



Asian J. of Management Science and Applications

ISSN online: 2049-8691 - ISSN print: 2049-8683

<https://www.inderscience.com/ajmsa>

Estimation of average TSP distance considering zone length-width ratio: implications for territory design

Daisuke Hasegawa, Naoshi Shiono

DOI: [10.1504/AJMSA.2025.10073006](https://doi.org/10.1504/AJMSA.2025.10073006)

Article History:

Received:	01 May 2024
Last revised:	15 July 2024
Accepted:	20 July 2024
Published online:	01 October 2025

Estimation of average TSP distance considering zone length-width ratio: implications for territory design

Daisuke Hasegawa*

Center for Real Estate Innovation,
The University of Tokyo,
7-3-1, Hongo, Bunkyo-ku,
Tokyo, 113-0033, Japan
Email: hasega60@g.ecc.u-tokyo.ac.jp
*Corresponding author

Naoshi Shiono

School of Business Administration,
Senshu University,
2-1-1, Higashimita, Tama, Kawasaki,
Kanagawa, 214-8580, Japan
Email: na-shiono@isc.senshu-u.ac.jp

Abstract: In vehicle routing, estimating route lengths using continuous approximation models can be valuable for delivery planning, particularly for tour cost estimation and territory design, because it eliminates the computational costs associated with solving TSP and VRP directly. In this study, we propose a route-length estimation formula based on rectilinear distances by considering the shape of the area. We calibrate the parameters through numerical experiments and apply them to a territorial decision problem. Thus, our proposed formula can estimate the average tour length over a rectilinear distance with high accuracy. As the number of points increases, the influence of the shape decreases.

Keywords: travelling salesman problem; continuous approximation models; territory design; rectilinear metric.

Reference to this paper should be made as follows: Hasegawa, D. and Shiono, N. (2025) 'Estimation of average TSP distance considering zone length-width ratio: implications for territory design', *Asian J. Management Science and Applications*, Vol. 8, No. 2, pp.218–232.

Biographical notes: Daisuke Hasegawa is a Project Lecturer at the Center for Real Estate Innovation, the University of Tokyo. He received his PhD in Engineering from the University of Tsukuba in March 2019. He worked at DeNA Co., Ltd. from April 2019 and at the Institute of Industrial Science, the University of Tokyo, from July 2020. He joined his current position at the University of Tokyo in April 2022.

Naoshi Shiono is a Professor at Senshu University. He earned his PhD in Business Administration from the University of Tsukuba in November 2016. After working at Tokyo Gas Co., Ltd. and the Kanagawa Institute of Technology, he joined Senshu University in April 2025. His research interests include optimisation modelling and data-driven infrastructure planning.

This paper is a revised and expanded version of a paper entitled 'Average TSP tour length approximations for territory design' presented at The 7th Asian Conference of Management Science and Applications (ACMSA2023), Okinawa, Japan, 15 December 2023.

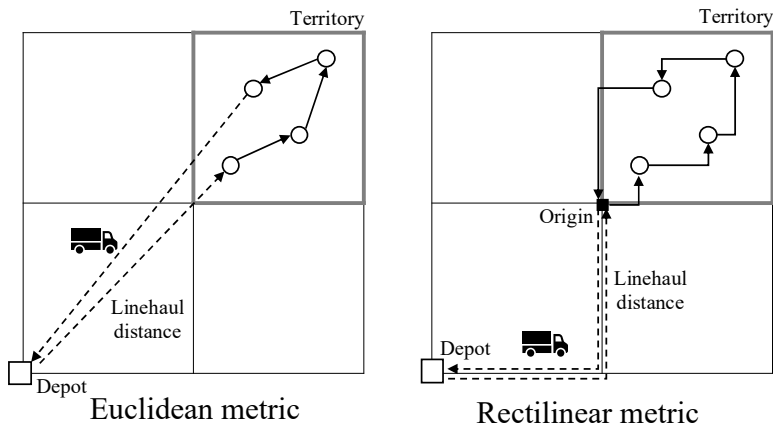
1 Introduction

In logistics, suppliers often designate separate pick-up and delivery territories for each driver. There are two reasons to set driver territories (Kovacs et al., 2014). One is customer satisfaction. Customers who are visited frequently want to be delivered by the same familiar driver. Another reason is that designating each driver's territory not only improves employee satisfaction, but also employee productivity by eliminating the need to learn new routes and customers. Consequently, managers periodically review their territories to minimise the total working time of all drivers. For example, in Japan, an average customer replaces a gas cylinder ten times a year. Then, for a driver, the customers to whom cylinders are delivered in each trip differ; that is, the customers are stochastic. A typical driver changes 30–40 cylinders per trip. To design the territories, we estimated the average tour length in the rectangle for a small number of customers (less than 100 customers) by the continuous approximation approach and the rectilinear metric, where the origin is located at the corner of the rectangle.

