Original Paper *Forma*, **16**, 47–63, 2001

Model Analysis for Formation of Population Distribution

Osamu HIRAYAMA[1], Ryuji TAKAKI[2], Takeshi KOSHIZUKA[3] and Hiroshi YANAI[4]

[1]*Tokyo Metropolitan Agricultural High School, Kotobuki-cho 1-11, Fuchu, Tokyo 183-0056, Japan*
[2]*Faculty of Technology, Tokyo University of Agriculture and Technology, Nakamachi 2-24-16, Koganei, Tokyo 184-8588, Japan*
[3]*Department of Social Engineering, Tsukuba University, Tsukuba, Ibaraki 305-8573, Japan*
[4]*Faculty of Science and Technology, Keio University, Yokohama 223-8522, Japan*

(Received September 8, 2000; Accepted February 5, 2001)

Keywords: Population Density Distribution, Railway Network, Traveling Time, Uncomfortableness

Abstract. A model analysis is made for a mechanism of formation of population distribution in case where travel times between two points are assumed to be given in two-dimensional space. Two factors are taken into consideration; one is the tendency that population gathers to the regions from which an average traveling time to other regions is smaller, another is the tendency to avoid the regions with high population density. Based upon this model population distributions were computed numerically in both cases without railroad and with railroad of various shapes.

1. Introduction

As operational research problems there are many analyses to find optimum distributions of service points or optimum traffic networks (OKABE and SUZUKI, 1992). In these analyses population densities are assumed to be given, and mathematical frameworks have been constructed based on this assumption. However, in designing social structures such as service points or traffic networks, the population density is often formed after their constructions and the estimation of amenity of the region can be first estimated after formation of the population density. Therefore, a mathematical tool to determine population density is wanted.

Distributions of the population in real cities and suburbs are not simple, but as a general tendency it is concentrated at regions of the city to which people can access easily. Clark's law (CLARK, 1967) is known as a model to describe such tendency. It assumes that the density ρ decreases exponentially with the distance r from the center.

$$\rho = Ae^{-br}, \quad 0.1 < b < 1.5 \text{ km}^{-1}.$$

This model, however, seems to be based on empirical observations and does not describe the mechanism of the formation of the distribution. In fact, many factors may exist to form the distribution. People will gather at places which can be accessed easily. In general, the shorter the distance is, the easier is to access. But if some traffic facilities like railroad are available, the situation will be different. Topographical, and even climatological conditions must be also taken into account. Moreover, people also have tendency to avoid crowded regions with high population density.

Now, it is necessary to establish a schematised but concrete model to describe the process of formation of the population distribution taking the traffic network into consideration. In this regard, we propose such a model in this paper.

A method of quantitative estimation of a region is proposed by KOSHIZUKA (1997) by introducing a concept of inner distance distribution. This method is suggestive for our purpose.

In the present model it is assumed that a region in the space is evaluated as a place for residence by a function which is the sum of two terms, the former of which corresponds to the uncomfortableness caused by the difficulty of accessibility, while the latter to the uncomfortableness caused by high density of the population. The total uncomfortableness in the space is evaluated by the integration of the uncomfortableness which each person feels.

We assume further that the density of the distribution is so determined that the total uncomfortableness is minimized (or the total amenity is maximized) under the condition that the total population is fixed. Since this problem can be reduced to an equivalent problem in physics of energy minimization, analogous treatment is possible.

2. Model of Formation of Population Distribution

In order to construct the model we set up the following assumptions.

(i) The space is two-dimensional, and the boundary is given. Each point in the space is denoted by the coordinates (x,y).

(ii) Total population N is given, namely,

$$N = \int \rho(x,y)dS, \tag{1}$$

where $\rho(x,y)$ is the population density at the point (x,y) and $dS = dxdy$.

(iii) We introduce here an uncomfortableness function $\psi(x,y)$ at each point in the space which a single person feels. As noted above the total uncomfortableness E of the space is assumed to be expressed by the following integral:

$$E = \iint \psi(x,y)\rho(x,y)dxdy. \tag{2}$$

Uncomfortableness function $\psi(x,y)$ is assumed to be composed of two factors, one is an average access time to other regions in the space, and another a high density of population. It is natural to consider that ψ is an increasing functions of the average access time and the local population density. Since the dependences of ψ on these factors are not clarified enough, it is assumed here that ψ is a linear function of these factors, i.e.

$$\psi(x,y) = U(x,y) + b\rho(x,y), \tag{3}$$

where b is a constant and $U(x,y)$ is a term proportional to average access time to other regions and its concrete expression is given later.

At this stage we can derive a general expression for the density ρ by applying a variational principle to E. By substituting Eq. (3) into Eq. (2), we have

$$E = \int U(x,y)\rho(x,y)dxdy + b\int \rho(x,y)^2 dxdy. \tag{4}$$

The population distribution ρ is determined so that the value of E is minimized under the condition (1). This problem is solved by minimizing $I = E - \lambda N$, where λ is a Lagrange's coefficient. By the use of a technique of variational method we calculate the variation δI, and put it equal to 0, i.e.

$$\delta I = \int U\delta\rho dS + 2b\int \rho\delta\rho dS - \lambda\int \delta\rho dS = 0, \tag{5}$$

which yields to

$$U + 2b\rho - \lambda = 0, \quad \rho = (\lambda - U)/2b. \tag{6}$$

The unknown parameter λ is determined by substituting Eq. (6) into Eq. (1), as

$$\lambda = \frac{1}{S}\left(2bN + \int UdS\right) = 2b\langle\rho\rangle + \langle U\rangle, \tag{7}$$

where

$$S = \int dS \ \text{(total area)}, \quad \langle\rho\rangle = N/S \ \text{(average population density)},$$
$$\langle U\rangle = \int UdS/S \ \text{(average value of convenience)}.$$

Then, substituting Eq. (7) into Eq. (6), we obtain

$$\rho = \langle\rho\rangle + (\langle U\rangle - U)/2b. \tag{8}$$

Now, we define the function $U(x,y)$ by introducing the following assumption:

(iv) The traveling time $T(P,P')$ between two arbitrary points $P(x,y)$ and $P'(x',y')$ in the space is given, and $U(x,y) \equiv U(P)$ is proportional to the average value of traveling time from the corresponding point P to other points P' and the number of occasions to go to a point is proportional to the fraction of population of the point P', i.e. $\rho(P')/N$. Then, we can express U and its average as

$$U(x,y) = \frac{a}{N}\int T(P,P')\rho(P')dS', \quad \langle U\rangle = \frac{a}{NS}\iint T(P,P')\rho(P')dS'dS. \tag{9}$$

Finally, by substituting Eqs. (7) and (9) into Eq. (6), we obtain the following integral equation for $\rho(P)$:

$$\rho(P) = \langle\rho\rangle + \frac{a}{2bN}\left[\frac{1}{S}\iint T(P,P')\rho(P')dS'dS - \int T(P,P')\rho(P')dS'\right]. \tag{10}$$

Note that this equation contains an unknown parameter $a/(2bN)$, which is a ratio of the effect of the traveling convenience to the effect of high density of inhabitants. This parameter is looked upon as an adjustable parameter.

