlying factors (Scott, 1992; Bailey and Gatrell, 1995).

Let us consider the distribution of type i establishments in a region S. We denote the location of establishment j by $\mathbf{x}_{ij}$, and the number of type i establishments by n_i. To estimate the PDF of type i establishments from their locational data, we employ the kernel density estimation where the kernel function is defined by the two-dimensional Gaussian distribution (Silverman, 1986; Scott, 1992). Let $f_i(\mathbf{x})$ be the PDF of type i establishments. Its estimator $\hat{f}_i(\mathbf{x})$ is then given by

$$\hat{f}_i(\mathbf{x}) = \frac{1}{2n_i \pi h_i^2} \sum_j \exp\left(-\frac{1}{2h_i^2} \left| \mathbf{x} - \mathbf{x}_{ij} \right|^2 \right), \tag{1}$$

where h_i is the smoothing parameter of the kernel called *window width*. Given the locational data of establishments, we can estimate the window width h_i from the data by various procedures (Scott, 1992). We adopt the least-square cross-validation method (Rudemo, 1982; Bowman, 1984; Silverman, 1986) which is attractive for its high tractability and simplicity. Outline of the method is as follows.

Consider the mean integrated square error (MISE) of $f_i(\mathbf{x})$, a measure of the illness of model fitting, which is given by

$$\begin{aligned} \mathrm{MISE}\left[f_i \right] &= \mathrm{E}\left[\int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - 2\int_{\mathbf{x}\in S} \hat{f}_i(\mathbf{x}) f_i(\mathbf{x})^2 \mathrm{d}\mathbf{x} + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} \right] \\ &= \mathrm{E}\left[\int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - 2\int_{\mathbf{x}\in S} \hat{f}_i(\mathbf{x}) f_i(\mathbf{x})^2 \mathrm{d}\mathbf{x} \right] + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} \end{aligned}. \tag{2}$$

Suppose an estimator $\hat{f}_{i,-k}(\mathbf{x})$ constructed from the locational data set $\{\mathbf{x}_{ij} \mid j \neq k\}$, that is,

$$\hat{f}_{i,-k}(\mathbf{x}) = \frac{1}{2\pi h_i^2 (n_i - 1)} \sum_{j \neq k} \exp\left(-\frac{1}{2h_i^2} \left| \mathbf{x} - \mathbf{x}_{ij} \right|^2 \right). \tag{3}$$

Using Equation 3 we have

$$\begin{aligned} \mathrm{MISE}\left[f_i \right] &= \mathrm{E}\left[\int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - 2\int_{\mathbf{x}\in S} \hat{f}_i(\mathbf{x}) f_i(\mathbf{x}) \mathrm{d}\mathbf{x} \right] + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} \\ &= \mathrm{E}\left[\int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - 2\, \hat{f}_{i,-n_i}\left(\mathbf{x}_{n_i} \right) \right] + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} \\ &= \mathrm{E}\left[\int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - \frac{2}{n_i} \sum_k \hat{f}_{i,-k}\left(\mathbf{x}_k \right) \right] + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} \end{aligned} \tag{4}$$

(for details, see Silverman, 1986). Unfortunately, the expectation term of the above equation is not computable. This makes it impossible to obtain directly the best estimator of h_i that minimizes the MISE[f_i].

We then turn to the quantity $I(h_i)$ defined by

$$I(h_i) = \int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - \frac{2}{n_i} \sum_k \hat{f}_{i,-k}(\mathbf{x}_k) + \int_{\mathbf{x}\in S} \left\{ f_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x}. \quad (5)$$

Comparing Equations 4 and 5, we notice that $I(h_i)$ is an unbiased estimator of MISE[f_i] which can be calculated from the locational data and h_i. We thus choose the window width h_i that minimizes $I(h_i)$ since it is supposed that the minimizer of $I(h_i)$ would be close to the minimizer of its expectation MISE[f_i].

Practically, it is not necessary to evaluate the third term of Equation 5. Thus h_i can be estimated by solving the following minimization problem.

$$\min_{h_i} \int_{\mathbf{x}\in S} \left\{ \hat{f}_i(\mathbf{x}) \right\}^2 \mathrm{d}\mathbf{x} - \frac{2}{n_i} \sum_k \hat{f}_{i,-k}(\mathbf{x}_k) \quad (6)$$

We can reach the optimum h_i by a nonlinear programming method such as the descent method (Fiacco and McCormick, 1968; Gill *et al*., 1981).

We applied the above method to the individual retail distributions in Yokohama, and obtained 78 PDFs. Figures 4, 5, and 6 illustrate the PDFs of banks, convenience stores, and used car shops, respectively.

Figure 4. The distribution of banks and PDF.

Figure 5. The distribution of convenience stores and PDF.

Figure 6. The distribution of used car shops and PDF.

In addition to the PDFs, we calculated the average PDF of all retail types, that is,

$$\bar{f}(\mathbf{x}) = \frac{1}{N} \sum_i f_i(\mathbf{x}) \quad (7)$$

where N is the number of retail categories (Figure 7). The average PDF $\bar{f}(\mathbf{x})$ indicates the degree of retail agglomeration around the location **x**. The average PDF, however, can be regarded to represent not only the degree but the size of retail agglomeration. In large shopping areas establishments are usually clustered, and the definition of shopping areas has used the size as well as the density of retail establishments (Leeming, 1959; McEvoy, 1972; Potter, 1982; Morrill, 1987). This supports a close relationship between the degree and the size of retail agglomeration.

In this paper, we use the average PDF instead of strictly defining shopping areas as regions with clear boundaries. This permits us to treat shopping areas as they are, namely, regions with fuzzy boundaries, and enables us to avoid the difficulties arising in the shopping area definition that were mentioned earlier.

Figure 7. The average PDF of 78 retail types.

4 Agglomerative tendencies of retail establishments

In analyzing retail distributions, we frequently have the question "whether the establishments tend to be clustered or dispersed?" This is because the degree of agglomeration is an important key to understand the location behavior of a retail activity. If establishments tend to form retail clusters, it is probable that they try to capture customers by providing opportunities for comparison shopping or one-stop shopping. In contrast to this, if establishments are uniformly distributed, they may be homogeneous in the quality or price of goods. In this section, in order to measure the degree of retail agglomeration of individual retail types, we propose two indices calculated from an estimated PDF.

We first consider the degree of retail agglomeration within a single retail type: "do the type *i* establishments tend to be clustered or dispersed?" This can be measured by applying the Gini coefficient of the PDF. The Gini coefficient is defined as twice the area between the Lorenz curve (cumulative PDF) and the diagonal, and is often used to summarize the degree of concentration for discrete distributions (Erickson and Straussfogel, 1986; Lloyd, 1991). We redefine the Gini coefficient for continuous distributions to represent the degree of spatial concentration of PDFs.

Let us suppose the function $\delta_i(\mathbf{x}; a)$ for retail type *i* which is defined by

$$\delta_i(\mathbf{x};a) = \begin{cases} 1 & \text{if } f_i(\mathbf{x}) \le a \\ 0 & \text{if } f_i(\mathbf{x}) > a \end{cases}. \tag{8}$$

The cumulative PDF is then given by

$$F_i(\mathbf{x}) = \int_{\mathbf{t}\in S} f_i(\mathbf{t})\delta_i\left(\mathbf{t}; f_i(\mathbf{x})\right)\mathrm{d}\mathbf{t} \tag{9}$$

(note that the function $f_i(\mathbf{t})$ represents a probability density distribution). The cumulative area corresponding to $F_i(\mathbf{x})$ is

$$A(\mathbf{x}) = \int_{\mathbf{t}\in S} \delta_i\left(\mathbf{t}; f_i(\mathbf{x})\right)\mathrm{d}\mathbf{t}. \tag{10}$$

Now we define the index γ_i, the continuous version of the Gini coefficient, in order to measure the degree of spatial concentration of type *i* establishments:

$$\gamma_i = 1 - \int_{\mathbf{x}\in S} \frac{F_i(\mathbf{x})}{A(\mathbf{x})}\mathrm{d}\mathbf{x}. \tag{11}$$

The index γ_i takes a value from 0 to 1. A large γ_i represents the high degree of concentration, whereas a small value represents the dispersion or uniformity.

We next investigate the degree of type *i* establishments to be located in retail agglomerations consisting of a variety types of establishments. In other words, we try to answer the question "do the type *i* establishments tend to cluster with other types of establishments?" The index γ_i is not suitable for this question, though it tells us the degree of agglomeration within a single retail type. We hence propose another index to measure the degree of propensity for clustering with other types of retail activities.

Let us consider the PDF $f_i(\mathbf{x})$ and the average PDF $\bar{f}(\mathbf{x})$. The average PDF represents the average degree of retail agglomeration at the location **x**, that is, the relative density of all kinds of retail establishments. Therefore, if type *i* establishments tend to be located in retail agglomerations, $f_i(\mathbf{x})$ shows a positive correlation with $\bar{f}(\mathbf{x})$. In contrast to this, the PDF of establishments that prefer isolated locations has a negative correlation with $\bar{f}(\mathbf{x})$.

To describe this relationship between $f_i(\mathbf{x})$ and $\bar{f}(\mathbf{x})$ quantitatively, we calculate the Bhattacharyya distance (Fukunaga, 1972; Basseville, 1989) between these functions, and define the index η_i which represents the degree of propensity for agglomerations of type *i* establishments by the equation below.

$$\eta_i = -\log\left\{\int_{\mathbf{x}\in S}\sqrt{f_i(\mathbf{x})\bar{f}(\mathbf{x})}\mathrm{d}\mathbf{x}\right\} \tag{12}$$

The index η_i shows a large value if the establishments prefer retail agglomerations. This index is comparable among different retail types in *S* because the integration of $f_i(\mathbf{x})$ in *S* is always equal to 1.

We should note that the indices γ_i and η_i are defined without spatial aggregation based on shopping areas or spatial units. Theoretically they are given by integrating functions in the whole region. Though the calculation usually requires numerical integration, we can decrease numerical errors by using fine grids in the calculation to obtain almost true values.

The index values calculated for the individual retail types are shown in Table 2. Let us first examine the index γ_i which represents the degree of agglomerative tendency within type *i* establishments. We notice that apparel activities such as men's clothing and accessory show large values. This suggests that spatial agglomeration increases their benefit, and supports that they are comparison-shopping activities. On the other hand, transport related activities such as car parts, motorcycle, and gasoline station have small values. This implies that these kinds of establishments are widespread in Yokohama. Food stores have a wide variation in γ_i, and there we cannot find any significant pattern. Though restaurants also show diversity in γ_i, there seems a tendency that expensive restaurants such as Japanese-style restaurant, sushi restaurant, and steakhouse are more concentrated than inexpensive restaurants such as Chinese noodle, pizza, and Japanese noodle.

Let us move to the index η_i. Table 2 seems to indicate a close relationship between γ_i and η_i. Retail types having large γ_i such as apparel activities again show large η_i, whereas those of small γ_i have small η_i. The Spearman's correlation coefficient between the two indices is 0.9098, which is significant at the level of 1%. This implies that retail activities that tend to gather within the same type of establishments also gather with other types of establishments. Apparel activities typically shows this tendency. In contrast to this, we can say that retail activities that avoid agglomeration with the same type of establishments do not prefer agglomeration with other types of establishments. Transport related activities show such a tendency.

Table 2 Agglomerative tendencies of retail establishments.

5 Spatial structure of retailing

The indices proposed in the previous section reveal agglomerative tendencies of retail establishments, and they are useful for analyzing individual retail activities. However, they are not enough to understand the whole structure of retailing composed of a variety of retail activities. We hence propose a method for analyzing the spatial structure of retailing in this section.

5.1 Classification of retail distributions

The spatial structure of retailing consists of numerous types of retail activities in metropolises as seen Section 2, and there is a very wide diversity of spatial distributions. Gasoline stations, for instance, are located along main roads and highways. Apparel activities, as seen in the previous section, prefer retail agglomeration. Department stores are located in downtowns.

In the analysis of the retail structure in metropolises, a huge amount of information to be handled often conceals spatial patterns to be detected. It is thus desirable to categorize retail distributions into some classes in advance in order to reduce the information and make the underlying patterns recognizable. To this end, we propose a method for classifying retail distributions using PDF which helps us to detect typical patterns of retail distributions.