Finding the shortest travel distance is known as the travelling salesman problem (TSP) (Applegate et al., 2006) and the vehicle routing problem (VRP). With improvements in computer performance and algorithms, the number of problems that can be optimally solved has increased. However, multiple calculations are still required in cases where the number of demand points is excessively high, when the exact locations of the demand points are unknown, or when the locations of the demand points change daily. In such cases, a continuous approximation model can be used. A continuous approximation model is characterised by its use of continuous representations of input data and decision variables as density functions over time and space, and the goal is to approximate the objective function into an expression that can be optimised by relatively simple analytical operations (Carlsson et al., 2023). In this study, we apply this model to estimate not only the average routing distance for each territory, but also the total routing distance in the area with high accuracy. Franceschetti et al. (2017) reviewed approximation models for freight transportation management. They also suggested that most applications of continuous approximation models can be categorised into districting, location, fleet composition, and vehicle routing. Note that the definitions of districting and territory design for our purpose are the same as those used in their study. Ansari et al. (2018) also reviewed them for the last 20 years. As recent studies, Álvarez-Miranda and Pereira (2021) have proposed a hybrid method for districting with a quality-of-service objective. Banerjee et al. (2022) have applied the model for same-day delivery systems to decide fleet size and territory design. Zhang et al. (2024) have provided the application for integrated truck and drone systems. In their previous studies, the models mainly assumed that customers are generated uniformly and randomly in the planar area. Although this assumption is a primitive condition, it has been providing insight for applying for the delivery plan in the city and each area.

Several previous studies have approximated tour distances using the Euclidean distance metric (Kwon et al., 1995). However, we focused on using the rectilinear distance metric because of its practical ease of use to evaluate the shape of each territory. Figure 1 shows an example of territory design. If we assume a Euclidean metric between customers and the depot, we need to estimate the linehaul distance, which is the distance from the depot to the first or last customers in the rectangle. According to the position of depot and territory, the accuracy of tour length estimations is sometimes poor. Therefore, in this study, we used rectilinear distance estimation. This distance metric does not require distance estimation for linehaul distances, imitating travel in a real urban road network such as New York, Kyoto and Barcelona. After calculating the linehaul distance from the depot to the origin, which is the nearest corner of the rectangle as shown in Figure 1, we added the calculated distance and estimated tour length to visit all customers and return to the origin in the rectangle. However, for the distance between any two points, the ratio of the rectilinear distance to the Euclidean distance is 1.280 for a rectangle (Koshizuka and Kobayashi, 1983) and 1.273 for a circle (Kurita, 2001). Therefore, the differences in distance metrics affect the TSP tour distance through multiple points.

Figure 1 Tour from depots to stochastic customers



Therefore, in this study, we focus on estimating the tour length from the origin to the customers randomly generated in the territory. To design the territories, we estimated the average tour length in the rectangle for a small number of customers (less than 100 customers) using the continuous approximation model by the rectilinear metric, which considers the horizontal and vertical differences as distances, and the origin is located at the corner of the rectangle. The remainder of this paper is organised as follows: Section 2 provides a literature review of the TSP length using the continuous approximation approach. Section 3 describes the estimation procedure. Section 4 describes the experimental design and the validation results. Section 5 presents the application to territory design and analysis in accordance with the number of territories and length-width ratio. Finally, the conclusions are presented in Section 6.

2 Literature review

In this study, we first introduce the previous studies to estimate the optimal TSP tour length, the length of the shortest tour covering all points, in continuous approximation models. Let L_1^* and L_2^* denote the approximate tour lengths for the rectilinear and Euclidean metric, respectively, in each previous study. The approximation of the optimal length has been studied since the 1940s. Several previous studies have assumed that the area is a unit square $[0, 1] \times [0, 1]$ and n points are uniformly and independently distributed. Beardwood et al. (1959) showed that

$$\lim_{n \rightarrow \infty} \frac{L_2^*}{\sqrt{n}} = \beta_{L2}, \quad (1)$$

with probability one for a constant β_{L2} . Several researchers have given different estimates of β_{L2} using various methodologies (Vig and Palekar, 2008). Applegate et al. (2006) estimated β_{L2} using ordinary least squares (OLS) regression for 100 to 2,000 points and obtained β_{L2} ranging from 0.7256 to 0.7765. Lei et al. (2015) applied a similar approach to Applegate et al. (2006) for 20 to 90 points and obtained β_{L2} ranging from 0.7774 to 0.8584, i.e., β_{L2} has a downward trend as points increase. Then, Choi et al. (2024) have proposed installing an adjustment factor depending on n for β_{L2} .

Some researchers also estimated β_{L1} , the constant in the rectilinear metric. Daganzo (1984) derived upper bounds of β_{L1} and β_{L2} are 1.15 and β_{L2} respectively. Also, Johnson et al. (1996) estimated β_{L1} to be 0.8943 ± 0.0007 while β_{L2} was 0.7124 ± 0.0002 . However, they did not show the calculation of β_{L1} in detail. Vig and Palekar (2008) showed that:

$$L_1^* = 0.765 + 0.892\sqrt{n}, \quad (2)$$

and

$$L_2^* = 0.662 + 0.710\sqrt{n} \quad (3)$$

using OLS regression for 11 to 2,000 points and proposed that $\beta_{L1} = 0.892$ and $\beta_{L2} = 0.710$.

Çavdar and Sokol (2015) proposed the tour length formula for the dispersion of any points in the Euclidean metric. The points are scattered in a rectangle with central horizontal and vertical axes. Using regression by tour length for 3,000–8,000 points in 27 rectangles, they proposed the tour length formula L_2^* as:

$$L_2^* = 2.791\sqrt{n(\sigma_{lx}\sigma_{ly})} + 0.267\sqrt{\frac{nA(\sigma_x\sigma_y)}{\bar{c}_x\bar{c}_y}} \quad (4)$$

where $\bar{c}_x$ and $\bar{c}_y$ are the average distance of points to the central horizontal and vertical axes, σ_x and σ_y are the standard deviation of the horizontal and vertical coordinates of the points, σ_{lx} and σ_{ly} are the standard deviation of the absolute distances of the points from the central horizontal and vertical axes, and A is the area of a rectangle.