3. Numerical Solutions in Some Cases

3.1. Method of numerical computation

Since it is difficult to obtain analytical solutions of Eq. (10), we try to solve it numerically. We express the two-dimensional space by two-dimensional square grids with uniform intervals, where the area of the grid is denoted by ΔS (see Fig. 1). We replace the

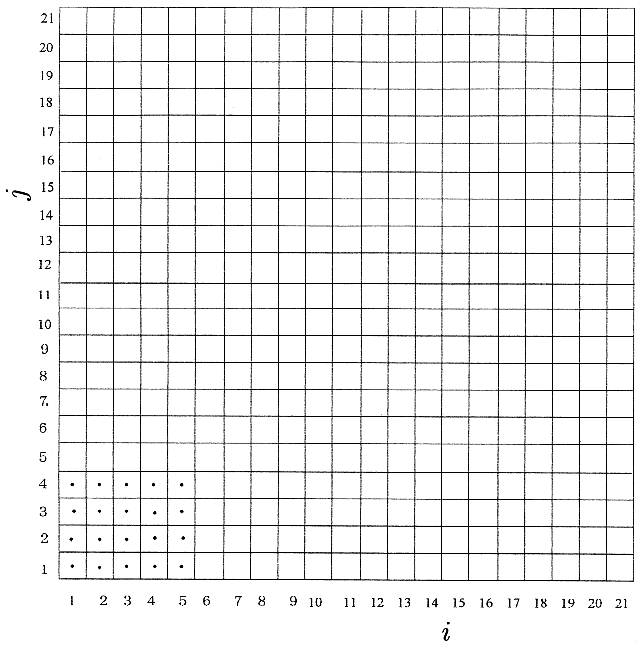

Fig. 1. The grid system applied in this analysis. In cases II and III shown in Table 1 the number of grid points are more.

integral equation (10) by a set of simultaneous algebraic equations for population densities at the grid points ρ_1, ρ_2, ..., ρ_{N_g}, where N_g is the total number of grid points. This manipulation is straight forward and its process is abreviated here. The result is given below.

$$\begin{aligned}
A_{11}\rho_1 + A_{12}\rho_2 + \cdots + A_{1N_g}\rho_{N_g} &= \langle \rho \rangle \\
A_{21}\rho_1 + A_{22}\rho_2 + \cdots + A_{2N_g}\rho_{N_g} &= \langle \rho \rangle \\
&\cdots \\
&\cdots \\
A_{N_g 1}\rho_1 + A_{N_g 2}\rho_2 + \cdots + A_{N_g N_g}\rho_{N_g} &= \langle \rho \rangle,
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
&A_{11} = 1 - q_2 \sum_{L \neq 1} T(1,L), \quad A_{12} = q_1 T(1,2) - q_2 \sum_{L \neq 2} T(2,L), \\
&\qquad \cdots, \quad A_{1N_g} = q_1 T\left(1, N_g\right) - q_2 \sum_{L \neq N_g} T\left(N_g, L\right) \\
&A_{21} = q_1 T(2,1) - q_2 \sum_{L \neq 1} T(1,L), \quad A_{22} = 1 - q_2 \sum_{L \neq 2} T(2,L), \\
&\qquad \cdots, \quad A_{2N_g} = q_1 T\left(2, N_g\right) - q_2 \sum_{L \neq N_g} T\left(N_g, L\right) \\
&\cdots \\
&A_{N_g 1} = q_1 T\left(N_g, 1\right) - q_2 \sum_{L \neq 1} T(1,L), \quad A_{N_g 2} = q_1 T\left(N_g, 2\right) - q_2 \sum_{L \neq 2} T(2,L), \\
&\qquad \cdots, \quad A_{N_g N_g} = 1 - q_2 \sum_{L \neq N_g} T\left(N_g, L\right),
\end{aligned} \tag{12}$$

$$q_1 = c\Delta S = \frac{cS}{N_g}, \quad q_2 = \frac{c(\Delta S)^2}{S} = \frac{c\Delta S}{N_g}, \quad c = \frac{a}{2bN}. \tag{13}$$

Here, we must give an expression of traveling time between two points $P(x,y)$, $P'(x',y')$,

$$T(P,P') \equiv T(x,y;x',y') = (|x - x'| + |y - y'|)/v_0, \tag{14}$$

where v_0 is the walking velocity. If a railroad is constructed in the region, the traveling time between two points $P(x,y)$ and $P'(x',y')$ can be expressed as follows, by partly using the railroad:

$$T(P;P') = \min(T(P,A) + L(A,B)/v + T(B,P')), \tag{15}$$

where A and B are the points of stations, $L(A,B)$ is a distance between A and B, and v is the train velocity. The symbol "min" means that we choose the minimum value for all possible combinations of (A,B).

Procedure of numerical solution is as follows. First, we give the railroad planning (location of stations and distances between any two stations). Second, we compute the traveling time $T(P,P')$ between two arbitrary points, by the use of Eq. (15). Third, we solve the simultaneous algebraic equations (11) by the Gauss-Jordan method.