Classification method usually requires a measure of similarity or dissimilarity between elements to be classified. As the measure of dissimilarity between retail distributions, we employ the Jeffrey-Matsushita distance (Fukunaga, 1972; Basseville, 1989) between their PDFs. Suppose the PDFs of retail types i and j, that is, $f_i(\mathbf{x})$ and $f_j(\mathbf{x})$. The Jeffrey-Matsushita distance between $f_i(\mathbf{x})$ and $f_j(\mathbf{x})$ is defined by

$$JM_{ij} = \sqrt{\int_{\mathbf{x}\in S}\left\{\sqrt{f_i(\mathbf{x})} - \sqrt{f_j(\mathbf{x})}\right\}^2 \mathrm{d}\mathbf{x}}\,. \qquad (13)$$

If the types i and j are similar in the retail distribution, JM_{ij} shows a small value. Thus, we can classify retail distributions by the cluster analysis where the distance between elements is given by the Jeffrey-Matsushita distance (Romesburg, 1989; Everitt, 1993).

We performed the unweighted pair-group clustering method using arithmetic averages (UPGMA), and obtained the dendrogram shown in Figure 8. The structure of the dendrogram appears to indicate six major classes (the dotted line). We hence classified the 78 retail categories into six types: location types L-A_1, L-A_2, L-A_3, L-B_1, L-B_2, and L-C. Figure 9 depicts the value of indices γ_i and η_i categorized by the location type. In addition to this, in order to examine the spatial characteristics of each location type, we calculated the average PDF of each type and draw Figure 10 showing the regions where the average cumulative PDF is equal to 0.95.

Figure 8 The dendrogram for retail distributions.

Figure 9 The indices γ_i and η_i categorized by the location type.

Figure 10 Regions where the average cumulative PDF is equal to 0.95. a) overlaid with the railway network, b) overlaid with the road network.

Spatial characteristics of each location type are summarized as follows.

Types L-A_1, L-A_2, and L-A_3: Types L-A_1, L-A_2, and L-A_3 are very similar in the degree of concentration as shown in Figure 9. They tend to gather not only within the same but different types of stores. They are also similar in the propensity for railway stations. The retail activities of these types are in higher order, that is, relatively expensive and infrequently purchased. A key to distinguish these types lies in the stations surrounded by the establishments; the L-A_1 establishments are highly concentrated around railway stations of a great number of passengers such as Yokohama, Kannai, Tsurumi, Tamapuraza, and Kamioka; the L-A_2 establishments can be found only around stations in the central Yokohama, that is, Yokohama, Kannai, and Tsurumi; the L-A_3 establishments are located around not only the stations mentioned above, but also those of fewer number of passengers such as Mitsuzakai, Futamatagawa, Shinsugita, and Kanazawa-Bunko.

Type L-B_1: Type L-B_1 shows less degree of concentration than Types L-A_1, L-A_2, and L-A_3. The establishments are located around railway stations that are very similar to those preferred by the L-A_3 establishments. Type L-B_1 activities are distinctive in that they are not frequently used though not so expensive as those of Types L-A_i.

Type L-B_2: The L-B_2 establishments are less concentrated than all of those mentioned above. Retail activities in the upper column of Type L-B_2 in Figure 8, say, camera, sporting goods, and toy, show preference for railway stations. Retail activities in the lower column, say, meat, green grocery, and pharmacy do not show such a clear tendency. Type L-B_2 is a mixture of higher and lower order retail activities.

Type L-C: Type L-C is distinctive for its spatial dispersion. The L-C establishments are distributed over the whole area of Yokohama. Retail activities in the upper column of Type L-C in Figure 8 tend to gather around railway stations, whereas those in the lower column are located along main roads and highways. This type consists of retail activities much varying in order.

The above results indicate that retail activities similar in goods price or purchase frequency do not always show a similar distribution. Though retail activities of types L-A_i are consistent to some extent, other types consist of a variety of activities ranging from lower to higher order. This implies that the location of stores is fairly affected by other factors such as the quality of goods, agglomeration effects, and so forth.

Comparing Figures 10a and 10b, we notice that the road network does not influence the spatial structure of retailing as much as railway stations. Though one might think this is very strange, it is usual with big cities in Japan. This is mainly because the railway system works as a major transportation system in Yokohama and other big cities, and the road network is of secondary importance. Retail stores concentrate around railway stations where numerous people gather.

5.2 Relationship between the agglomeration degree of establishments and retail function

Relationship between the size and function of retail agglomerations has drawn much attention in geography (Potter, 1982; Morrill, 1987; Jones and Simmons, 1990; Brown, 1992; Okabe and Sadahiro, 1996). In metropolitan areas, however, the size of retail agglomeration is not measurable in a strict sense because of the indeterminate boundaries of shopping areas. We hence employ the degree of retail agglomeration instead, and analyze it in relation to the retail function. Though this approach is not the same as used in existing studies, it will provide a useful information because a close relationship usually exists between the degree and the size of retail agglomeration as mentioned earlier.

Let us denote the summation of Type L-Θ PDFs at $\mathbf{x}$ as $f_\Theta(\mathbf{x})$. Consider the ratio of $f_\Theta(\mathbf{x})$ to the summation of all PDFs, that is,

$$r(\Theta, \mathbf{x}) = \frac{f_\Theta(\mathbf{x})}{\sum_i f_i(\mathbf{x})}. \qquad (14)$$

We investigate the relationship between the ratio $r(\Theta, \mathbf{x})$ and the average PDF $\bar{f}(\mathbf{x})$ which represents the degree of retail agglomeration.

Figure 11 shows the relationship between $\bar{f}(\mathbf{x})$ and $r(\Theta, \mathbf{x})$. It is clear that there is a close relationship between the degree of agglomeration and retail function, though it is continuous rather than discrete. When the average PDF $\bar{f}(\mathbf{x})$ is less than 1.0×10^{-5}, Type L-C activities are predominant. This is supported by the fact that Type L-C shows small values of γ_i and η_i, which indicates the propensity for isolated locations. As $\bar{f}(\mathbf{x})$ increases, the ratio of Type L-C decreases consistently and Types L-B_1 and L-B_2 appear successively. This suggests that local shopping centers whose $\bar{f}(\mathbf{x})$ is under $1.0*10^{-3.5}$ consist of these types of retail activities (see also Figures 7 and 8). When $\bar{f}(\mathbf{x})$ exceeds $1.0*10^{-3.5}$, Type L-A_2 appears first, and then do Types L-A_3 and L-A_1. The ratios of Types L-B_1, L-A_2, and L-A_1 increase consistently with $\bar{f}(\mathbf{x})$. At the highest level retail agglomerations mainly consist of the L-B_1, L-A_2, and L-A_1 establishments.

Figure 11 Relationship between the average PDF $\bar{f}(\mathbf{x})$ and the ratio of PDFs $r(\Theta, \mathbf{x})$.

5.3 Spatial structure of retail agglomeration

Having analyzed the relationship between the size and function of retail agglomerations, we then proceed to the spatial structure of retail agglomerations. How many levels of agglomerations do exist? How do they form the retail structure in a metropolitan area? We propose an analyzing method to answer these questions.

We first divide the region S into a square lattice. Let us denote the region covered by cell i as C_i. We then calculate the ratio of $f_\Theta(\mathbf{x})$ to the summation of all PDFs in C_i as

$$\bar{r}(\Theta, C_i) = \frac{\int_{\mathbf{x} \in C_i} r(\Theta, \mathbf{x}) d\mathbf{x}}{\int_{\mathbf{x} \in C_i} d\mathbf{x}}. \tag{15}$$

As seen in the above equation, the ratio $\bar{r}(\Theta, C_i)$ is the average of the ratio $r(\Theta, \mathbf{x})$ in cell i.

We next perform the UPGMA cluster analysis to classify the cells by the retail functions. The similarity between cells i and j is defined by

$$d_{ij} = \sqrt{\sum_{\Theta} \left\{ \bar{r}(\Theta, C_i) - \bar{r}(\Theta, C_j) \right\}^2}. \tag{16}$$

Having obtained classified cells, we then examine their spatial distribution.

We applied the above method to the retail distributions in Yokohama. For the size of cells, we tried various sizes ranging from 200 to 1000 meters, and found that the cell size does not greatly affect the results. We thus describe the result based on the square lattice of side 300 meters.

The obtained dendrogram for the cells is illustrated in Figure 12. We show only a part of the dendrogram where seven major classes are indicated. The probability distributions of the average PDF are depicted in Figure 13. This figure indicates that the C-C, C-D, C-E, C-F, and C-G cells are located where the average PDF is fairly small. Examining the cells in detail, we found that most of them contain only one or two isolated stores. We thus excluded these cells in further analysis in order to focus on retail agglomerations.

Figure 12 The dendrogram for cells.

Figure 13 The probability distributions of the average PDF.

To analyze the characteristics of retail agglomerations more in detail, we reclassified the Types C-A and C-B cells into 3 subclasses respectively: C-A_1, C-A_2, C-A_3, C-B_1, C-B_2, and C-B_3 (Figure 14). We then calculated the average of $\bar{f}(\mathbf{x})$ and $\bar{r}(\Theta, C_i)$ for each cell type (Table 3). The spatial distribution of the cells is shown in Figure 15.

Figure 14 The dendrogram for cells representing retail agglomerations.

Figure 15 The spatial distribution of the cells representing retail agglomerations.

Table 3 The averages of the average PDF $\bar{f}(\mathbf{x})$ and the ratio $\bar{r}(\Theta, C_i)$ (%).

From Figure 15 and Table 3, we summarize the characteristics of each cell type as follows.

Type C-A_1: Type C-A_1 corresponds to entertainment districts. This type is distinctive for the highest degree of agglomeration (the average PDF) and the high ratios of the L-A_2 and L-B_1 retail activities, especially taverns, bars, Japanese-style restaurants (which serve alcoholic drinks), video games, and mahjong games. The C-A_1 cells are clustered close to railway stations such as Yokohama, Kannai, and Shin-Yokohama, which are surrounded by business districts.

Type C-A_2: In contrast to Type C-A_1, Type C-A_2 is distinctive for the smallest average PDF. Examining the C-A_2 cells in detail, we found that most of them contain only a few establishments. The L-B_1 and L-C activities are predominant. The C-A_2 cells are scattered in the suburban area of Yokohama, and they prefer residential areas at a distance from railway stations. Considering the above, we regard the C-A_2 cells as very small retail clusters in residential areas.

Type C-A_3: Type C-A_3 mainly corresponds to unplanned local shopping areas around railway stations. Many of C-A_3 cells are concentrated around railway stations, and others surround the C-A_1 and C-B_3 cells. The ratio of lower level activities, say, L-B_1, L-B_2, and L-C, seems relatively high. Those indicate that Type C-A_3 represents traditional unplanned retail agglomerations frequently found around railway stations in Japan.

Type C-B₁: Numerical indices of Type C-B_1 are similar to those of Type C-A_3, though the average PDF is somewhat smaller and the ratios of Types L-A_3 and L-B_1 activities are somewhat different. A clear distinction between the two cell types can be made by their spatial distributions. While the C-A_3 cells are highly concentrated around stations, the C-B_1 cells are scattered over the whole area of Yokohama off railway stations. Another difference between the cell types is in the size of cell clusters: the C-B_1 cells form smaller clusters than those of the C-A_3 cells. We hence consider that Type C-B_1 represents local retail agglomerations in residential areas, which are also frequently found in Japan.

Type C-B_2: Type C-B_2 is very similar to Type C-B_1 both in the numerical attributes and the spatial distribution. The difference between them lies in the composition of retail activities. Many of C-B_1 cells contain either a green grocery store or a pharmacy, whereas most of C-B_2 cells have a beauty shop. This suggests that green grocery stores and pharmacies worked as the nucleus of the C-B_1 agglomerations, and so did pharmacies for the C-B_2 agglomerations.

Type C-B_3: Type C-B_3 corresponds to the highest level of retail agglomerations. The degree of agglomeration is considerably high, and the ratio of L-A_1 activities shows the largest among the six cell types. The spatial distribution of the C-B_3 cells is also distinctive. They are concentrated only around the stations having a great number of passengers, that is, Yokohama, Kannai, Shin-Yokohama, Tamapuraza, Konandai, and Kanazawa-Bunko.