Meanwhile, some researchers have suggested that n is small during actual deliveries. Vig and Palekar (2008) demonstrated that repair people travel between 15 and 20 stops in

a single trip. Holguín-Veras and Patil (2005) suggested that 35% of all trip chains have only two stops and that the maximum number of stops in a trip chain is 23 for commercial vehicles in Denver. In Japan, as mentioned in the first section, a driver replaces 30–40 gas cylinders per trip. In addition, the person for periodic inspection for gas appliances travels at most ten stops. A home delivery driver for an online supermarket travels a maximum of 20 stops. Therefore, to discuss each driver's territory in practice, tour length formula should be created for a small n . In fact, Choi and Schonfeld (2022) presented a comprehensive review to approximate the TSP tour length for a small n .

Kou et al. (2022) created 90 instances, precisely ten instances for each of the three square sizes (1×1 , 2×2 and 5×5) and each of three kinds of points (10, 20 and 50), and estimated the TSP tour length by a regression analysis as

$$\ln L_2^* = 1.506 + 0.940 \ln std_{RT}, \quad (5)$$

$$L_2^* = 1.564 + 0.904\sqrt{nA}, \quad (6)$$

and

$$\ln L_1^* = 1.440 + 0.933 \ln std_{RT}, \quad (7)$$

where std_{RT} the standard deviation of tour lengths for 1,000 random tours. That is, std_{RT} is needed to compute L_1^* .

Two well-known studies have discussed the approximation of the TSP tour length for a small n in a rectangle. Chien (1992) estimated the TSP tour length using the Euclidean metric in eight rectangles with a length-width ratio λ of 1 to 8 and eight sectorial-shaped regions. They produced 4,160 instances in which 5–30 customers were randomly distributed, and the origin was one of the corners of the region. For the ease of convenience in the following section, we now use the tour length $t_{L2}^*(\lambda, n)$ instead of L_2^* . After seven tour length estimators are considered and evaluated, $t_{L2}^*(\lambda, n)$ is proposed as:

$$t_{L2}^*(\lambda, n) = 2.10\bar{r} + 0.67\sqrt{(n-1)A_1}, \quad (8)$$

where $\bar{r}$ is the average Euclidean distance from customers to the depot, and A_1 is the area of the smallest rectangle that covers only the customers. $t_{L2}^*(\lambda, n)$ was obtained with high accuracy (adjusted $R^2 = 0.99$). In addition, they presented

$$t_{L2}^*(\lambda, n) = a_1\bar{r} + a_2\sqrt{(n-1)A_1}, \quad (9)$$

in the rectangle corresponding to λ , as shown in Table 1.

Table 1 Results of $t_{L2}^*(\lambda, n)$

Coefficient	$t_{L2}^*(1, n)$	$t_{L2}^*(2, n)$	$t_{L2}^*(3, n)$	$t_{L2}^*(4, n)$
a_1	2.39	2.26	2.25	2.41
a_2	0.57	0.59	0.58	0.54
Coefficient	$t_{L2}^*(5, n)$	$t_{L2}^*(6, n)$	$t_{L2}^*(7, n)$	$t_{L2}^*(8, n)$
a_1	2.45	2.38	2.53	2.57
a_2	0.54	0.57	0.54	0.54

Kwon et al. (1995) also estimated the tour length in the rectangle whose length-width ratio λ is 1 to 8 and whose customers from 10 to 89 are randomly distributed. Then, the TSP tour length $t_{L2}^*(\lambda, n, A)$ of the rectangle was proposed as

$$t_{L2}^*(\lambda, n, A) = \left(0.833 - 0.001n + \frac{1.115\lambda}{n} \right) \sqrt{nA}, \quad (10)$$

or

$$t_{L2}^*(\lambda, n, A) = 0.414\bar{r} + \left(0.775 - 0.0008n + \frac{0.903\lambda}{n} \right) \sqrt{nA}, \quad (11)$$

where A is the area of the rectangle with high accuracy ($R^2 = 0.998$). Inspired by Kwon et al. (1995), our research is a more practical and general formulation of the rectilinear distance metric, taking into account the length-width ratio, for application to territory design.

3 Our approach

3.1 Notation

We make use of the following notation throughout the paper:

N	set of the number of points (demand points and the origin)
Λ	set of length-width ratios
A	area of a driver's territory
n, n'	number of points (demand points and the origin)
m	number of instances
λ	length-width ratio
a_1, a_2, b	coefficient for independent variables
$t_{\lambda, n, i}$	rectilinear tour length for n points with the length-width ratio λ in instance i
$\bar{t}_{\lambda, n}$	average tour length for n points with the length-width ratio λ for all instances
$t_{L1}^*(\lambda, n)$	approximate rectilinear tour length for n points with the length-width ratio λ and an area of one
$t_{L1}^*(\lambda, n, A)$	approximate rectilinear tour length for n points with the length-width ratio λ and an area of A .

3.2 Experimental procedures

Several methods can be used to approximate the TSP tour length. In particular, there are two main methods. One method is to perform OLS regression on all data, whereas the

other is to summarise the data, for example, by averaging and then performing OLS regression. Considering our objective, we used the average TSP tour length in a rectangle. Our procedure was based on that of Choi and Schonfeld (2022).