Values of parameters c (psychological parameter), the total area S, N_g ($=S/\Delta S$), v_0, v must be specified before computation. We must fix appropriate values of parameters so that resulting values of $\rho(P)$ might not be negative. We treated 14 cases as listed in Table 1. We fix $N = 441$ for all cases except II and III, and grid distance is fixed at 1 for all cases hence $\Delta S = 1$. Shapes of the region is assumed to be a square with side length 21, hence $N_g = 441$, except for rectangular cases II (42×21, $N_g = 882$) and III (63×21, $N_g = 1323$). The mean population density is fixed unity for all cases. We fixed $a = 1$, $b = 5$, hence $c = 0.0002267$ in the all cases except II and III. The walking and train velocities v_0 and v are fixed as $v_0 = 1$ and $v = 10$ in all cases.

3.2. *Results of numerical solutions*

Figures 2(a) and (b) show the two-dimensional and three-dimensional contour map of population distribution in case I respectively. The contours are nearly concentric circles, where the center of the region has the highest density, as is expected. It is remarkable that the contours near the periphery are circular in spite of the square boundary.

Table 1. Cases treated in numerical solution.

Case	Number of stations	N	c	E	(E/N)	Figure
I	none	441	0.0002267	7994.77	18.1287	Figs. 2(a) and (b)
II	none	882	0.0000708	24368.82	27.6290	Fig. 3(a)
III	none	1323	0.0000472	43636.07	32.9827	Fig. 3(b)
IV	2	441	0.0002267	7566.15	17.1568	Fig. 4(a)
V	3	441	0.0002267	7133.97	16.1768	Fig. 4(b)
VI	4	441	0.0002267	6881.71	15.6048	Fig. 4(c)
VII	5	441	0.0002267	6800.43	15.4204	Fig. 4(d)
VIII	5	441	0.0002267	6569.43	14.8967	Fig. 5(a)
IX	5	441	0.0002267	6507.92	14.7572	Fig. 5(b)
X	5	441	0.0002267	6531.74	14.8112	Fig. 5(c)
XI	5	441	0.0002267	6852.35	15.5382	Fig. 5(d)
XII	5	441	0.0002267	6544.50	14.8401	Fig. 5(e)
XIII	5	441	0.0002267	6586.68	14.9358	Fig. 5(f)
XIV	4	441	0.0002267	6900.93	15.6483	Fig. 5(g)

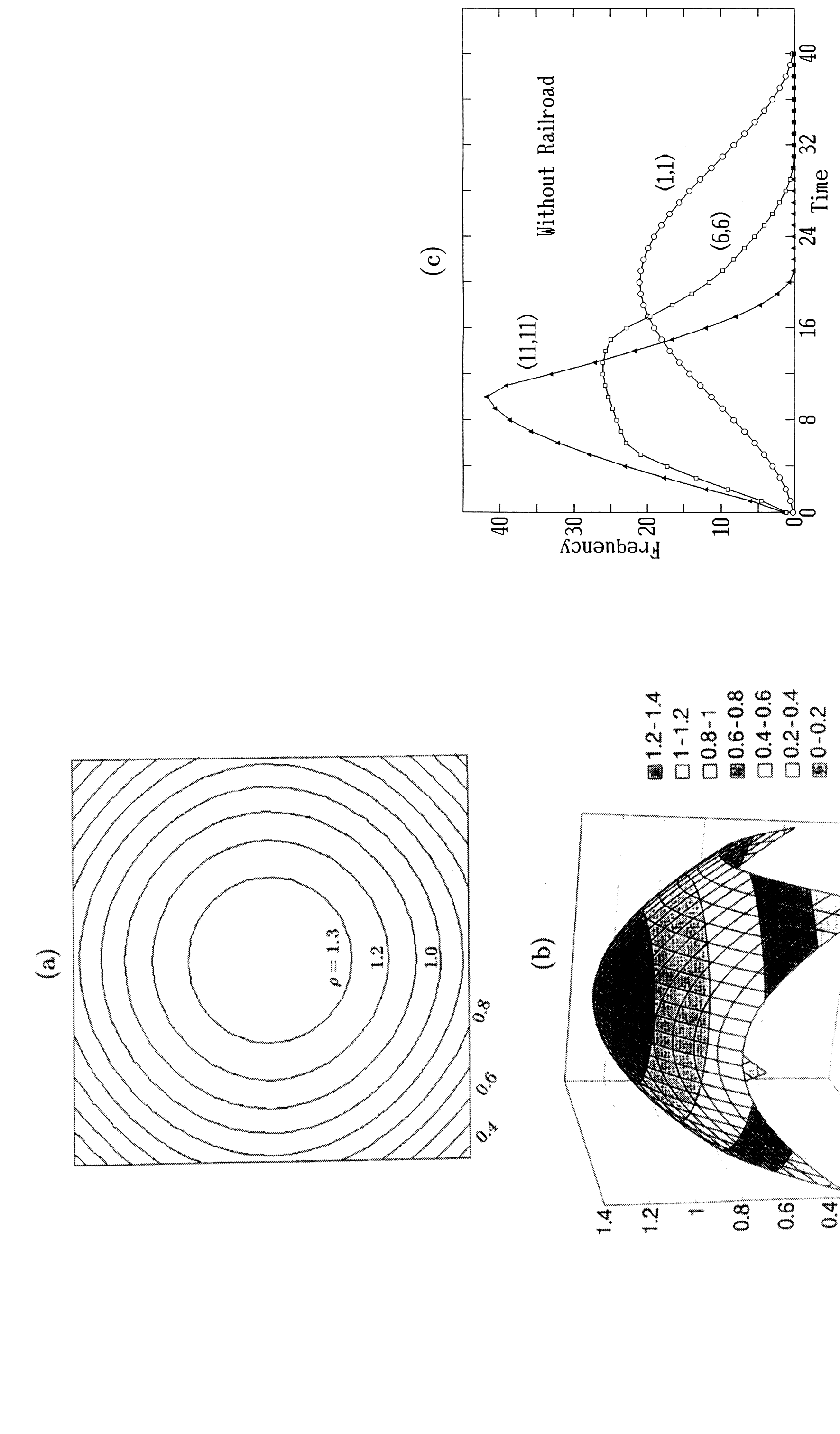

Fig. 2. Results in case I. (a) Two-dimensional contour map of population density. (b) Three-dimensional contour map of population density. (c) Histograms of traveling times from the three representative points in the space. The coordinates are shown in the figure.

Figure 2(c) shows the distribution of traveling time from representative three points (indicated in the figure) in this case. This figure shows that the central point (11,11) is most convenient as to the traveling time.