Besides the above types of cells, the C-C, C-D, and C-E cells often contain one or two isolated establishments. We thus conclude that there are four levels of retail agglomerations and isolated establishments in Yokohama:

1) 1-st level agglomerations (Type C-B_3): Located around the biggest railway stations. Contain department stores and women's clothing stores. Highest degree of agglomeration. Often accompanied by entertainment-specialized areas (Type C-A_1), or the 2-nd level agglomerations (Type C-A_3).
2) 2-nd level agglomerations (Type C-A_3): Located around most of the railway stations and the 1-st level agglomerations. Consist of a wide variety of stores from shoppers goods to convenience goods. The second highest degree of agglomeration.
3) 3-rd level agglomerations (Types C-B_1 and C-B_2): Distributed in residential areas off the railway stations. Usually contain a green grocery store, a pharmacy, or a beauty shop as the nucleus. Frequently purchased goods.

4) 4-th level agglomerations (Type C-A_2): Scattered in the suburban area of Yokohama. Consist of a few establishments.
5) Isolated establishments (Types C-C, C-D, and C-E): Scattered over the whole area of Yokohama. Some prefer main roads and highways.

The spatial hierarchy of retailing system in Yokohama seems significant, though its pattern is not so regular as that of the central place theory. The 1-st and 2-nd level agglomerations are fewer than those of lower levels, and located only around bigger stations. The 3-rd and 4-th level agglomerations are dispersed over the whole area of Yokohama.

The relationship between levels in the hierarchy and types of store is fairly clear. Higher level agglomerations mainly consist of stores of expensive and infrequently purchased goods, whereas lower level agglomerations consist of shoppers goods stores. The hierarchical system is similar to Berry's typology of shopping areas to some extent: general shopping areas, ribbon developments, and specialized shopping areas. The 1-st and 2-nd level agglomerations seem to correspond the general shopping areas, whereas the 3-rd and 4-th levels are the ribbon developments. However, specialized shopping areas are not detected, and isolated establishments do not have their counterpart in Berry's typology. The hierarchy of retailing system in Yokohama does not perfectly agree with that of 1960's. On the other hand, it does not completely support the Brown's typology proposed in 1991 (Brown, 1991). The system of Yokohama does not have planned shopping centers such as megastores and speciality shopping areas which are often found in suburban areas in the UK and the USA. This is mainly because railway system works as a major transportation system in Yokohama, as stated earlier. We can conclude that the retailing system in Yokohama is a hybrid of systems advocated by Berry and Brown.

6 Conclusions

In the present paper, we have proposed a method for analyzing the retail structure in a metropolis and applied it to the retail structure in Yokohama, Japan. The method is based on the probability distribution function (PDF) of establishments estimated from the their locational data, and it is better than the existing methods using spatial aggregation data in several points as discussed earlier.

The two indices γ_i and η_i drawn from PDF were proved to be useful for summarizing agglomerative tendencies of retail establishments. Though a PDF itself fully describes the spatial characteristics of establishments, the summary indices are more convenient especially in comparing a wide variety of retail distributions.

The classification of retail distributions also worked well in the empirical study: it revealed the six major types of location in Yokohama. Though we knew by experience that there are some typical patterns in retail distributions, there are few methods for investigating it quantitatively. The PDF-based method provides a useful tool for the classification of retail distributions.

The spatial structure of retail agglomerations was then analyzed by use of the cells classified by the retail functions. The classification of cells indicated four levels of retail agglomerations and isolated stores. The spatial structure of the agglomerations showed a close relationship with the location of railway stations. Though the result may not be so surprising, the proposed method appears to be valid for analyzing retail agglomerations.

Finally we should note some limitations of the present study. First, in the empirical study we used only the locational information of retail establishments, without taking account of other attributes such as the floor size, the quality and price of goods, and so forth. That is because such detailed data are unavailable in general. However, if detailed data are accessible, the proposed method can be easily modified to take various factors into account. The floor size, for instance, can be reflected in the weight of the kernels. This assures us applications of the method in other countries where large supermarkets and megastores are predominant and the size factor is of great importance. The data of goods types permit us to classify the establishments more strictly or to analyze directly the distribution of individual goods through the same procedure as used in the present paper. Second, a quantitative method should be developed to analyze the spatial relationship among the distributions of retail establishments, railway stations, and residential areas. The present paper analyzed these relationships only by qualitative methods, thus it remains unknown whether they are really meaningful. A new quantitative method to treat these relationships should be developed in further researches.

References

Applebaum W, 1966, "Methods for Determining Store Trade Areas, Market Penetration, and Potential Sales" *Journal of Marketing Research* **3** 127-141

Ashworth G J, De Vries J, 1985, "Delimiting the central areas of west European cities: the example of Colmar" *Geograficky Casopis* **37** 61-79

Bailey T C, Gatrell A C, 1995 *Interactive Spatial Data Analysis* (Longman, Harlow)

Berry B J L, 1967 *Geography of Market Centers and Retail Distribution* (Prentice-Hall, Englewood Cliffs)

Bohman-Carter G F, 1994 *Geographic Information Systems for Geoscientists (*Pergamon, Oxford)

Bowman A W, 1984, "An alternative method of cross-validation for the smoothing of density estimates" *Biometrica* **71** 353-360

Bracken I, 1993, "An extensive surface model database for population-related information: concept and application" *Environment and Planning B*, **20**, 13-27

Bracken I, Martin D, 1989, "The generation of spatial population distributions from census centroid data" *Environment and Planning A*, **21**, 537-543

Bracken I, Martin D, 1995, "Linkage of the 1981 and 1991 UK Censuses using surface modelling concepts" *Environment and Planning A*, **27**, 379-390

Brown S, 1991, "Retail location: the post hierarchical challenge" *International Review of Retail Distribution and Consumer Research*, **1**, 367-381

Brown S, 1992 *Retail Location: A Micro-Scale Perspective* (Avebury, Aldershot)

Carter H, Rowley G, 1966, "The morphology of the central business district of Cardiff" *Transactions of the Institute of British Geographers* **38** 119-134

Clark W A V, 1967, "The spatial structure of retail functions in a New Zealand city" *New Zealand Geographer* **23** 23-33

Davies R L, Rogers D, 1984 *Store Location and Store Assessment Research* (John Wiley & Sons, New York)

Erickson R A, Straussfogel D L, 1986, "The spatial patterns of employment change in large American metropolises: 1947-1977" *Urban Geography*, **7**, 385-396

Everitt B S, 1993 *Cluster Analysis* (Arnold, London)

Fiacco A V, McCormick G P, 1968 *Nonlinear Programming: Sequential Unconstrained Minimization Technique* (John Wiley & Sons, New York)

Fotheringham A S, Rogerson P, 1994 *Spatial Analysis and GIS* (Taylor and Francis, London)

Getis A, Getis J M, 1968, "Retail store spatial affinities" *Urban Studies* **5** 317-332

Gill P E, Murry W, Wright M H, 1981 *Practical Optimization* (Academic Press, London)

Jones K, Simmons J, 1990 *The Retail Environment* (Routledge, London)

Lee Y, Koutsopoulos K, 1976, "A locational analysis of convenience food stores in metropolitan Denver" *Annals of Regional Science* **10** 104-117

Leeming F A, 1959, "An experimental survey of retail shopping and service facilities in part of North Leeds" *Transactions and Papers of the Institute of British Geographers* **26** 133-152

Logan A, 1968 *The Pattern of Service Centers in Warringah Shire* (University of Sydney Planning Research Centre, Sydney)

Longley P, Clarke G, 1995 *GIS for Business and Service Planning* (Cambridge, GeoInformation International)

Lloyd W J, 1991, "Changing suburban retail patterns in metropolitan Los Angeles" *Professional Geographer* **43** 335-344

McEvoy D, 1972, "Vacancy rates and the retail structure of cities", in *The Retail Structure of Cities*, Institute of British Geographers, Occasional Publication No. 1, London, 59-68

Miki F, 1980, "Analysis of the retail distribution of the convenience goods" *Papers of the Annual Conference of the City Planning Institute of Japan* **15** 157-162 (in Japanese)

Miki F, 1983, "A study of the relationship between population distribution and retail distribution of the convenience goods" *Papers of the Annual Conference of the City Planning Institute of Japan* **18** 19-24 (in Japanese)

Morrill R, 1987, "The Structure of Shopping in a Metropolis" *Urban Geography* **8** 97-128

Okabe A, Boots B, Sugihara K, 1992 *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams* (John Wiley & Sons, New York)

Okabe A, Miki F, 1984, "A conditional nearest-neighbor spatial-association measure for the analysis of locational interdependency" *Environment and Planning A* **16** 163-171

Okabe A, Sadahiro Y, 1994, "A statistical method for analyzing the spatial relationship between the distribution of activity points and the distribution of activity continuously distributed over a region" *Geographical Analysis* **26** 152-167

Okabe A, Sadahiro Y, 1996, "An illusion of spatial hierarchy: spatial hierarchy in a random configuration" *Environment and Planning A* **28** 1533-1552

Openshaw S, 1984 *The Modifiable Areal Unit Problem* (Geobooks, Norwich)

Potter R B, 1982 *The Urban Retailing System: Location, Cognition and Behaviour* (Gower, Aldershot)

Romesburg H C, 1989 *Cluster Analysis for Researchers*. (Krieger, Malabar)

Rudemo M, 1982, "Empirical choice of histograms and kernel density estimators"

Scandinavian Journal of Statistics **9** 65-78

Sadahiro Y, 1994, "A statistical method for analyzing the relationship between the spatial distribution of retail stores and that of railway stations" *Papers on City Planning* **29** 523-528 (in Japanese)

Scott P, 1959, "The Australian CBD" *Economic Geography* **35** 290-314

Scott P, 1970 *Geography and Retailing* (Hutchinson, London)

Scott D W, 1992 *Multivariate Density Estimation* (John Wiley & Sons, New York)

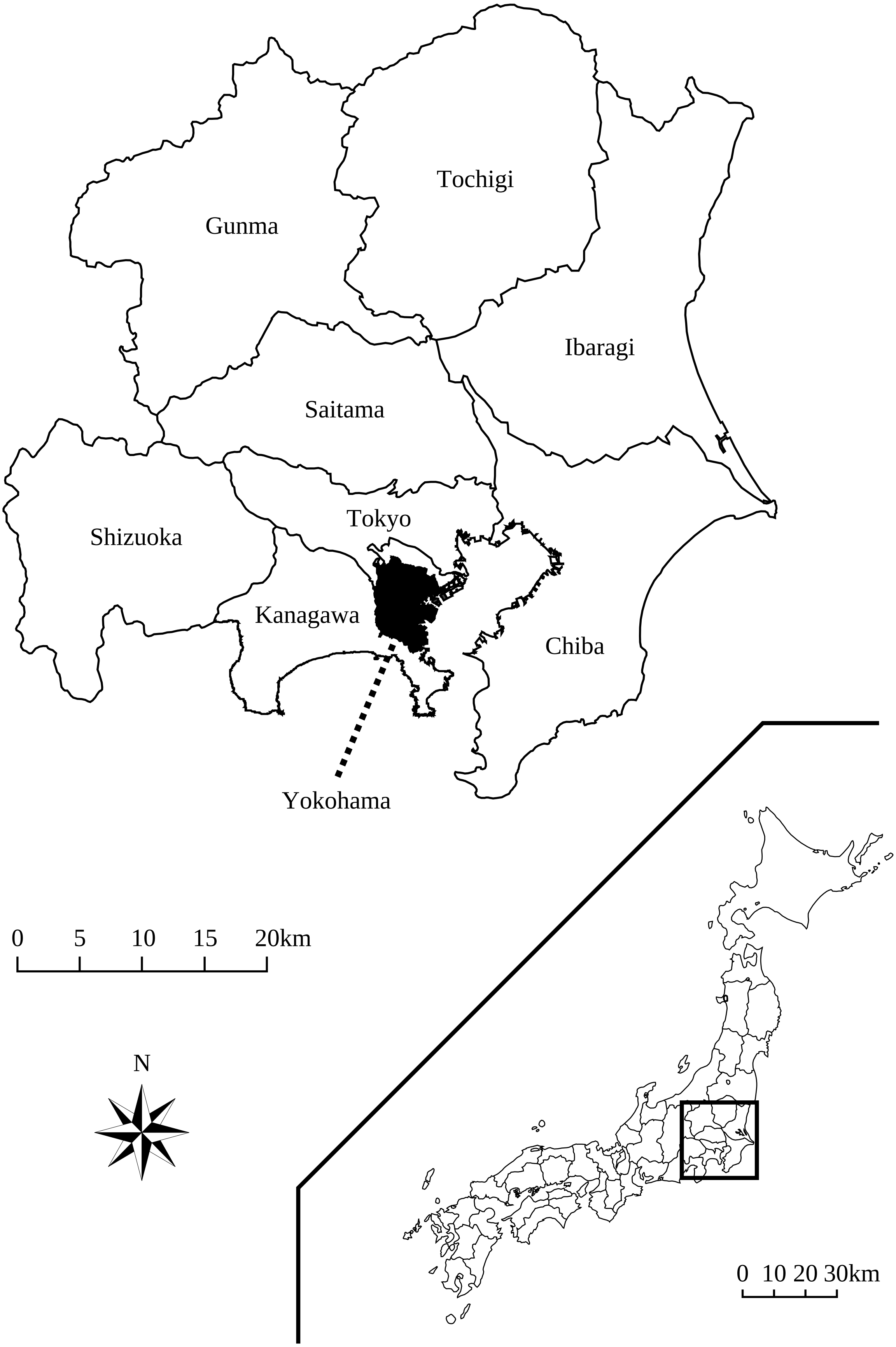

Tochigi
Gunma
Ibaragi
Saitama
Tokyo
Shizuoka
Kanagawa
Chiba
Yokohama
0 5 10 15 20km
N
0 10 20 30km