First, $n - 1$ demand points were generated according to a given distribution in a rectangle with an area of one. For $\lambda \in \Lambda$, by generating two random numbers uniformly distributed in the interval $[0, \sqrt{\lambda}]$ and $[0, 1/\sqrt{\lambda}]$, respectively, the numbers are regarded as the x - and y -coordinates of a demand point in the rectangle, i.e., each demand point in the rectangle $[0, \sqrt{\lambda}] \times [0, 1/\sqrt{\lambda}]$ is equally likely to be selected. In addition, one point is generated at the corner of the rectangle as the origin.

Second, we selected a solution method to compute the rectilinear TSP tour length for n points. In this study, we computed the rectilinear tour length using Helsgaun's implementation of the Lin-Kernighan heuristic (LKH) algorithm by Helsgaun (2000). Vig and Palekar (2008) also applied this algorithm because Johnson and McGeoch (2002) showed that the solutions from this algorithm are typically very close to the optimal solution.

Third, m instances of the given n points and the length-width ratio λ were produced. After the rectilinear TSP tour length $t_{\lambda,n,i}$ were computed by for every instance, the rectilinear tour lengths for each n and λ were averaged, i.e.,

$$\bar{t}_{\lambda,n} = \frac{1}{m} \sum_{i=1}^m t_{\lambda,n,i}. \quad (12)$$

Subsequently, the repeated runs moved to other n and λ . Finally, after the averaged rectilinear tour length was fitted using OLS regression, the tour length formula $t_{L1}^*(\lambda, n)$ was obtained.

3.3 Features of the tour length

In this study, we assumed that $N = \{11, 21, \dots, 91, 101\}$, $\Lambda = \{1, 2, 3, 4, 5\}$, i.e., the maximum length-width ratio is five and 500 instances for each $n \in N$ and $\lambda \in \Lambda$.

Figure 2 shows the result of $\bar{t}_{\lambda,n}$. This result indicates that the average length is proportional to the concave function of n . Previous studies used the factor $\sqrt{n}$, which we considered applying in the TSP tour approximation. Additionally, as the number of demand points increased, the effect of the length-width ratio on the average length decreased. In fact, $\bar{t}_{5,101} - \bar{t}_{1,101} = 0.3$ while $\bar{t}_{5,11} - \bar{t}_{1,11} = 1.1$. We must then consider the factor proportional to the length-width ratio and inversely proportional to the points.

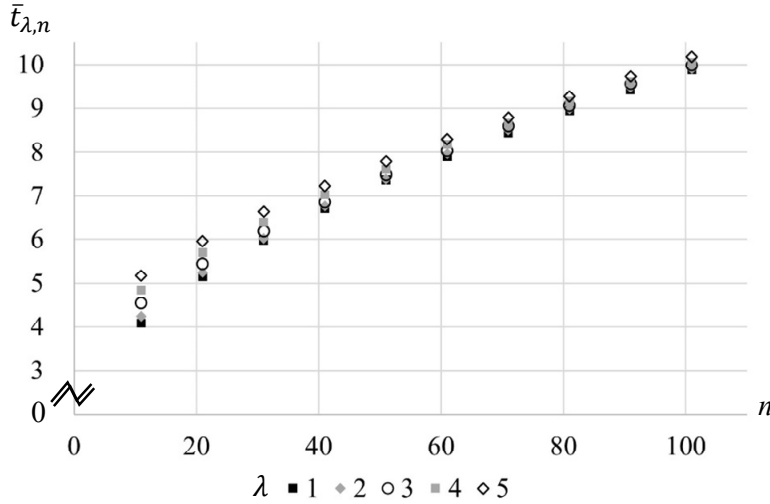
4 Regression-based approximations

4.1 Formula in the unit rectangle

In this study, we investigated a simple approximation formula with high accuracy. Since our objective is territory design, we need the simple tour length formula. According to the previous studies described in Section 2, candidate independent variables are $\sqrt{n}$, $\sqrt{n-1}$, $n\sqrt{n}$, $\lambda/\sqrt{n}$, $1/\sqrt{\lambda}$, $\sqrt{\lambda}$ and constant term. In addition, $\sqrt{\lambda/n}$, λ/n and λ are also candidate variables. Then, we sought independent variables where the number of them is

minimum and the adjusted coefficient of determination (adjusted R^2) is more than 0.99 by running the OLS regression.

Figure 2 Average tour length



As a result, we discovered the following TSP tour length approximation formula (model 1):

$$t_{L1}^*(\lambda, n) = a_1 \sqrt{n} + \frac{a_2 \lambda}{\sqrt{n}} + b, \quad (13)$$

where the area of the rectangle is one; the estimated parameters a_1 and a_2 are 0.880 and 0.872, respectively; and the constraint b is 0.875, with an adjusted $R^2 = 0.999$. In this study, we applied three measures to estimate quality, in addition to R^2 . The mean percentage error (MPE) is defined as

$$MPE = \frac{1}{|N||\Lambda|} \sum_{\lambda \in \Lambda, n \in N} \frac{\bar{t}_{\lambda,n} - t_{L1}^*(\lambda, n)}{\bar{t}_{\lambda,n}} \times 100\%; \quad (14)$$

the mean absolute percentage error (MAPE) is defined as

$$MAPE = \frac{1}{|N||\Lambda|} \sum_{\lambda \in \Lambda, n \in N} \frac{|\bar{t}_{\lambda,n} - t_{L1}^*(\lambda, n)|}{\bar{t}_{\lambda,n}} \times 100\%; \quad (15)$$

and the maximum absolute percentage error (MaxAPE) is defined as

$$MaxAPE = \max_{\lambda \in \Lambda, n \in N} \sum_{\lambda \in \Lambda, n \in N} \frac{|\bar{t}_{\lambda,n} - t_{L1}^*(\lambda, n)|}{\bar{t}_{\lambda,n}} \times 100\%. \quad (16)$$