Figures 3(a) and (b) show the contour maps of population density in cases II and III, which have shapes of concentric ellipses. The results shown in Figs. 2 and 3 seem to suggest that the shapes of contours depend much on the global nature of the regions, whether they are elongated or not.

In the rightmost column in the Table 1 the value of (E/N) is shown, because for the comparison of uncomfortableness of regions with different populations the value of uncomfortableness per person is more important than the value of E itself. By the comparison of the results of cases I, II and III we can see (E/N) has the lowest value in case I with square region. This result is convincing since the traveling time is important in the present analysis.

Figures 4(a)–(d) show the results in cases with straight railroad in cases IV–VII. As is expected the population densities adjacent to the railroad are high. The traveling time distribution in case VII is shown in Fig. 4(e). By comparing Fig. 2(c) and 4(e) we can see that the curves have the peaks at the smaller traveling time in case with railroad than in case

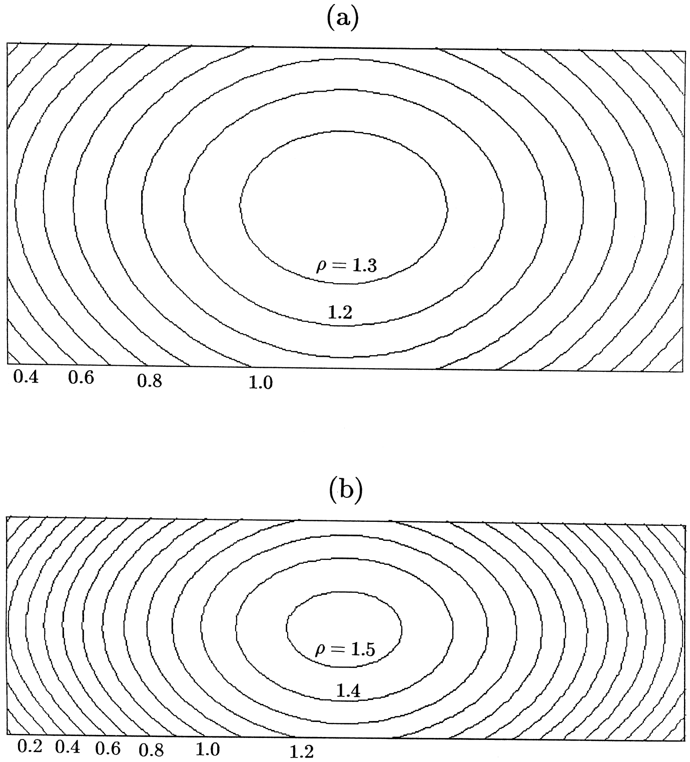

Fig. 3. Two-dimensional contour maps of population density in the rectangular cases. (a) Case II, (b) Case III.

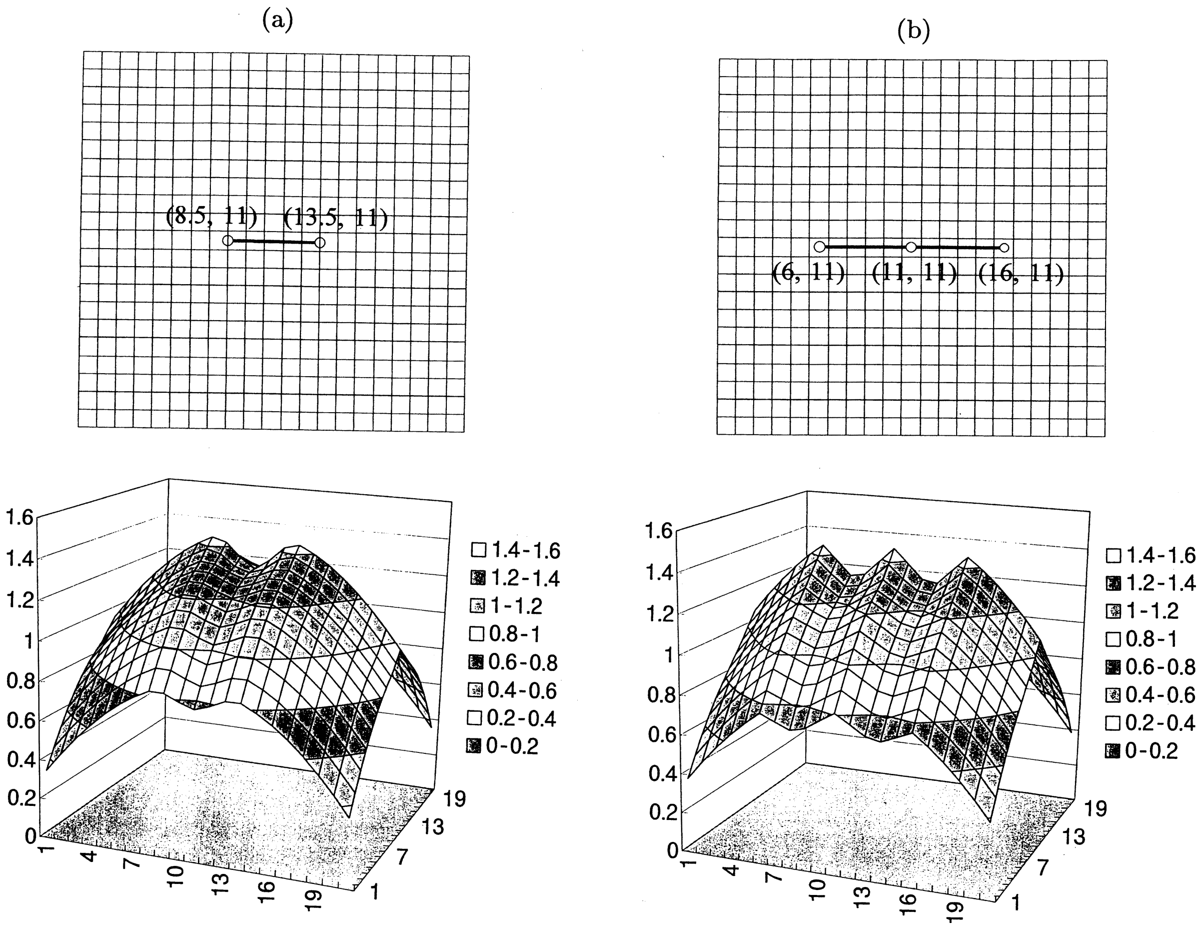

Fig. 4. Results in case IV–VII with straight railroad. (a) In the case with two stations. Above: railroad plannings, below: contour maps of population density. (b) In the case with three stations. (c) In the case with four stations. (d) In the case with five stations. (e) Traveling time distribution from the three representative points in case VII. (f) Railroad length dependence of total uncomfortableness E.