Figure 1

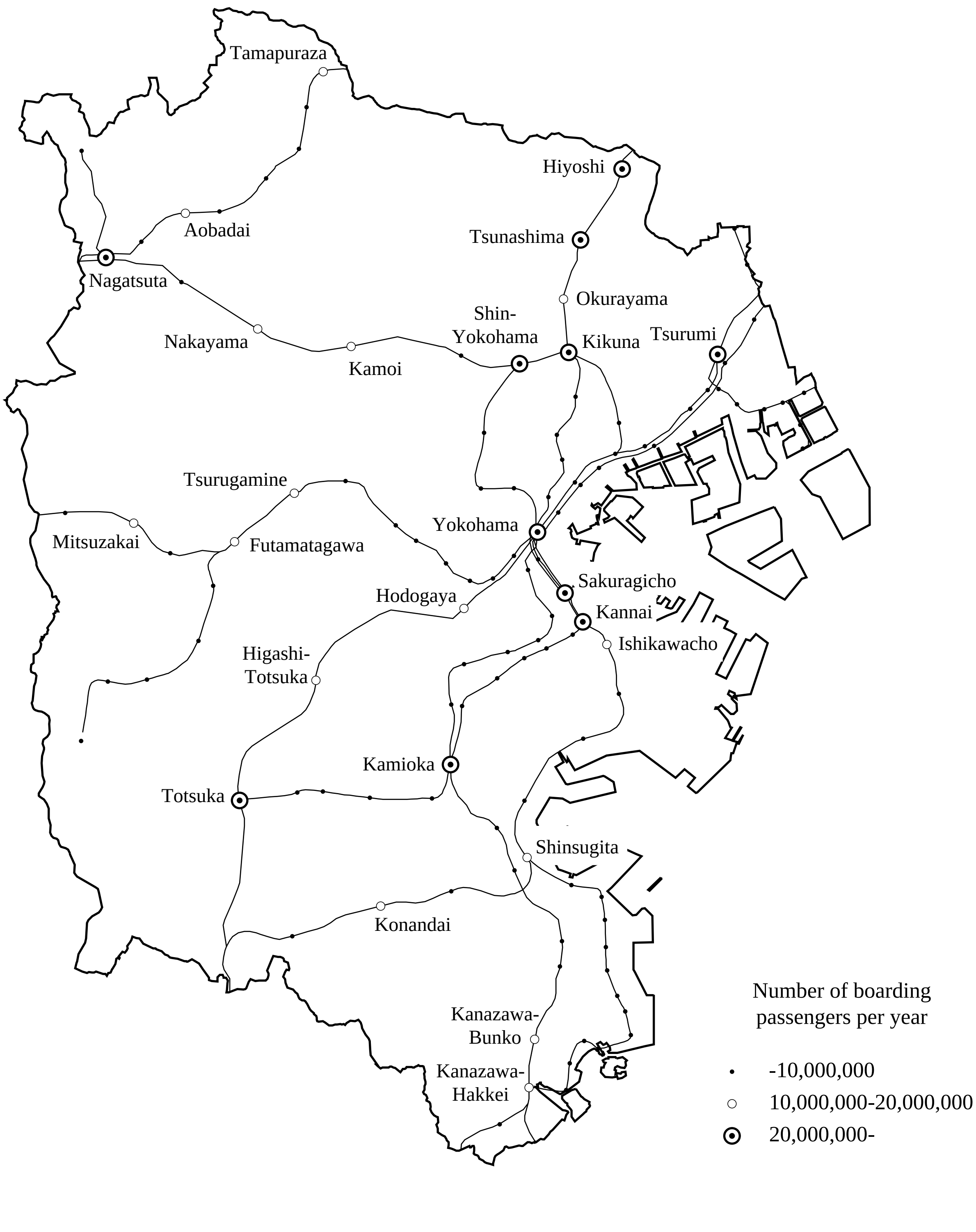

Tamapuraza
Hiyoshi
Aobadai
Tsunashima
Nagatsuta
Okurayama
Shin-
Yokohama
Nakayama
Kikuna
Tsurumi
Kamoi
Tsurugamine
Yokohama
Mitsuzakai
Futamatagawa
Sakuragicho
Hodogaya
Kannai
Ishikawacho
Higashi-
Totsuka
Kamioka
Totsuka
Shinsugita
Konandai
Kanazawa-
Bunko
Kanazawa-
Hakkei
Number of boarding
passengers per year
-10,000,000
10,000,000-20,000,000
20,000,000-

Figure 2

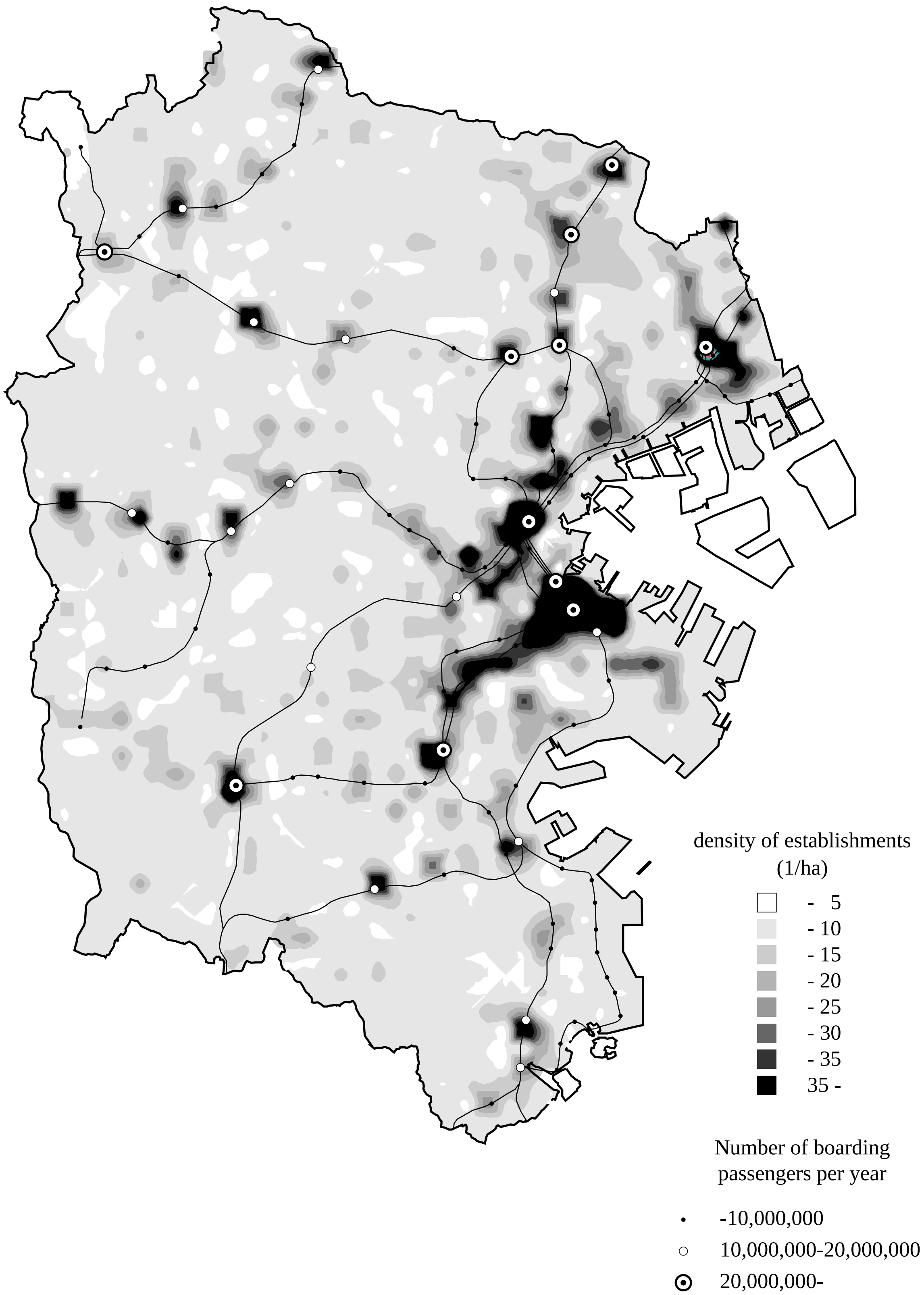

density of establishments
(1/ha)
- 5
- 10
- 15
- 20
- 25
- 30
- 35
35 -
Number of boarding
passengers per year
-10,000,000
10,000,000-20,000,000
20,000,000-