We also proposed another formula (model 2), which is a simplified version of model 1 since a_1 , a_2 and b are similar, i.e., a_2 and b in equation (13) are replaced by a_1 , as:

$$t_{L1}^*(\lambda, n) = a_1 \left(\sqrt{n} + \frac{\lambda}{\sqrt{n}} + 1 \right). \quad (17)$$

The statistics for model 1 are shown in Table 2, and the three measures for models 1 and 2 are listed in Table 3. Table 2 shows that the standard errors were small for all the coefficients. In addition, the MPE, MAPE, and MaxAPE were small for both equations. Moreover, model 1 had low measures. In particular, MaxAPE is an important factor when considering territory design; the values for both equations are 2.0% or less.

Table 2 Statistics of model 1

<i>Parameter</i>	<i>Coefficient</i>	<i>Standard error</i>	<i>t-value</i>	<i>Pr (> t)</i>
a_1	0.880	0.0353	24.8	<0.001
a_2	0.872	0.0036	240.0	<0.001
b	0.875	0.0259	33.7	<0.001

Table 3 Fit comparison for two equations

<i>Formula</i>	<i>MPE</i>	<i>MAPE</i>	<i>MaxAPE</i>
Model 1	0.006%	0.5%	1.8%
Model 2	0.1%	0.5%	2.0%

Table 4 Measures for other instances

<i>Formula</i>	<i>MPE</i>	<i>MAPE</i>	<i>MaxAPE</i>
Model 1	0.05%	0.6%	1.9%
Model 2	0.7%	0.6%	1.8%

To evaluate the robustness, we generated other $\bar{t}_{\lambda,n}$ from 200 instances each and computed three measures, as shown in Table 4. The MPE for both formulas become worse, however, the values were small. These results indicate that the proposed formulas perform well.

4.2 Formula in arbitrary rectangle

In the previous section, our experiment assumed that the area of the rectangle was one. Then, we finally revise models where the area of the rectangle is A . Since each demand point in the rectangle $[0, \sqrt{A\lambda}] \times [0, \sqrt{A/\lambda}]$ is equally likely to be selected, the approximation models become:

$$t_{L1}^*(\lambda, n, A) = \left(a_1 \sqrt{n} + \frac{a_2 \lambda}{\sqrt{n}} + b \right) \sqrt{A} \quad (18)$$

as model 1 and

$$t_{L1}^*(\lambda, n, A) = a_1 \left(\sqrt{n} + \frac{\lambda}{\sqrt{n}} + 1 \right) \sqrt{A}, \quad (19)$$

as model 2 respectively. The independent variables of equation (19) are similar to those of equation (10) and both include the term $\sqrt{nA}$, but equation (10) includes the term $n\sqrt{nA}$ while equation (19) includes the term $\sqrt{A}$.

As mentioned in the first section, our objective was to model to evaluate the territory design. For example, in Figure 3, we consider a territory divided into six equally sized rectangles. Since each size of territory is identical, we only need to calculate the sum of the linehaul distance and $a_1\lambda/\sqrt{n}$ for each territory. Therefore, our approximation model can be easily used in territory design.

5 Application for territory design

The design of the territory is one of the most important factors in devising for strategies for distribution company. When there are many customers to a given area, one tour route can provide the shortest tour distance which means only one driver is needed for a delivery. However, this assumption may not be possible due to driver's working hours or the size of the vehicle, in other words time and capacity constraints. In this section, we show examples of applications of our model to actual territory designs. Specifically, we analyses how much the total tour distance increases with the dividing of the zone using our proposed model.

5.1 Problem settings

In this section, the following additional notation is used.

P number of horizontal divisions in a square zone

Q number of vertical divisions in a square zone

n_{all} total number of demand points

A_{all} area of a distributor's square service zone (the sum of the areas of all territories).

Suppose that a distributor's square service zone is divided into PQ equal parts (territories), i.e., $A_{all} = PQA$. In each territory, $n - 1$ demand points are uniformly and randomly generated, resulting in a total of $n_{all} = (n - 1) PQ$ points in the service zone. In this section, we analysed how the shape of the territory, as defined P and Q , affects the total tour distance. This includes both the tour distance within the territory, as formulated in equation (19) in the previous section, and the linehaul distance where $P \leq Q$. Figure 3 provides an example of two shapes of territories with six equal parts. The shape of the territory becomes more and more square as P becomes closer to Q .

In Figure 3, the tour and linehaul distances are represented by solid and dashed lines respectively. Let $Z(P, Q, A_{all}, n_{all})$ where $n_{all} = (n - 1) PQ$ be the total tour distance in the service zone. Then, $Z(P, Q, A_{all}, n_{all})$ is the sum of tour distance function $t_{L1}^*(\lambda, n, A_{all})$ and linehaul distance function $l(p, q, P, Q, A_{all})$ as follows.