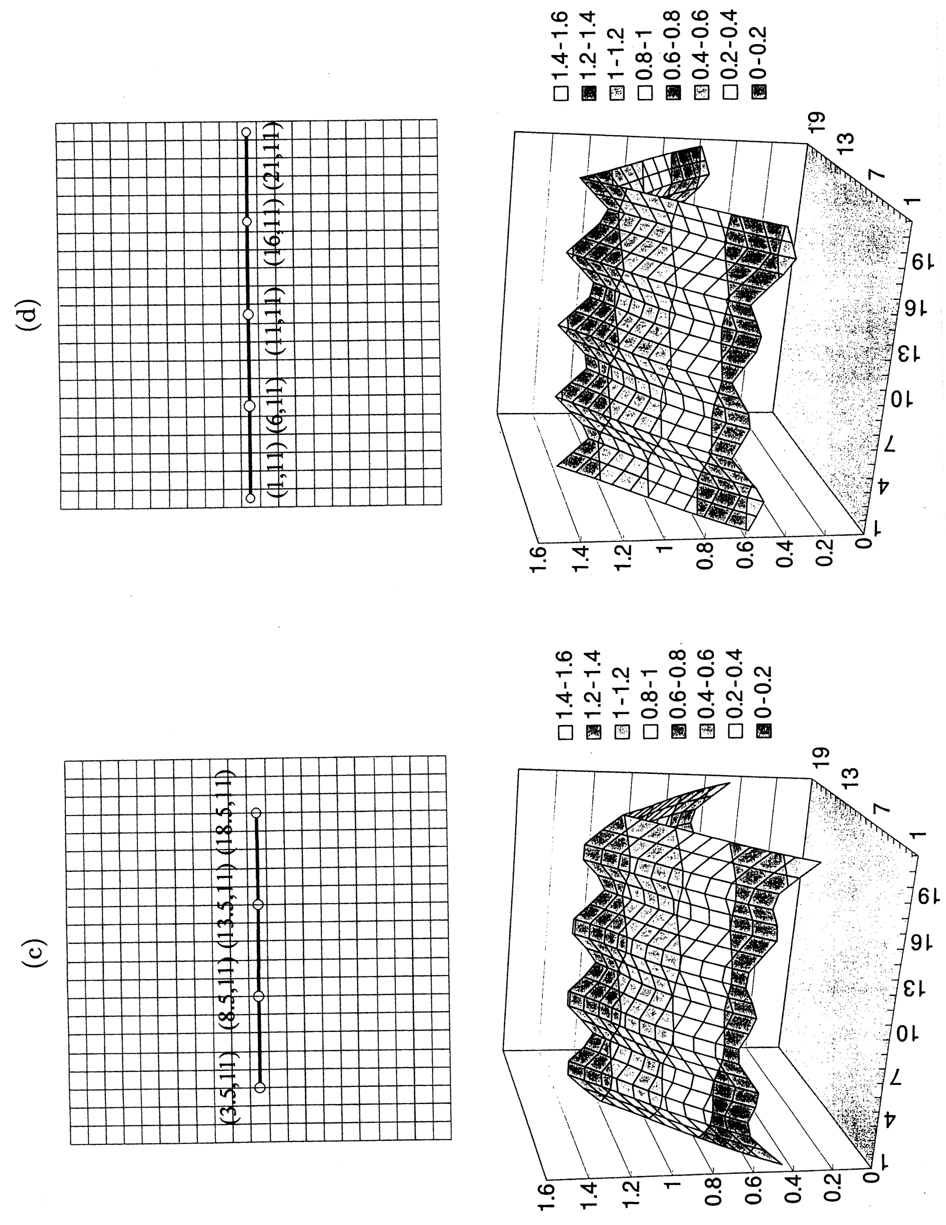

Fig. 4. (continued).

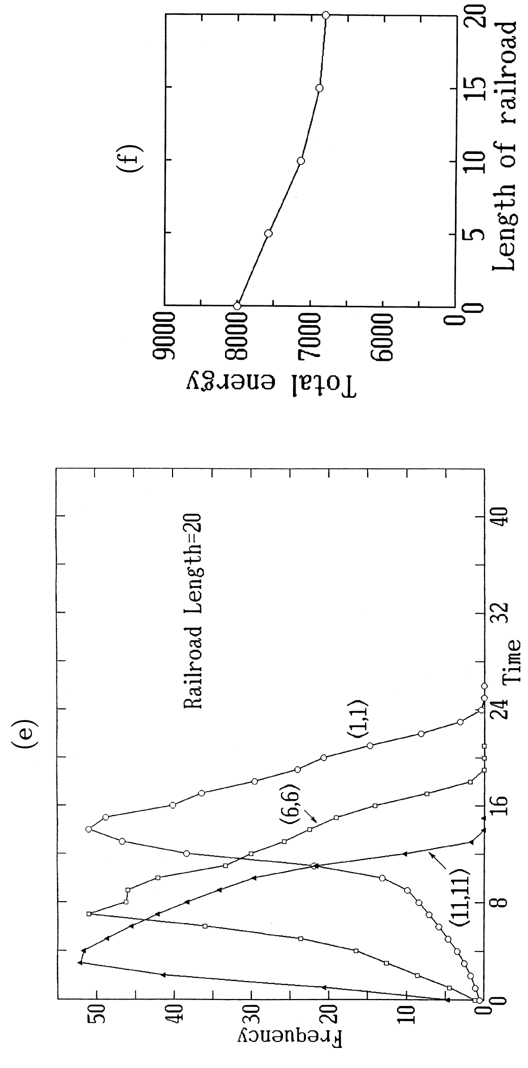

Fig. 4. (continued).