Figure 3

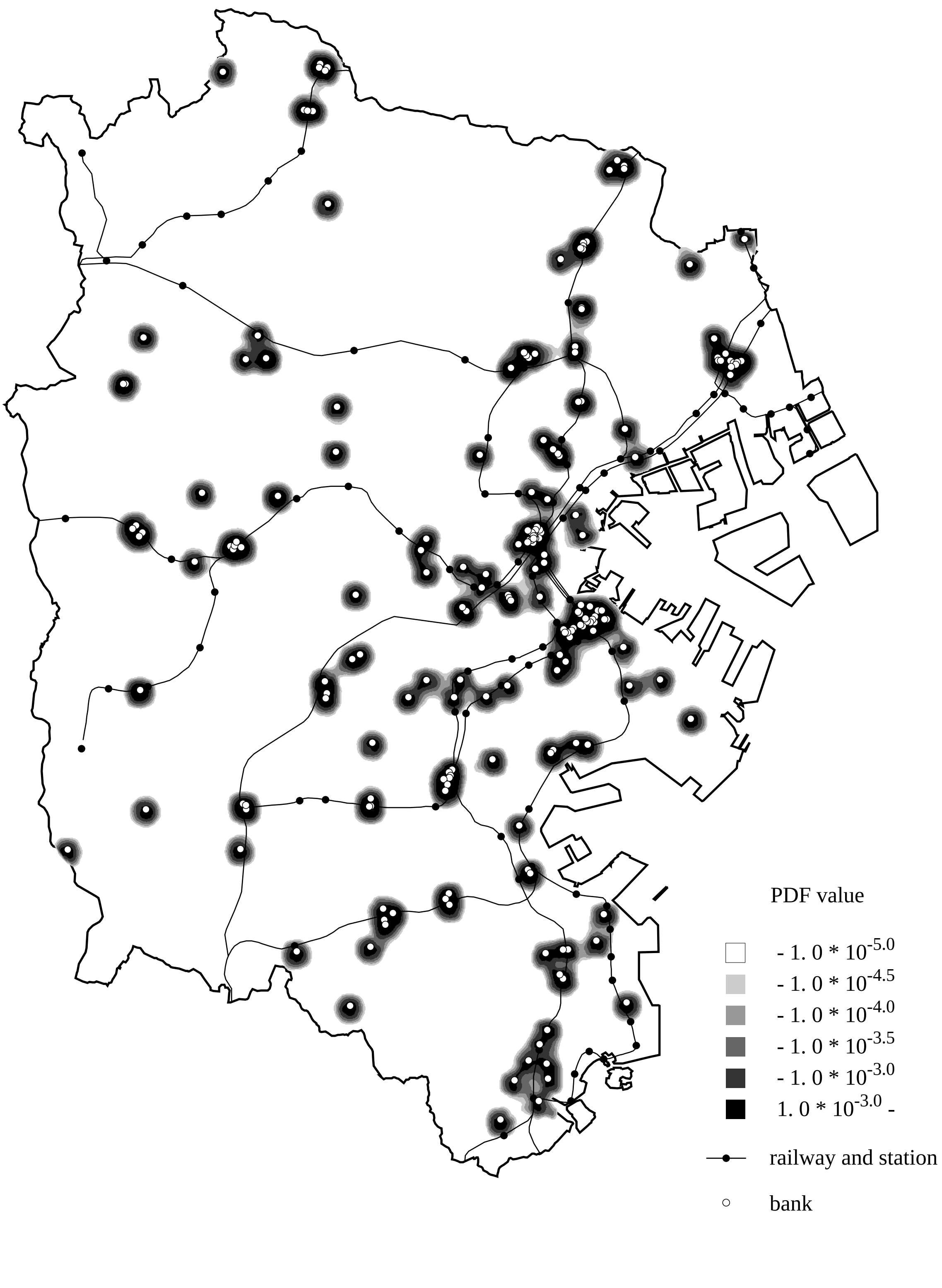

PDF value
- 1. 0 * $10^{-5.0}$
- 1. 0 * $10^{-4.5}$
- 1. 0 * $10^{-4.0}$
- 1. 0 * $10^{-3.5}$
- 1. 0 * $10^{-3.0}$
1. 0 * $10^{-3.0}$ -
railway and station
bank

Figure 4

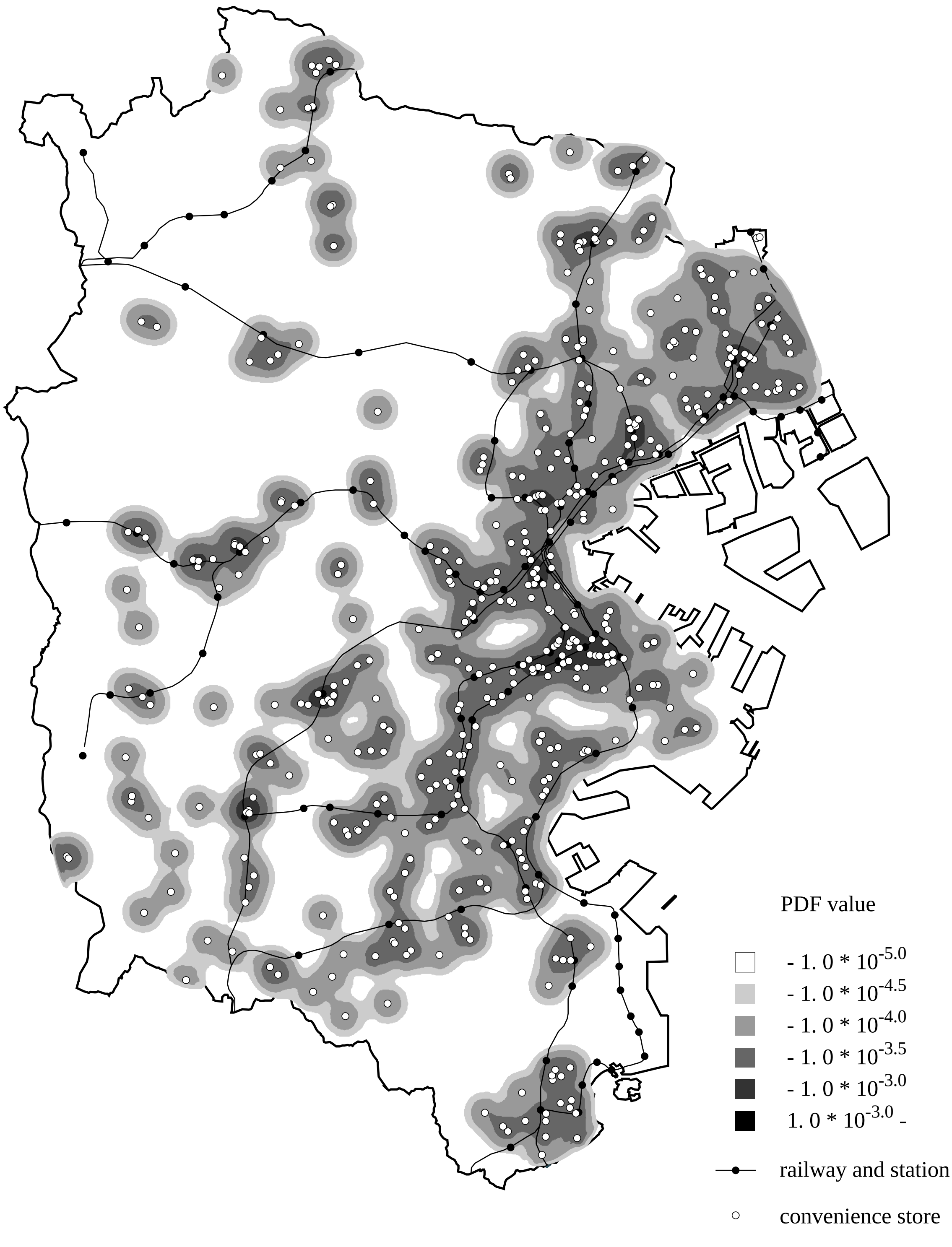

PDF value
- 1. 0 * $10^{-5.0}$
- 1. 0 * $10^{-4.5}$
- 1. 0 * $10^{-4.0}$
- 1. 0 * $10^{-3.5}$
- 1. 0 * $10^{-3.0}$
1. 0 * $10^{-3.0}$ -
railway and station
convenience store

Figure 5

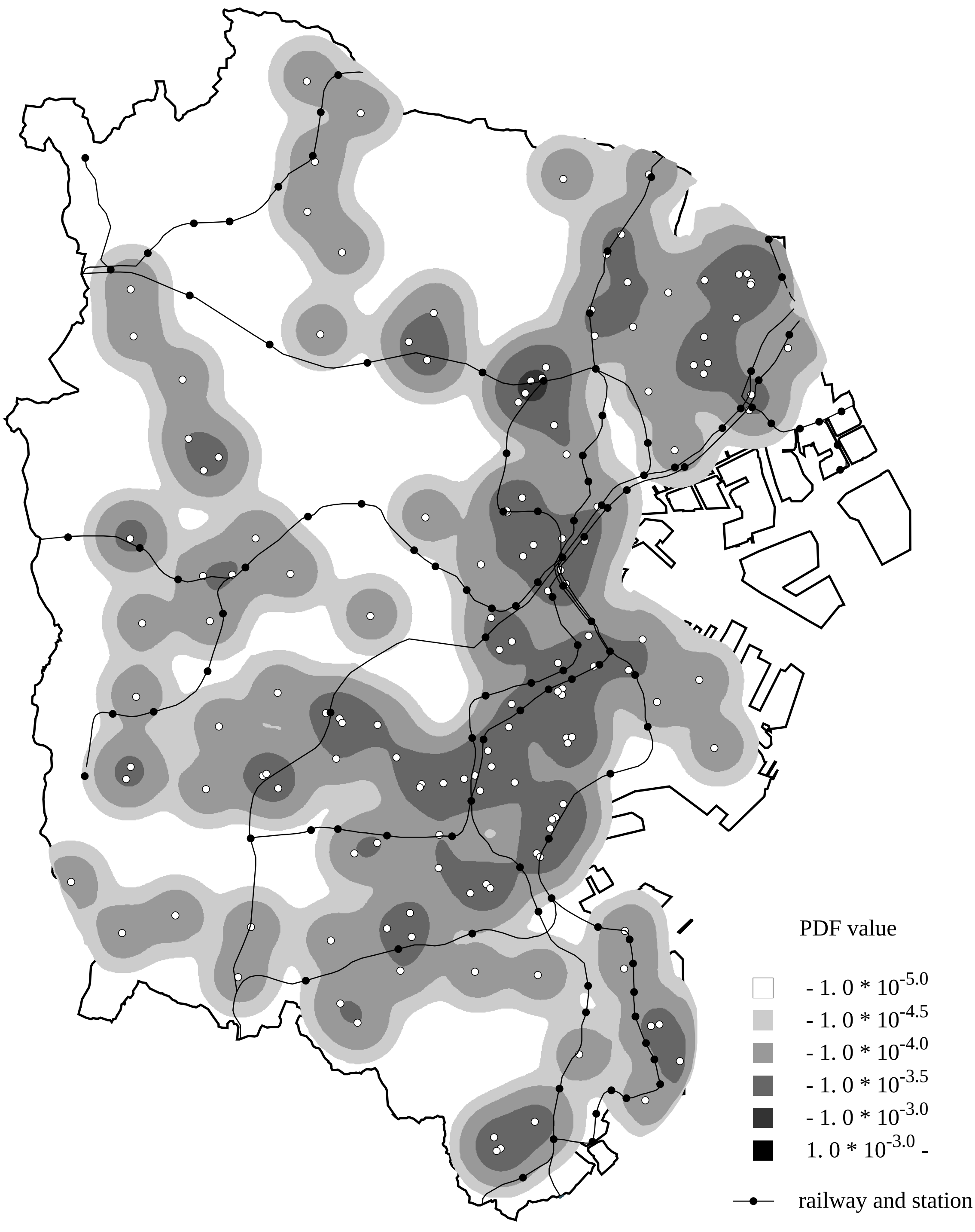

PDF value
- 1. 0 * 10^-5.0
- 1. 0 * 10^-4.5
- 1. 0 * 10^-4.0
- 1. 0 * 10^-3.5
- 1. 0 * 10^-3.0
1. 0 * 10^-3.0 -
railway and station
used car shop