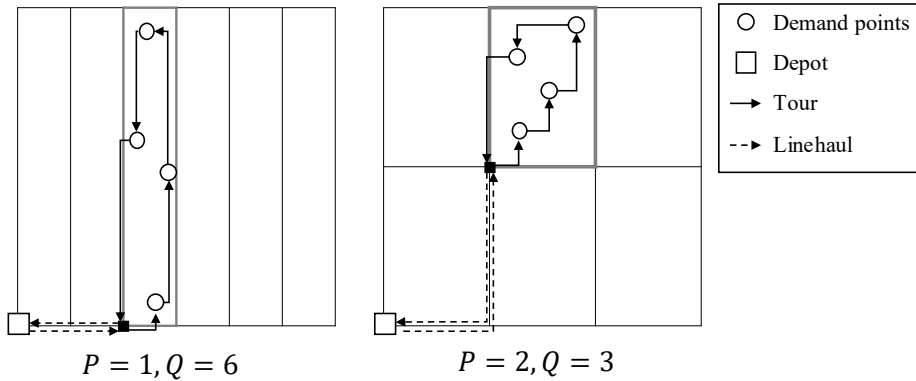
$$Z(P, Q, A_{all}, n_{all}) = \sum_{p=1}^{P-1} \sum_{q=0}^{Q-1} \left(2l(p, q, P, Q, A_{all}) + t_{L1}^* \left(\frac{Q}{P}, n, \frac{A_{all}}{PQ} \right) \right), \quad (20)$$

where

$$l(p, q, P, Q, A_{all}) = \frac{p\sqrt{A_{all}}}{P} + \frac{q\sqrt{A_{all}}}{Q}. \quad (21)$$

The linehaul distance $l(p, q, P, Q, A_{all})$ represents the rectilinear distance from the depot to the nearest node (corner) of territory from the depot.

Figure 3 Territory design



Then, equation (20) becomes

$$\begin{aligned} Z(P, Q, A_{all}, n_{all}) &= \sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} \left(\frac{2p\sqrt{A_{all}}}{P} + \frac{2q\sqrt{A_{all}}}{Q} + a_1 \left(\sqrt{n} + \frac{Q}{P\sqrt{n}} + 1 \right) \sqrt{\frac{A_{all}}{PQ}} \right) \\ &= \sqrt{A_{all}} (2PQ - P - Q) + a_1 \sqrt{A_{all}PQ} \left(1 + \sqrt{n} + \frac{Q}{P\sqrt{n}} \right), \end{aligned} \quad (22)$$

where

$$n_{all} = (n-1)PQ.$$

5.2 Impact of zone dividing on tour and linehaul distance

In this section, we present the effect of increasing the distance using territory design. We observed the total tour distance according to Q . If $P = 1$, then the total tour distance $Z(P, Q, A_{all}, n_{all})$ is given by

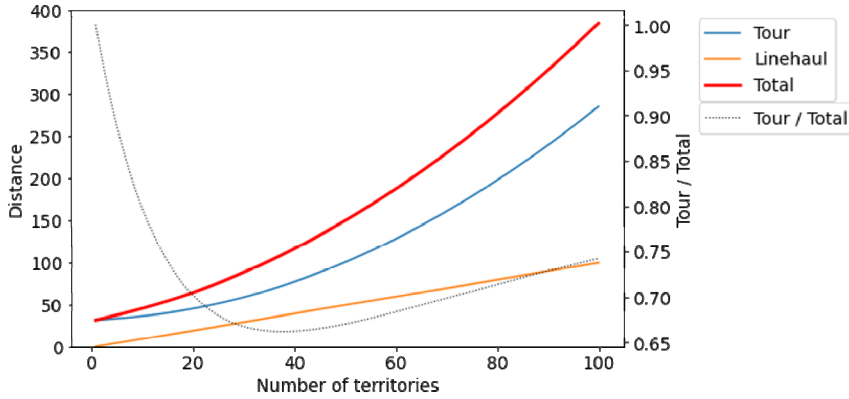
$$\begin{aligned} Z(1, Q, A_{all}, n_{all}) &= \sum_{q=0}^{Q-1} \left(\frac{2q\sqrt{A_{all}}}{Q} + a_1 \left(\sqrt{\frac{n_{all}}{Q}} + 1 + \frac{Q}{\sqrt{n_{all}/Q+1}} + 1 \right) \sqrt{\frac{A_{all}}{Q}} \right) \\ &= \sqrt{A_{all}} (Q-1) + a_1 \sqrt{A_{all}Q} \left(\sqrt{\frac{n_{all}}{Q}} + 1 + \frac{Q}{\sqrt{n_{all}/Q+1}} + 1 \right) \end{aligned} \quad (23)$$

The first term is the linehaul distance, and the second term represents the tour distance within every territory. Figure 4 shows the results of the tour, linehaul, total distance

which is the sum of them, and the ratio of tour and total distance where $a_1 = 0.88$, $n_{all} = 1,200$, $A_{all} = 1$, and number of territories (it is Q in this case) is between 1 and 100.

In Figure 4 there is a clear trend of increasing the total tour when the number of territories was expanded. And it was caused by expanding increase in tour distances. This is larger than the linear increase in linehaul distance. Therefore, the percentage of tour distance increased after a decrease.