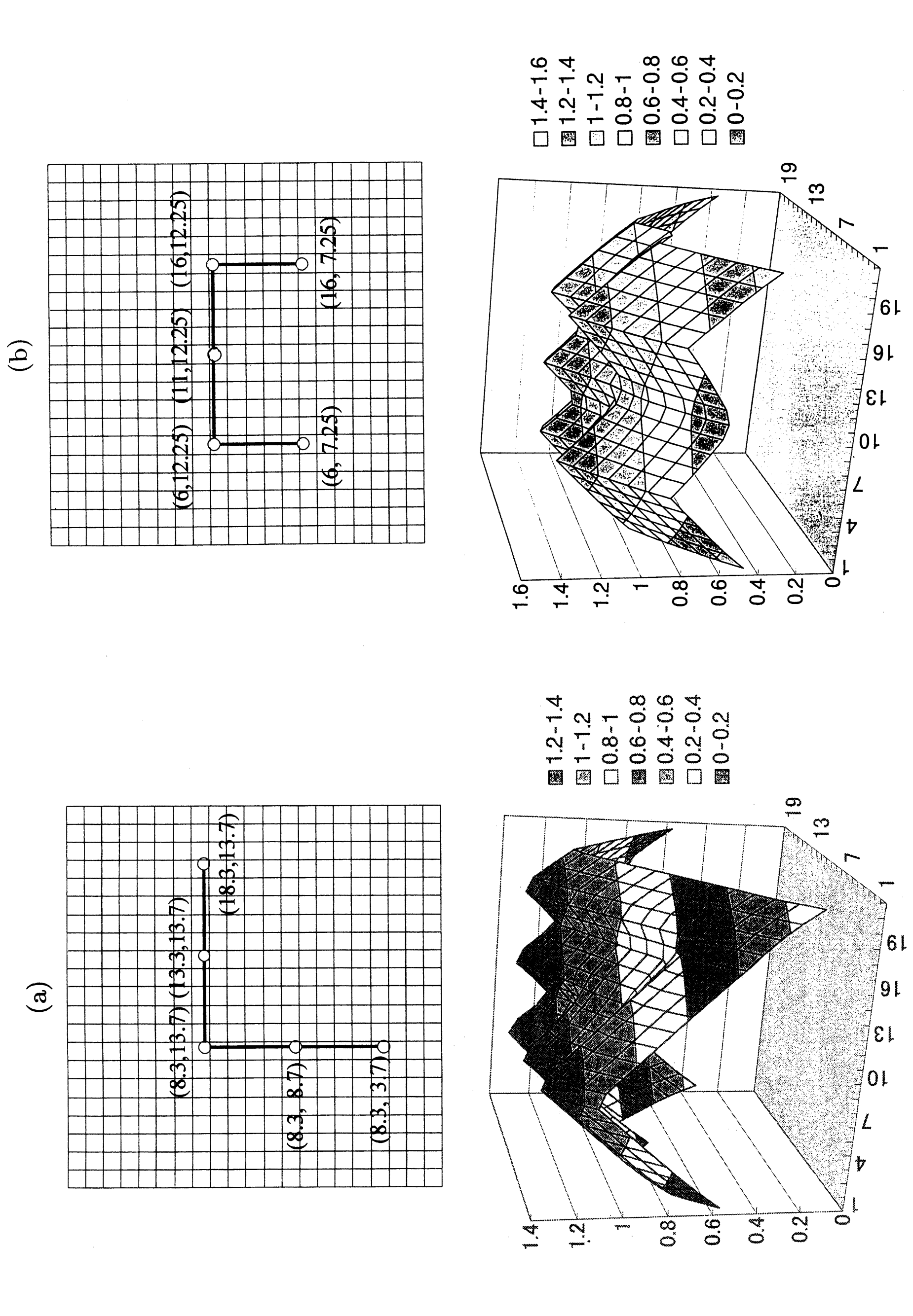

Fig. 5. Results in cases VIII–XIV. The railroad length is fixed 20 in all cases. Above: railroad plannings, below: contour map of population density.