Figure 6

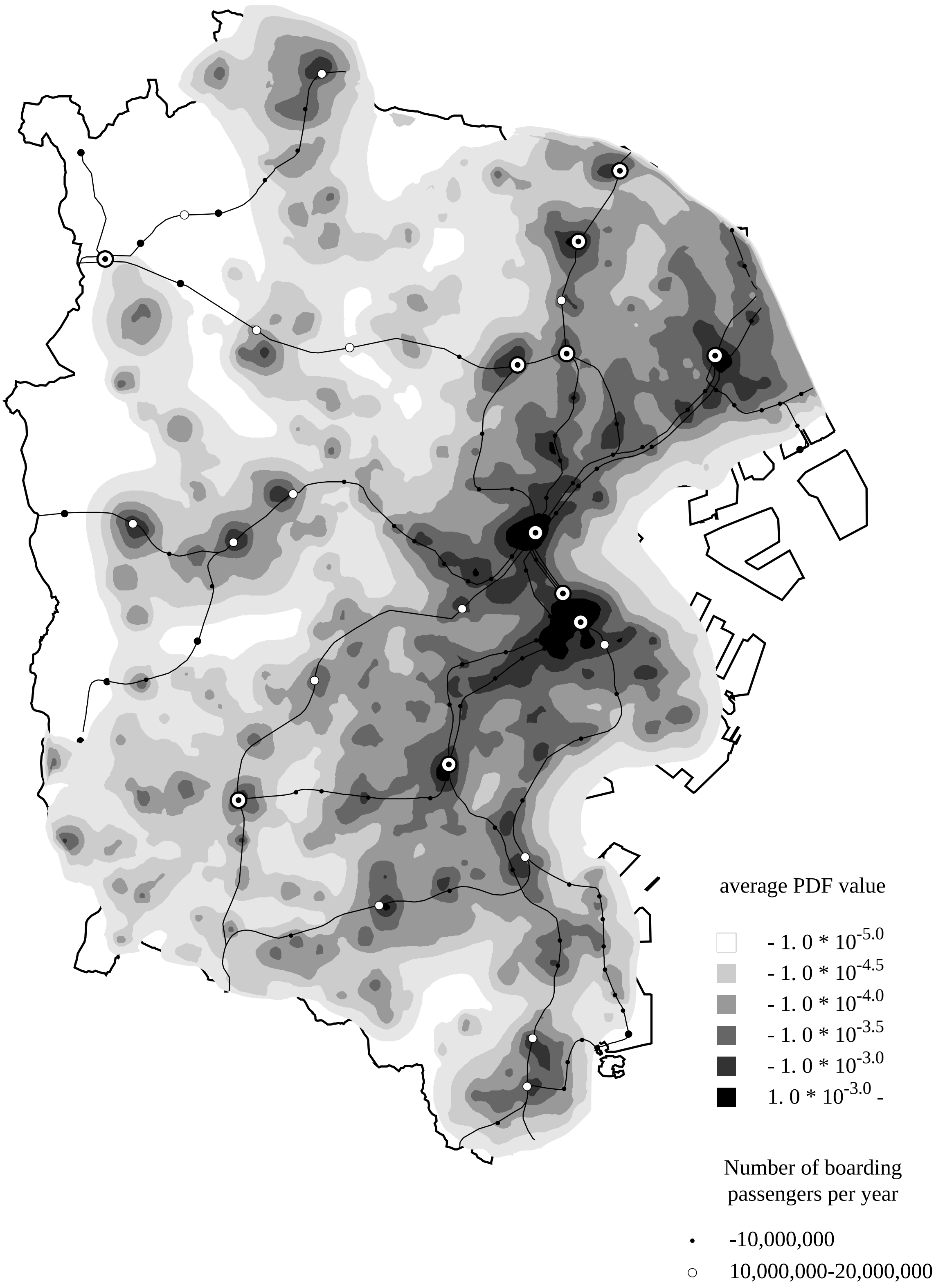

average PDF value
- 1. 0 * 10^-5.0
- 1. 0 * 10^-4.5
- 1. 0 * 10^-4.0
- 1. 0 * 10^-3.5
- 1. 0 * 10^-3.0
1. 0 * 10^-3.0 -
Number of boarding passengers per year
-10,000,000
10,000,000-20,000,000
20,000,000-

Figure 7

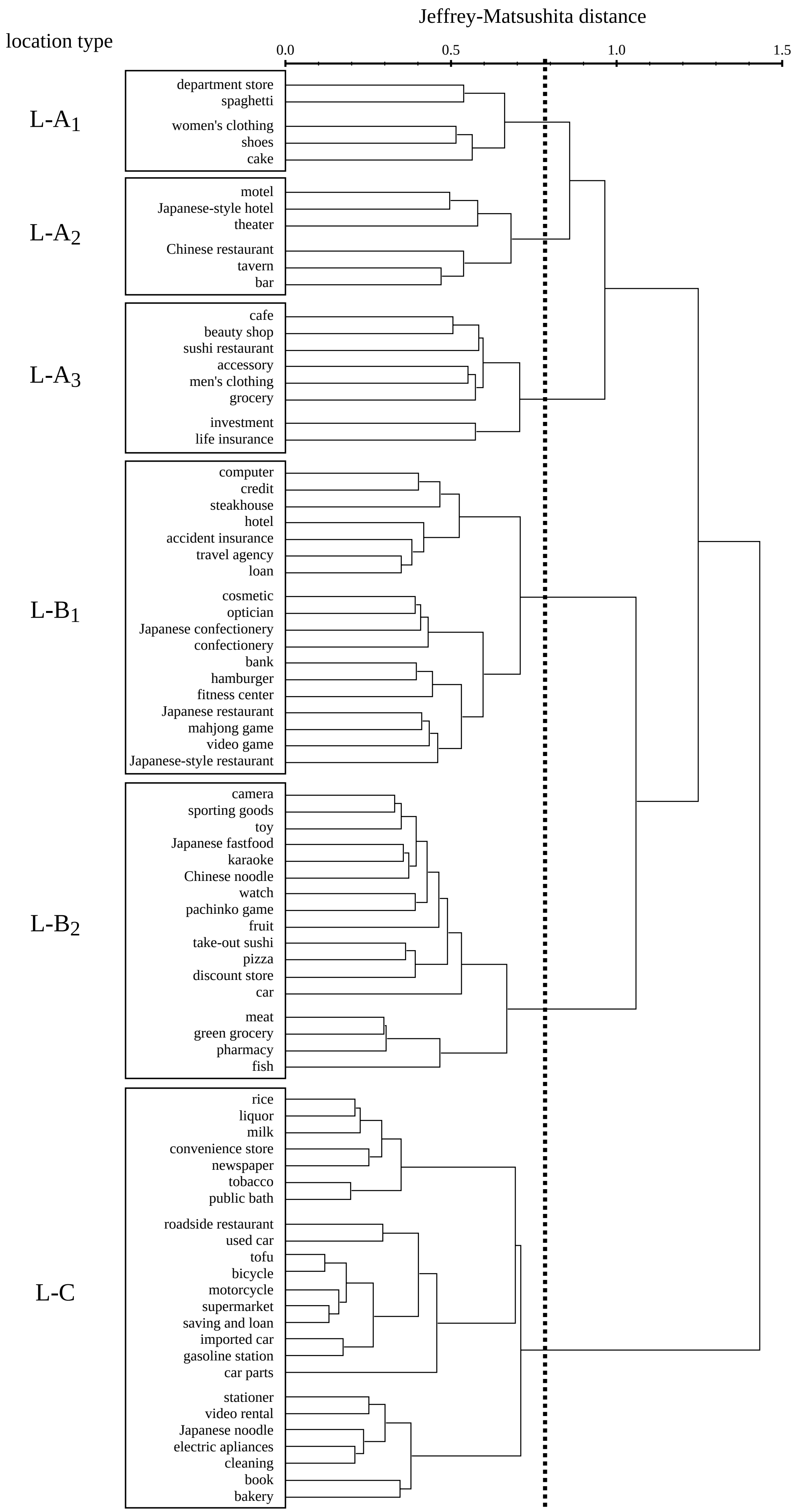
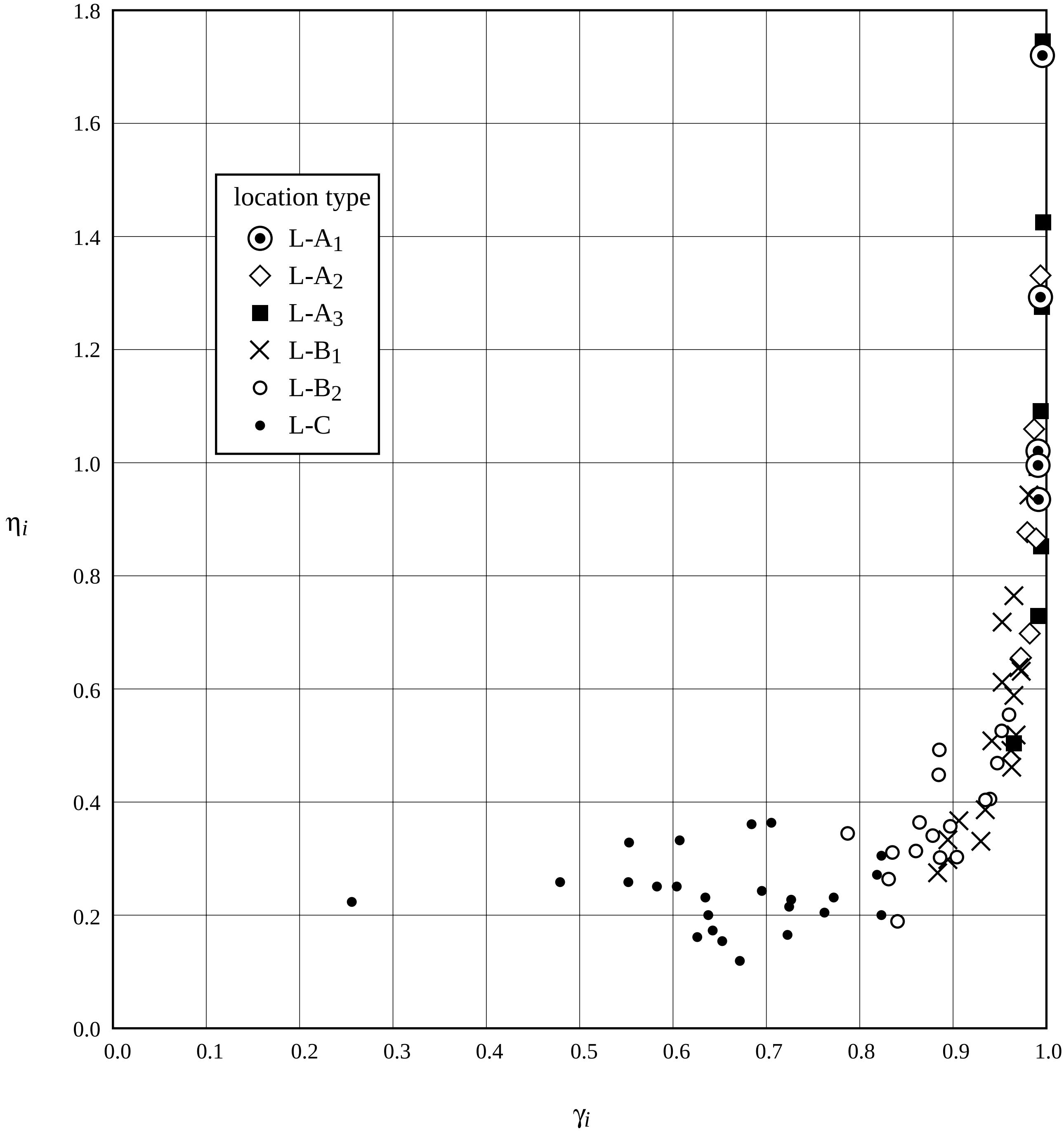

Jeffrey-Matsushita distance
location type
0.0
0.5
1.0
1.5
L-A1
department store
spaghetti
women's clothing
shoes
cake
L-A2
motel
Japanese-style hotel
theater
Chinese restaurant
tavern
bar
L-A3
cafe
beauty shop
sushi restaurant
accessory
men's clothing
grocery
investment
life insurance
L-B1
computer
credit
steakhouse
hotel
accident insurance
travel agency
loan
cosmetic
optician
Japanese confectionery
confectionery
bank
hamburger
fitness center
Japanese restaurant
mahjong game
video game
Japanese-style restaurant
L-B2
camera
sporting goods
toy
Japanese fastfood
karaoke
Chinese noodle
watch
pachinko game
fruit
take-out sushi
pizza
discount store
car
meat
green grocery
pharmacy
fish
L-C
rice
liquor
milk
convenience store
newspaper
tobacco
public bath
roadside restaurant
used car
tofu
bicycle
motorcycle
supermarket
saving and loan
imported car
gasoline station
car parts
stationer
video rental
Japanese noodle
electric apliances
cleaning
book
bakery

Figure 8



location type
L-A1
L-A2
L-A3
L-B1
L-B2
L-C
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
ηi
γi