Figure 4 Total and tour distance ($a_1 = 0.88$, $n_{all} = 1,200$, $A_{all} = 1$, $P = 1$) (see online version for colours)



Next, we calculated the distances for the opposite example with $P = 1$, where the territory is designed to be divided into squares as much as possible. We obtained the distance where the shape of territory is a square-like shape ($P = Q$ and $\lambda = 1.0$) such as

$$\begin{aligned}
 Z(\sqrt{Q}, \sqrt{Q}, A_{all}, n_{all}) &= 2\sqrt{A_{all}}(Q - \sqrt{Q}) \\
 &+ a_1 \sqrt{A_{all}Q} \left(\sqrt{\frac{n_{all}}{Q}} + 1 + \frac{1}{\sqrt{n_{all}/Q + 1}} + 1 \right),
 \end{aligned} \tag{24}$$

Figure 5 shows the same distances and ratio as in Figure 4 when $P = Q$. The increase in total distance with the number of territories is linear compared to the case of $P = 1$, as shown in Figure 4. In particular, the linear increase in tour distance is a significant difference.

At last, we compared Figure 5 to the result of the largest λ (largest Q where $P = 1$) in Figure 4. Figure 6 shows a comparison of the total distances between $P = 1$, $P = 2$ and $P = Q$. If the number of territories is small ($PQ < 27$ for this condition), $P = 1$ results in a shorter total distance. Nevertheless, the difference is not significant; therefore, the effect of the territory design which adjust of the length-width ratio is small when fewer territories are designed, while an increase in the number of territories has a larger impact. On the other hand, the difference in the total distance increases when the number of territories is large, and square-like territory travels less distance. However, the increase of total distance is reduced if $P = 2$. In the empirical calculations in Section 3, we showed that the total distance is shorter for square-like shapes, where the length-width ratio of the territory is smaller. Furthermore, this numerical analysis shows that the increase in tour distance increases as λ increases above 30, thus the total distance is shorter for small λ .

Figure 5 Total and tour distance ($a_1 = 0.88$, $n_{all} = 1,200$, $A_{all} = 1$, $P = Q$) (see online version for colours)

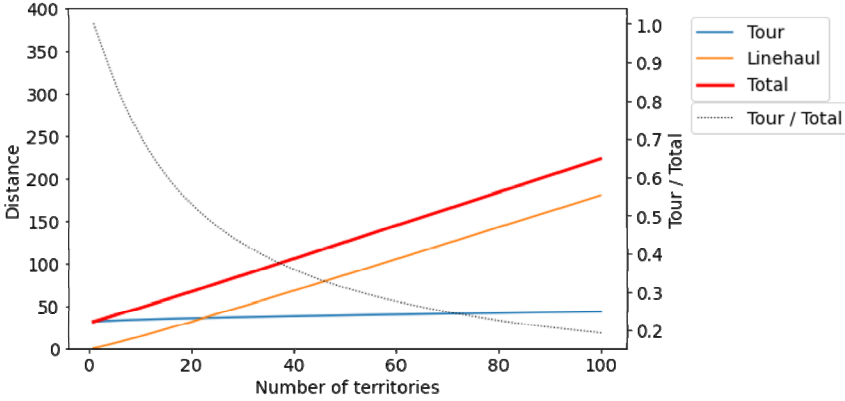
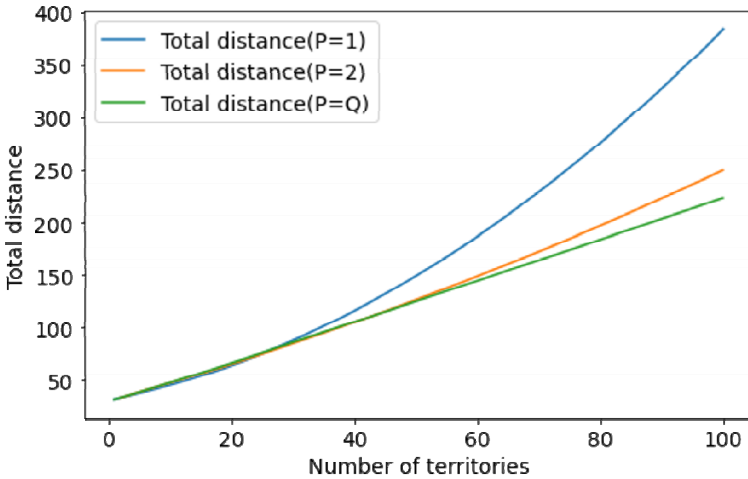


Figure 6 Comparison of total distance in shape of territories ($a_1 = 0.88$, $n_{all} = 1,200$, $A_{all} = 1$) (see online version for colours)



6 Conclusions

In this study, we proposed a continuous approximation model for touring distance based on the rectilinear distance metric, which has the advantage of reduced computational cost. To evaluate the accuracy of our model, we varied the number of tour points and length-width ratio of the region and then conducted numerical simulations.

The results indicate that our model can accurately estimate tour distances. In addition, we observed that the influence of the length-width ratio decreased as the number of demand points increased. The results are not only relevant to cities with gridded streets but also prove to be effective for developing delivery plans in various urban environments. Numerical experiments on territory designing demonstrated the

relationship between the number of territories, length-width ratio of the zone, and tour and linehaul distances. And, number of territories have an impact for increasing tour and linehaul distance, but these increase are lower if length-width ratio was less than 30.

In future studies, our model will be extended to support VRP. The model will be used for more realistic territory design and delivery planning by considering not only the way areas are divided but also the modelling of time specifications and capacity constraints.

Acknowledgements

This work was supported by JSPS KAKENHI [Grant Nos. JP21K14314, JP23K04269, and JP24K17413]. The authors appreciate their support.