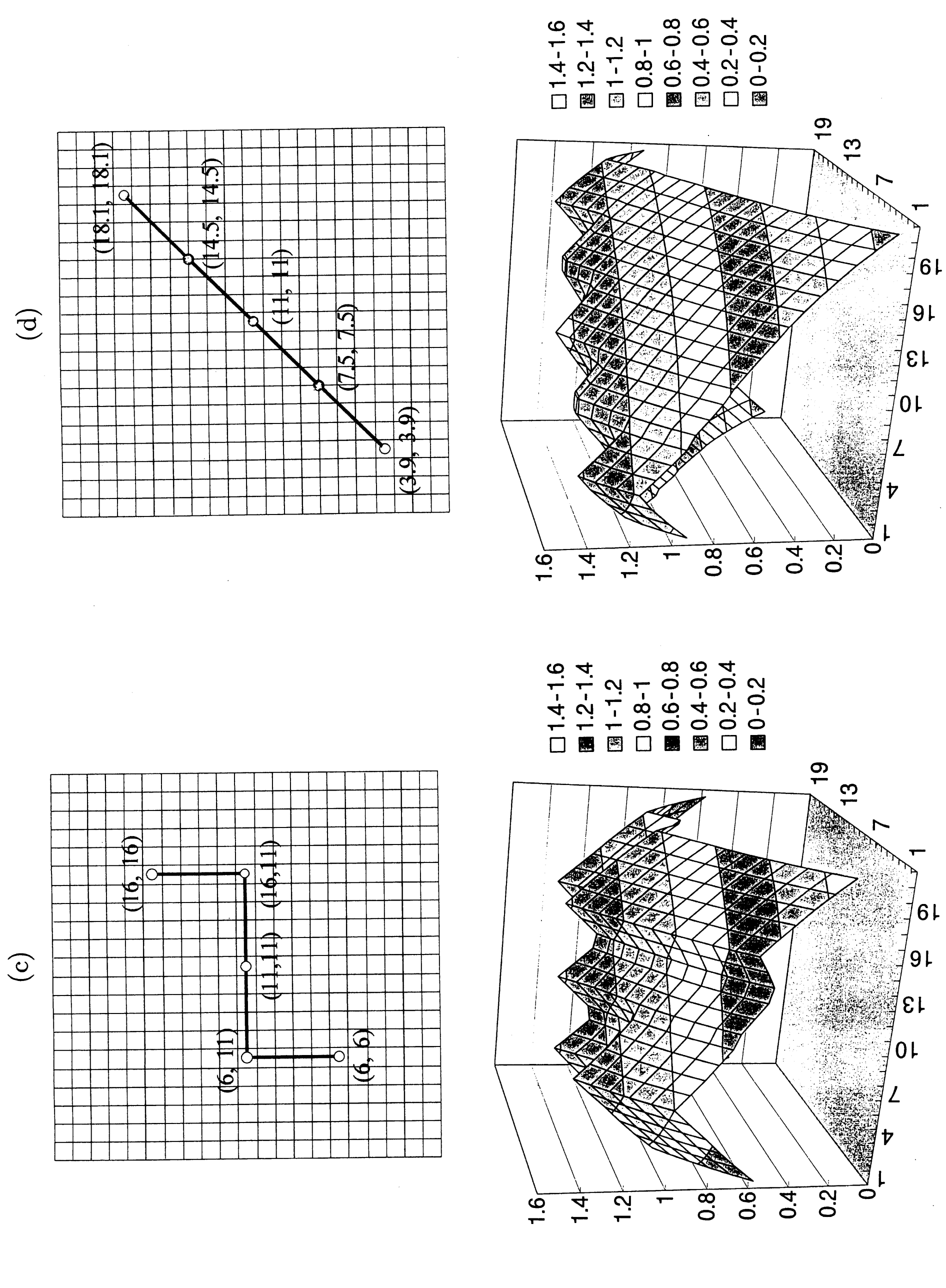

Fig. 5. (continued).

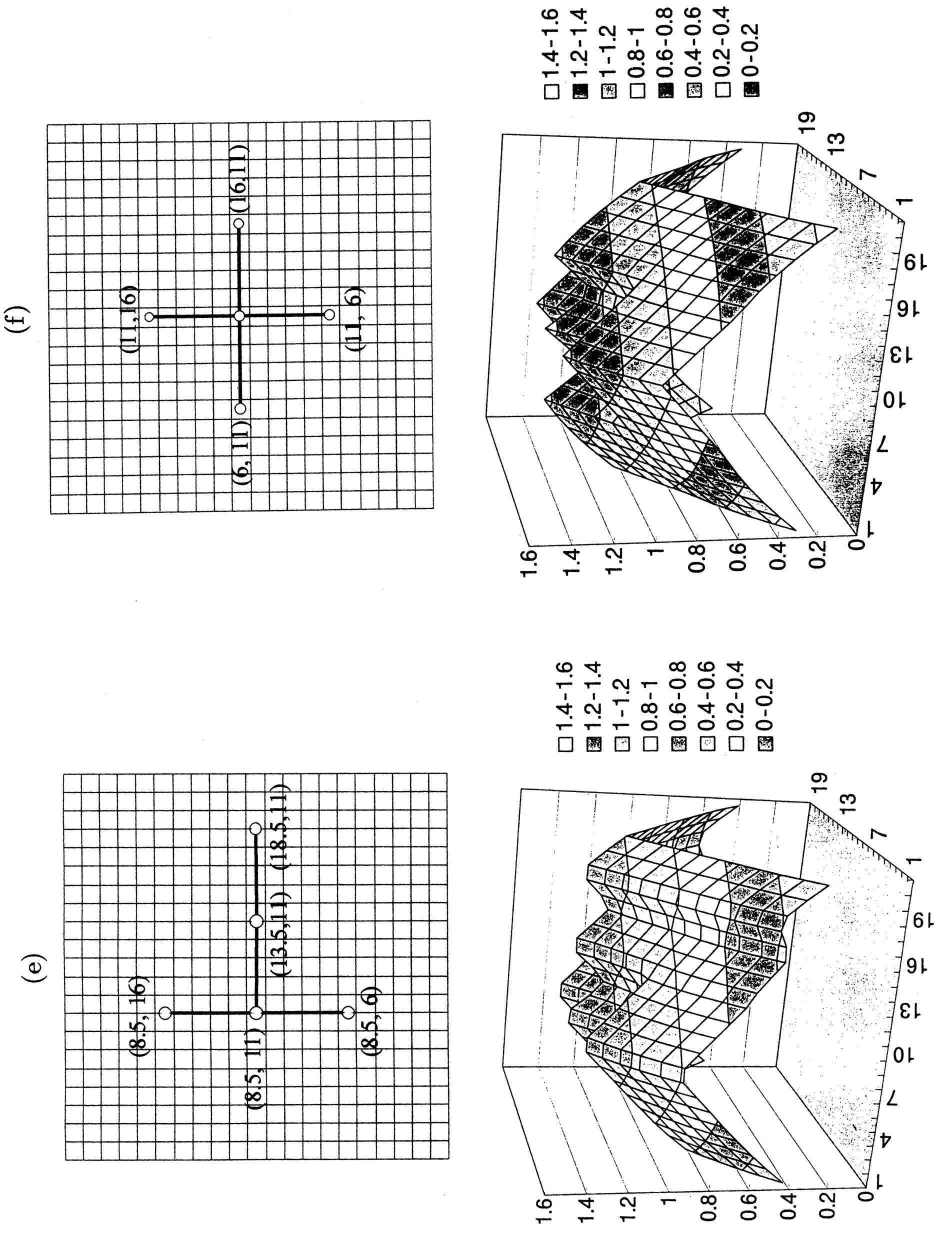

Fig. 5. (continued).