Figure 9

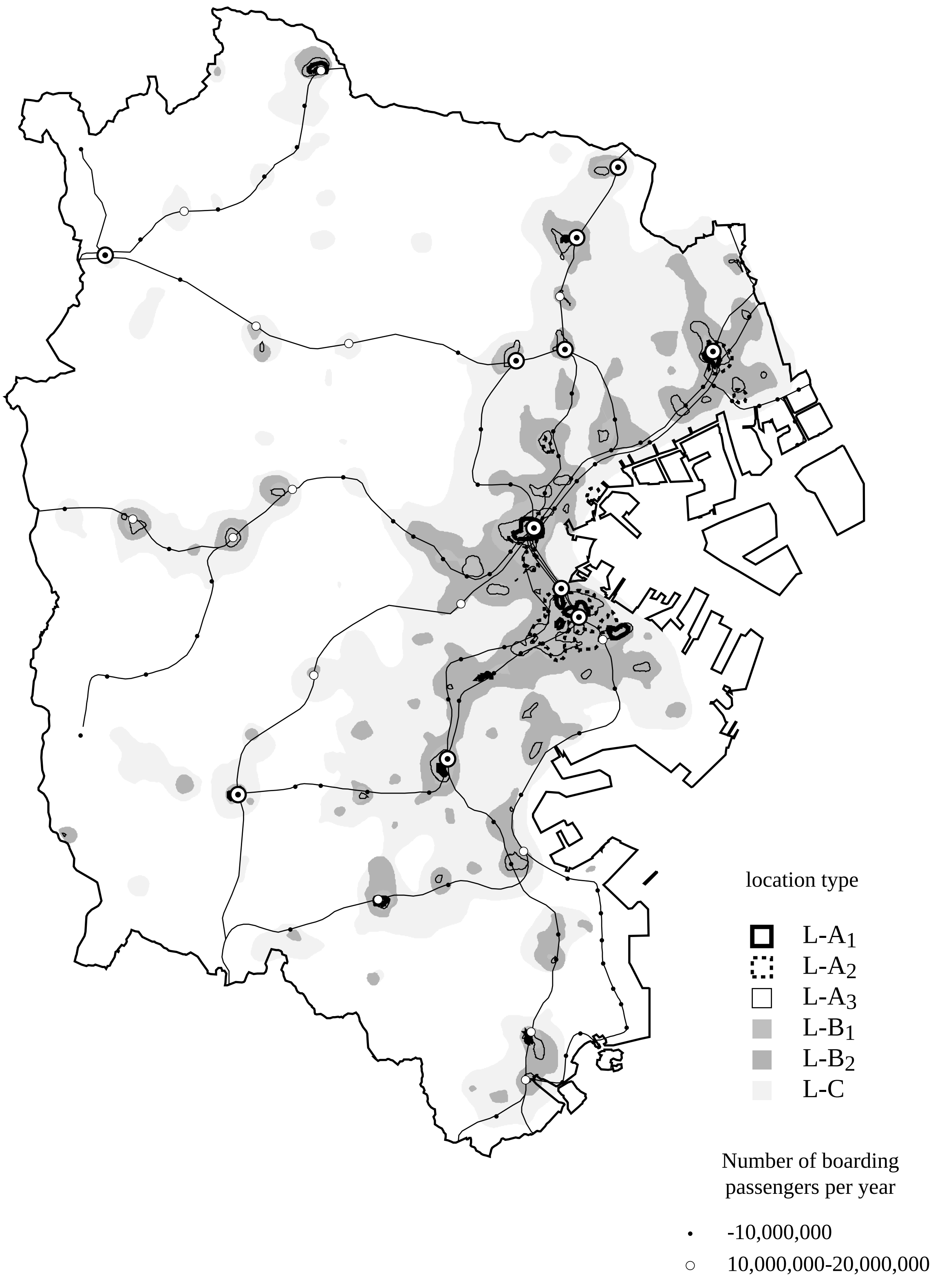

location type
L-A1
L-A2
L-A3
L-B1
L-B2
L-C
Number of boarding passengers per year
-10,000,000
10,000,000-20,000,000
20,000,000-

Figure 10a

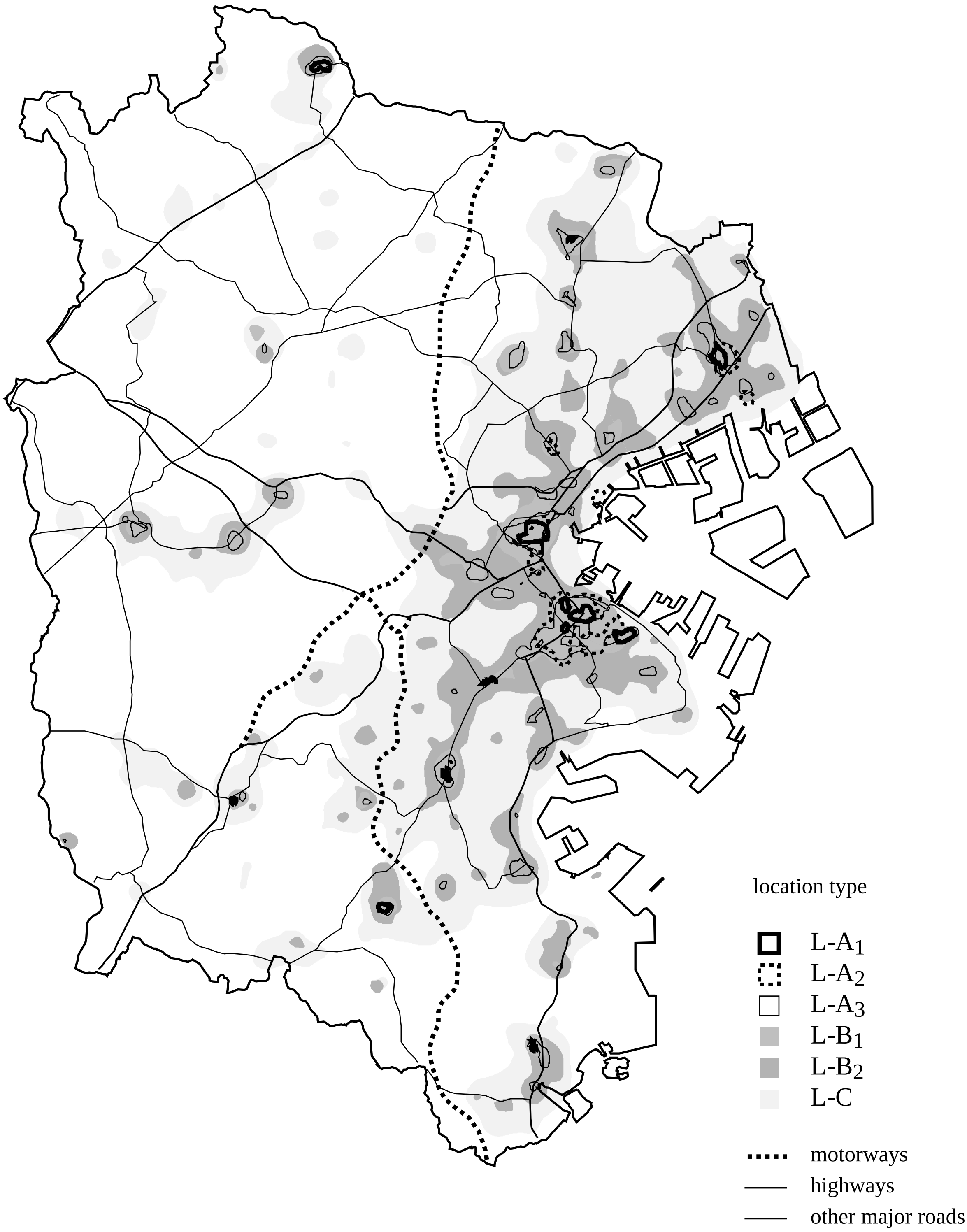
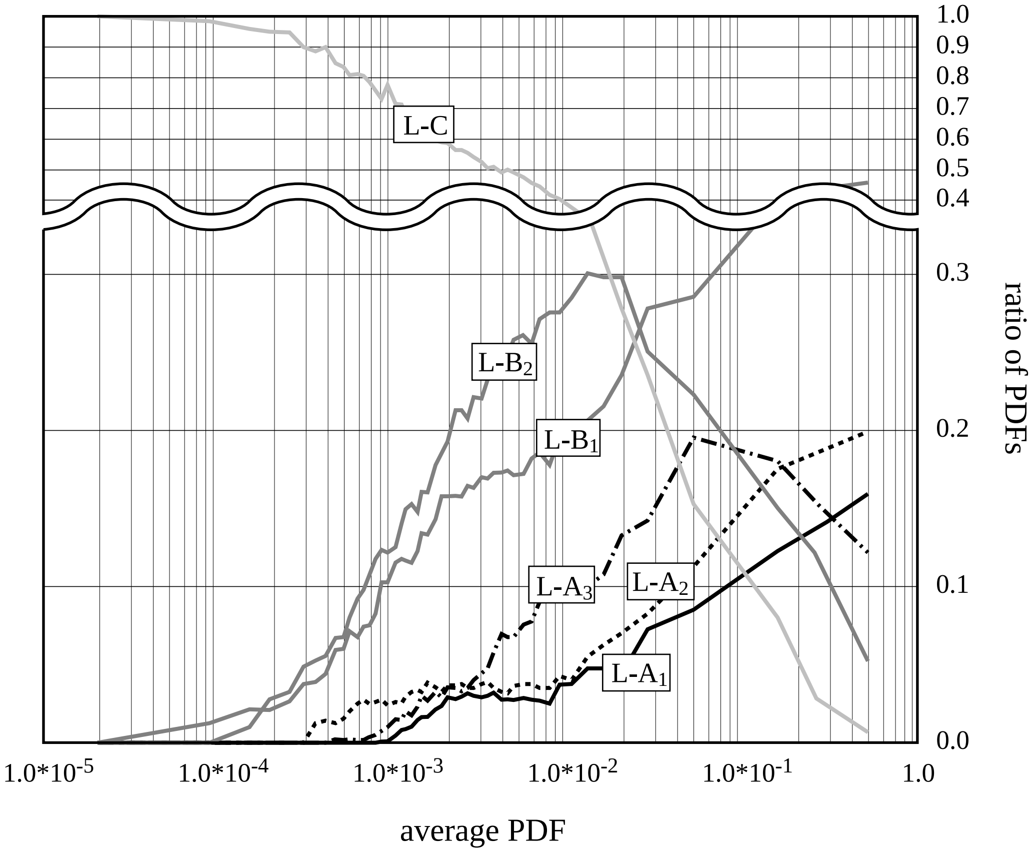

location type
L-A1
L-A2
L-A3
L-B1
L-B2
L-C
motorways
highways
other major roads

Figure 10b



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
ratio of PDFs
L-C
L-B2
L-B1
L-A3
L-A2
L-A1
1.0*10-5
1.0*10-4
1.0*10-3
1.0*10-2
1.0*10-1
1.0
average PDF

Figure 11

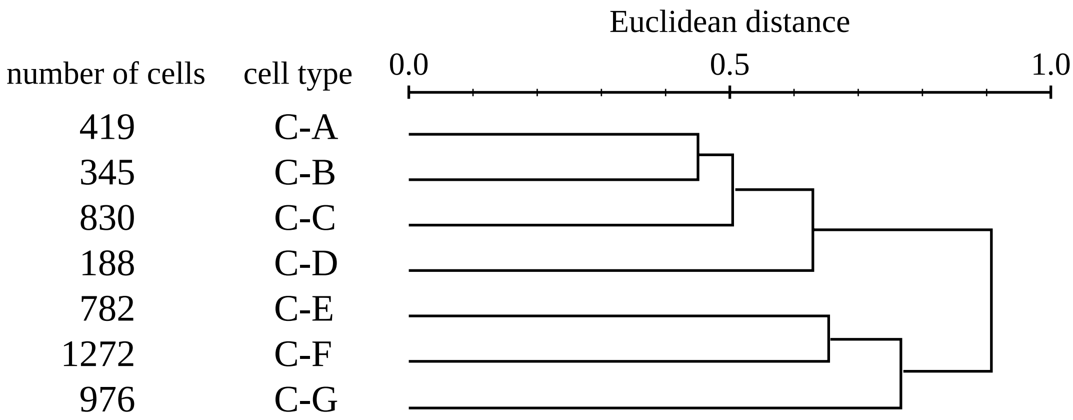
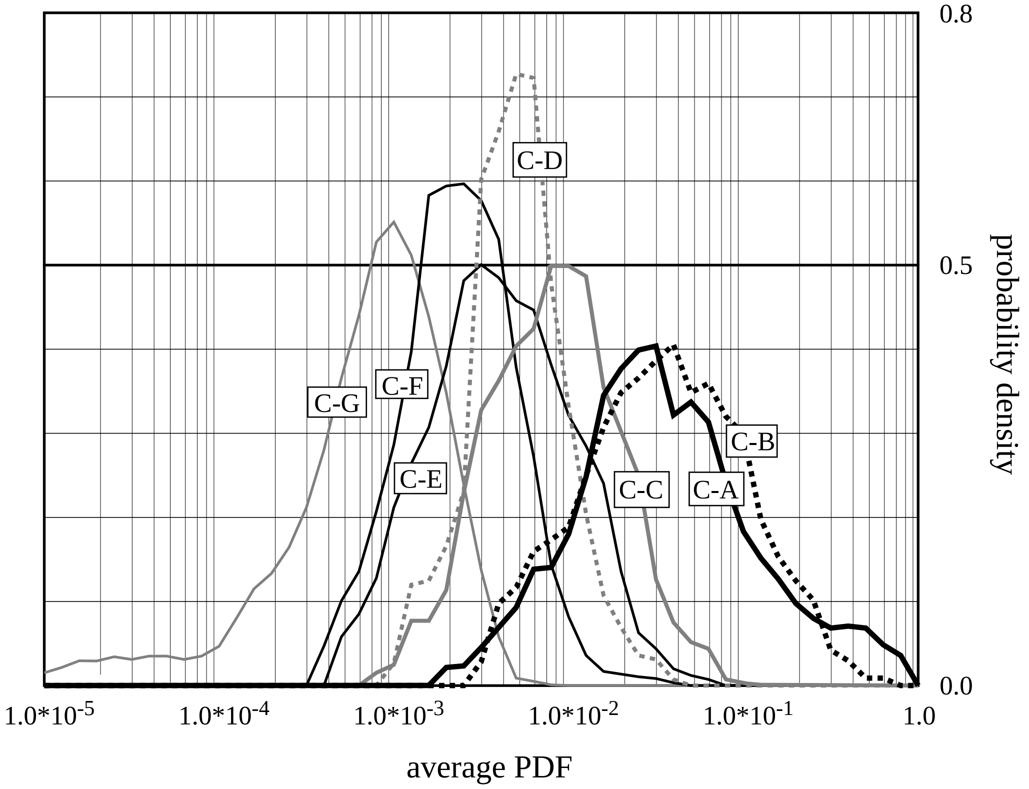