References

- Álvarez-Miranda, E. and Pereira, J. (2021) 'A districting application with a quality of service objective', *Mathematics*, Vol. 10, No. 1, p.13.
- Ansari, S., Başdere, M., Li, X., Ouyang, Y. and Smilowitz, K. (2018) 'Advancements in continuous approximation models for logistics and transportation systems: 1996–2016', *Transportation Research Part B: Methodological*, Vol. 107, pp.229–252.
- Applegate, D.L., Bixby, R.E., Chvátal, V. and Cook, W.J. (2006) *The Traveling Salesman Problem: A Computational Study*, Princeton University Press, Princeton, New Jersey.
- Banerjee, D., Erera, A.L. and Toriello, A. (2022) 'Fleet sizing and service region partitioning for same-day delivery systems', *Transportation Science*, Vol. 56, No. 5, pp.1327–1347.
- Beardwood, J., Halton, J.H. and Hammersley, J.M. (1959) 'The shortest path through many points', *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 55, No. 4, pp.299–327, DOI: 10.1017/S0305004100034095.
- Carlsson, J.G., Peng, Y., Center, M.T. and Region, P.S. (2023) *Continuous Approximation Models with Temporal Constraints and Objectives*, University of Southern California, No. PSR-22-19.
- Çavdar, B. and Sokol, J. (2015) 'A distribution-free TSP tour length estimation model for random graphs', *European Journal of Operational Research*, Vol. 243, No. 2, pp.588–598, DOI: 10.1016/j.ejor.2014.12.020.
- Chien, T.W. (1992) 'Operational estimators for the length of a traveling salesman tour', *Computers & Operations Research*, Vol. 19, No. 6, pp.469–478, DOI: 10.1016/0305-0548(92)90002-M.
- Choi, Y. and Schonfeld, P.M. (2022) 'Review of length approximations for tours with few stops', *Transportation Research Record*, Vol. 2676, No. 3, pp.201–213, DOI: 10.1177/03611981211049433.
- Choi, Y., Golden, B.L. and Schonfeld, P.M. (2024) 'Improving analytic approximations of TSP tour lengths with adjustment factors', *Transportmetrica A: Transport Science*, p.2346631, DOI: 10.1080/23249935.2024.2346631.
- Daganzo, C.F. (1984) 'The length of tours in zones of different shapes', *Transportation Research Part B: Methodological*, Vol. 18, No. 2, pp.135–145, DOI: 10.1016/0191-2615(84)90027-4.
- Franceschetti, A., Jabali, O. and Laporte, G. (2017) 'Continuous approximation models in freight distribution management', *Top*, Vol. 25, pp.413–433, DOI: 10.1007/s11750-017-0456-1.
- Helsgaun, K. (2000) 'An effective implementation of the Lin-Kernighan traveling salesman heuristic', *European Journal of Operational Research*, Vol. 126, No. 1, pp.106–130, DOI: 10.1016/S0377-2217(99)00284-2.

- Holguín-Veras, J. and Patil, G.R. (2005) 'Observed trip chain behavior of commercial vehicles', *Transportation Research Record*, Vol. 1906, No. 1, pp.74–80, DOI: 10.1177/0361198105190600109.
- Johnson, D.S. and McGeoch, L.A. (2002) 'Experimental analysis of heuristic for the STSP', in Gutin, G. and Punnen, A.P. (Eds.): *The Traveling Salesman Problems and Its Variations*, pp.369–443, Springer, Kluwer Academic Publishers, New York, New York.
- Johnson, D.S., McGeoch, L.A. and Rothberg, E.E. (1996) 'Asymptotic experimental analysis for the Held-Karp traveling salesman bound', in *The Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms*, Atlanta, Georgia, pp.341–350.
- Koshizuka, T. and Kobayashi, J. (1983) 'On the relation between road distance and Euclidean distance', *Journal of the City Planning Institute of Japan*, in Japanese, Vol. 18, pp.43–48, DOI: 10.11361/journalcpj.18.43.
- Kou, S., Golden, B. and Poikonen, S. (2022) 'Optimal TSP tour length estimation using standard deviation as a predictor', *Computers & Operations Research*, Vol. 148, p.105993, DOI: 10.1016/j.cor.2022.105993.
- Kovacs, A.A., Golden, B.L., Hartl, R.F. and Parragh, S.N. (2014) 'Vehicle routing problems in which consistency considerations are important: a survey', *Networks*, Vol. 64, pp.192–213, DOI: 10.1002/net.21565.
- Kurita, O. (2001) 'Theory of road patterns for a circular disk city', *Journal of the City Planning Institute of Japan*, in Japanese, Vol. 36, pp.859–864, DOI: 10.11361/journalcpj.36.859.
- Kwon, O., Golden, B. and Wasil, E. (1995) 'Estimating the length of the optimal TSP tour: an empirical study using regression and neural networks', *Computers & Operations Research*, Vol. 22, No. 10, pp.1039–1046, DOI: 10.1016/0305-0548(94)00093-N.
- Lei, H., Laporte, G., Liu, Y. and Zhang, T. (2015) 'Dynamic design of sales territories', *Computers & Operations Research*, Vol. 56, pp.84–92.
- Vig, V. and Palekar, U.S. (2008) 'On estimating the distribution of optimal traveling salesman tour lengths using heuristics', *European Journal of Operational Research*, Vol. 186, No. 1, pp.111–119, DOI: 10.1016/j.ejor.2006.12.066.
- Zhang, J., Campbell, J.F. and Sweeney II, D.C.(2024) 'A continuous approximation approach to integrated truck and drone delivery systems', *Omega*, Vol. 126, p.103067.