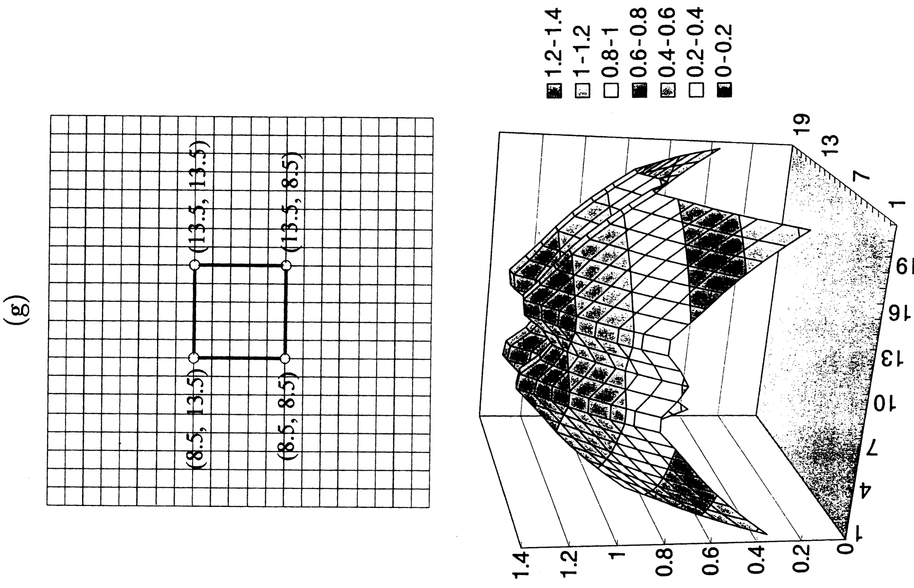

Fig. 5. (continued).

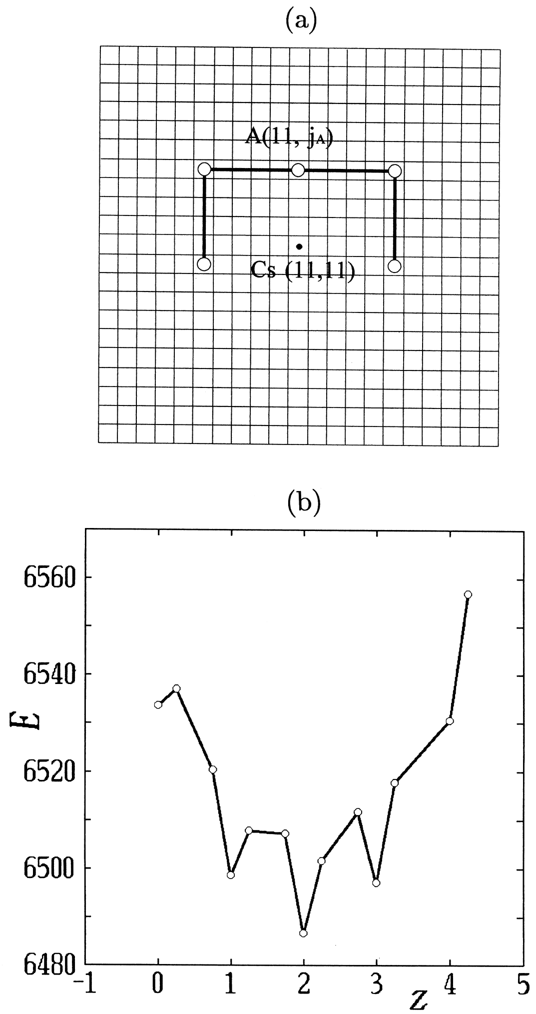

Fig. 6. Dependence of E to the vertical position of the center of the railroad in case IX. (a) Position of the center of the railroad. (b) Dependence of E to $z \equiv j_A - 11$.

without railroad. The railroad length dependence of the total energy E in cases I and IV–VII are shown in Fig. 4(f). We can see E decreases (amenity increases) monotonically with the increase of the railroad length, as is expected. Note that degree of the decrease of E is not so remarkable when the railroad length is larger than 15, which suggests that the extension of the railroad near the boundary of the space is not so effective for the increase of the convenience of the whole space.

Figures 5(a)–(g) show the results of case VIII–XIV, in which length of the railroad is 20 and the center of the mass of the railroad is fixed at the center of the space, i.e. (11,11). The population density are high in the regions along the railroad also in these cases, which is similar tendencies to that of Fig. 4 (cases IV–VII), however they loose spatial symmetry. Among these seven cases the total uncomfortableness E takes the least value in case IX. This fact suggests that case IX is the best choice for railroad shape with length 20.

The total uncomfortableness E varies also if the position of the railroad is shifted with its shape unchanged. Therefore, we examined the effect of the vertical position of the center of the railroad. The position is defined by the value of j_A for the station A (see Fig. 6(a)). We can see from Fig. 6(b) that E depends much on the distance between the station A and the center of the whole space C_s, $z \equiv j_A - 11$, and that E takes the lowest value at $z = 2$. Note that E takes the minimum values when the stations are located at the grid points.

4. Conclusions and Discussions

In this paper we have shown that the proposed mathematical method with variational principle is effective in estimating total uncomfortableness of the space on the analogy of the physical concept of energy. The method would also give a guideline in transportation network. Details of conclusions derived from the above results are listed below.

(1) In the case without railroad the population density contours are nearly concentric circles.

(2) The railroad attracts the population density towards itself and reduces the uncomfortableness of the region.

(3) The total uncomfortableness decreases with the length of the railroad.

(4) If the total length of railroad is fixed, the best planning of the railroad is determined by comparing the value of total uncomfortableness.

Some comments on the above conclusions must be added here. First, the L1 distance D between the two points is employed in the present analysis. We have tried to estimate the population densities by the use of L2 distance (Euclidean distance) in some cases, but patterns of population distribution are almost the same as the present results.

Second, we fixed $a = 1$, $b = 5$, hence $c = 0.0002267$ in the all cases except II and III, and $a = 1$, $b = 8$, hence $c = 0.0000708$, 0.0000472 in cases II and III. These values are determined by trial and error so that the values of population density of all the points are positive. If we employ some smaller value of c than that of present analysis, population densities in some points will become negative, which should be prohibited in this analysis.

Third, in the present analysis we have assumed a relatively simple situation, i.e. the space is uniform and the convenience of a point is determined only by the traveling time, and the space has one railroad line. Algorithm of numerical computation, allowing situations beyond these constraints will be designed in the future.

Fourth, in the present analysis the railroad planning is given at first and population density is given after that. However in the actual social development railroads are planned with the increase of the population. The present analysis could be extended to that more realistic situation, which is a subject of future work.

REFERENCES

CLARK, C. (1967) *Population Growth and Land Use*, Macmillan.

KOSHIZUKA, T. (1997) Analysis of road networks based on transportation, *Bull. of the Soc. Sci. Form,* **12**, 22–23 (in Japanese).

OKABE, A. and SUZUKI, A. (1992) *Mathematics of Optimum Configuration*, Asakura Syoten (in Japanese).