Euclidean distance
number of cells
cell type
0.0
0.5
1.0
419
C-A
345
C-B
830
C-C
188
C-D
782
C-E
1272
C-F
976
C-G

Figure 12



0.8
0.5
0.0
probability density
C-D
C-F
C-G
C-E
C-C
C-A
C-B
1.0*10-5
1.0*10-4
1.0*10-3
1.0*10-2
1.0*10-1
1.0
average PDF

Figure 13

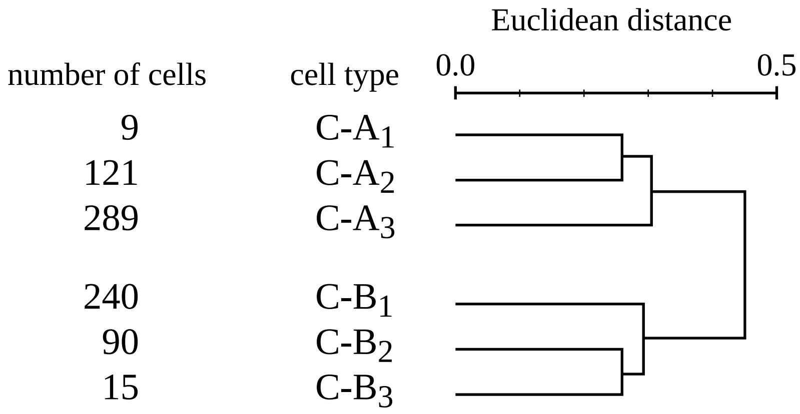

Euclidean distance
number of cells
cell type
0.0
0.5
9
121
289
240
90
15
C-A1
C-A2
C-A3
C-B1
C-B2
C-B3

Figure 14

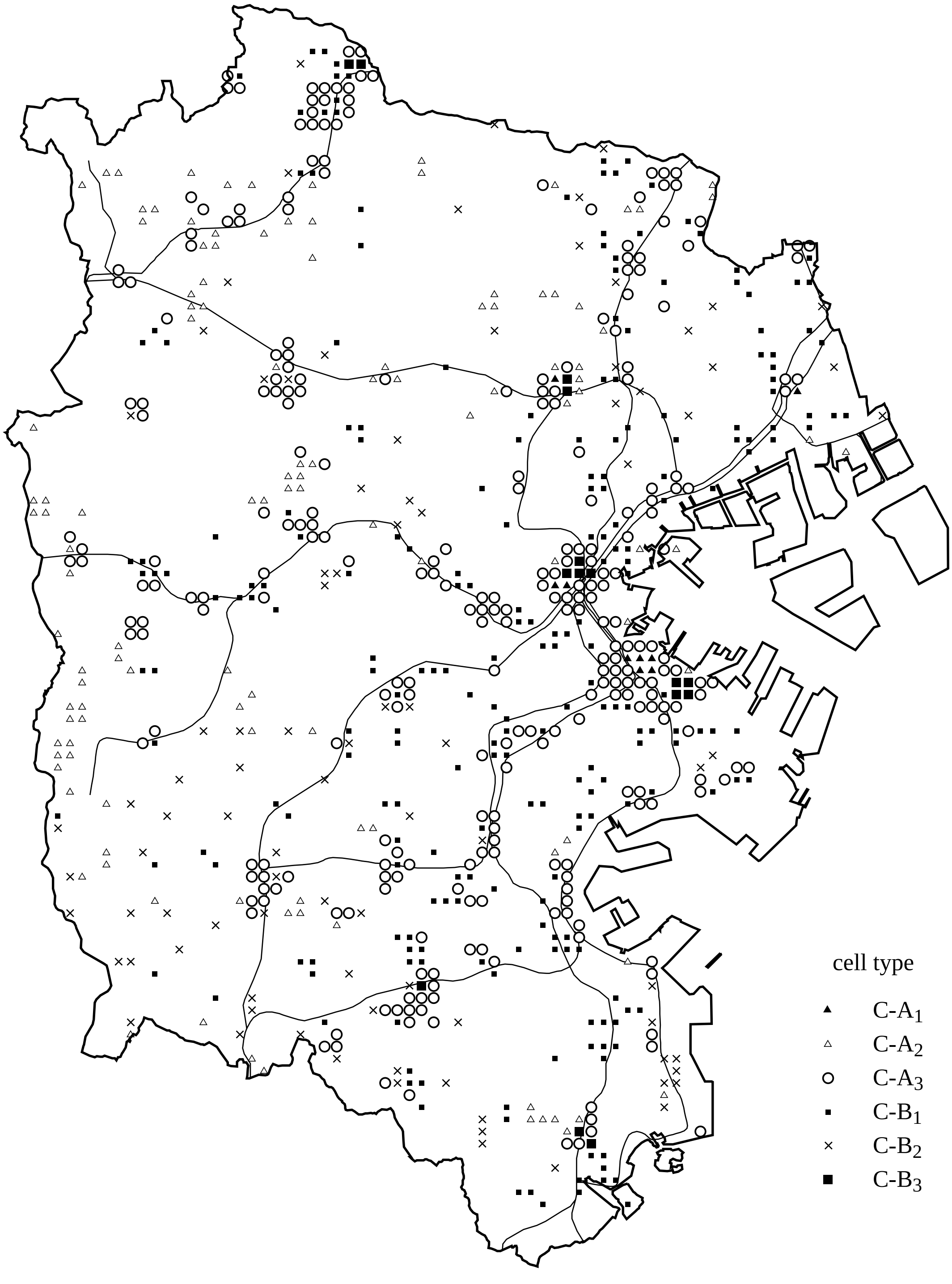


Figure 15

retail type	number of establishments
food stores	
grocery	642
green grocery	662
meat	461
fruit	84
fish	505
rice	458
tofu	209
milk	196
bakery	283
liquor	950
tobacco	216
cake	366
confectionery	546
Japanese confectionery	251
restaurants	
Japanese restaurant	313
Japanese-style restaurant	231
sushi restaurant	863
Chinese restaurant	1354
roadside restaurant	126
Japanese noodle	879
Chinese noodle	214
spaghetti	21
steakhouse	36
Japanese fastfood	161
hamburger	156
pizza	61
take-out sushi	107
cafe	1085
tavern	1663
bar	3705
apparels	
women's clothing	974
men's clothing	209
accessory	309
shoes	339
watch	181
transpot related activities	
car	356
used car	154
imported car	68
motorcycle	179
bicycle	143
auto parts	376
gasoline station	537

retail type	number of establishments
personal services	
cleaning	1886
beauty shop	1809
fitness center	160
public bath	185
travel	
travel agency	186
hotel	47
Japanese-style hotel	91
motel	55
entertainment	
theater	32
video rental	213
video game	144
pachinko game	138
mahjong game	459
karaoke	190
finance	
bank	230
saving and loan	89
loan	358
credit	37
investment	85
life insurance	304
accident insurance	419
multipurpose stores	
department store	50
supermarket	300
discount store	37
convenience store	486
unclassified	
sporting goods	138
electric apliances	1001
camera	171
toy	114
stationer	328
computer	21
book	396
newspaper	254
pharmacy	875
optician	305
cosmetic	642

Table 1

retail type	γ_i	η_i
investment	0.9997	1.7427
life insurance	0.9993	1.4211
spaghetti	0.9992	1.7190
men's clothing	0.9992	1.2743
cafe	0.9984	0.8504
accessory	0.9976	1.0867
theater	0.9975	1.3287
department store	0.9974	1.2911
grocery	0.9964	0.9935
shoes	0.9964	1.0181
beauty shop	0.9963	0.7265
women's clothing	0.9955	0.9328
cake	0.9943	0.9959
tavern	0.9926	0.8624
Japanese-style hotel	0.9911	1.0573
credit	0.9857	0.9389
bar	0.9851	0.6958
motel	0.9828	0.8745
Japanese-style restaurant	0.9797	0.6307
Chinese restaurant	0.9773	0.6532
fitness center	0.9748	0.6330
video game	0.9725	0.5170
sushi restaurant	0.9701	0.5001
bank	0.9698	0.5874
computer	0.9697	0.7625
loan	0.9680	0.4880
mahjong game	0.9671	0.4576
fish	0.9648	0.5501
hotel	0.9583	0.6108
steakhouse	0.9566	0.7157
fruit	0.9545	0.5259
travel agency	0.9501	0.4681
hamburger	0.9465	0.5052
watch	0.9441	0.4048
pachinko game	0.9387	0.4034
Japanese restaurant	0.9380	0.3850
optician	0.9339	0.3317
Japanese confectionery	0.9109	0.3659
camera	0.9087	0.3030
toy	0.9002	0.3588
green grocery	0.8978	0.2939
accident insurance	0.8973	0.3298
confectionery	0.8899	0.2978
meat	0.8892	0.2995
discount store	0.8889	0.4473
car	0.8887	0.4905
cosmetic	0.8866	0.2733
Chinese noodle	0.8818	0.3397
take-out sushi	0.8668	0.3608
karaoke	0.8619	0.3111
pharmacy	0.8432	0.1853
Japanese fastfood	0.8383	0.3082
sporting goods	0.8351	0.2663
public bath	0.8274	0.3048
book	0.8270	0.1974
tobacco	0.8225	0.2683
pizza	0.7890	0.3416
rice	0.7752	0.2281
convenience store	0.7650	0.2019
newspaper	0.7310	0.2273
bakery	0.7270	0.2129
liquor	0.7263	0.1645
roadside restaurant	0.7082	0.3601
milk	0.6991	0.2417
car parts	0.6885	0.3571
Japanese noodle	0.6742	0.1148
electric apliances	0.6553	0.1516
stationer	0.6450	0.1703
vodep rental	0.6397	0.1968
tofu	0.6380	0.2309
cleaning	0.6293	0.1591
used car	0.6088	0.3318
bicycle	0.6087	0.2470
motorcycle	0.5859	0.2495
saving and loan	0.5564	0.2584
imported car	0.5551	0.3272
supermarket	0.4824	0.2554
gasoline station	0.2572	0.2230

Table 2

			location type					
cell type	number of cells	average PDF	L-A_1	L-A_2	L-A_3	L-B_1	L-B_2	L-C
C-A_1	9	0.2234	1.03	12.24	19.24	52.94	7.26	7.29
C-A_2	121	0.0093	2.00	4.83	2.38	48.19	3.28	39.31
C-A_3	289	0.0649	7.03	6.51	5.86	34.42	20.12	26.06
C-B_1	240	0.0433	4.34	4.59	22.39	16.71	29.42	22.55
C-B_2	90	0.0323	1.04	1.97	41.76	14.14	12.92	28.18
C-B_3	15	0.1483	27.54	3.55	22.88	15.43	12.67	17.91

Table 3