https://doi.org/10.1016/j.trc.2019.09.004

Optimization of Traffic Count Locations for Estimation of Travel Demands with Covariance between Origin-Destination Flows

Hao Fu [a], William H. K. Lam [a*], Hu Shao [a,b], X.P. Xu [a,b], H.P. Lo[a], Bi Yu Chen [a,c], N. N. Sze [a], Agachai Sumalee [a]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[b] School of Mathematics, China University of Mining and Technology, Xuzhou, Jiangsu, 221116, China
[c] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China
[*]Corresponding author: cehklam@polyu.edu.hk

Abstract

Vehicular traffic between different Origin-Destination (OD) pairs for a typical hourly period may statistically correlate with each other. The covariance mainly generated from the daily variation of travel patterns, network topology, and trip chaining activities of household members can be particularly high during the morning peak hour. With the increasing attention on the OD demand variance and covariance in stochastic road networks, a new criterion is proposed in this paper for measuring the estimation accuracy of OD demand covariance. The mathematical properties of this proposed criterion are analyzed to better understand the relationship between the estimation errors of mean and covariance of OD demands. This paper aims to investigate how to optimize the traffic count locations for minimizing the weighted maximum deviation of estimated mean and covariance of OD demands from the observed values. To consider the effects of stochastic OD demands on the traffic count location problem in the proposed model, link choice proportions are regarded as stochastic variables and updated by an adapted traffic flow simulator in this study. Both the weighted-sum approach and bi-objective approach are examined with the adaption of the firefly algorithm (FA) to solve the single-objective and bi-objective problems. Numerical examples are presented to demonstrate the effects, with and without considering the covariance of the OD demands for the optimization of traffic count locations.

Keywords: Traffic count locations, OD demand estimation, OD demand covariance, Bi-objective optimization

1. Introduction

1.1 Background and literature review

Estimation of OD demands from traffic counts has been commonly used for transport planning and traffic management purposes in past decades (Antoniou et al., 2016; Caggiani et al., 2013). Most of the existing studies on the traffic count location problem focus on the optimization of traffic count locations for the best estimates of mean OD demand (Gentili and Mirchandani, 2012; Hu and Liou, 2014; Bianco et al., 2001; Chootinan et al., 2005). However, the point estimate model, without considering the variance and covariance, would lead to biased OD estimates (Zhao and Kockelman, 2002).

Attributed to both the daily and seasonal variations in travel activity patterns, OD demands for a typical hourly period are not deterministic but actually fluctuate from day to day (Clark and Watling, 2005; Fei et al., 2013; Yin et al., 2009; Zhang et al., 2010). Apart from the promotion of daily variation of travel patterns and network topology, trip chaining, carpooling and ridesharing are also emerging contributors for the OD pair covariance.

Therefore, in the stochastic traffic networks, not only the mean OD demands but also the **covariance of OD demands** should be considered for OD estimation (Haas, 1999; Shao et al., 2015, 2014). Based on the observed link flows from day to day, statistical methods were used to estimate a sequence of OD demands with considering the covariance between different links (Vardi, 1996; Hazelton, 2008). Cantelmo et al. (2014) proposed a two-step approach in dynamic demand estimation problems with taking account the covariance of OD pairs. Yang et al. (2018) preserved all stochastic information obtained from different types of traffic sensors (e.g. link counts and route counts).

1.1.1 Traffic count location and OD demand estimation problems

The related studies in literature are categorized into three different types: (1) OD estimation problem; (2) traffic count location problem; and (3) combination of traffic count location and OD estimation problems as shown in Table 1. Our proposed model falls within the third type because of the inherent nature of the modeling approach.

Table 1 Categories of OD estimation and traffic count location problems

	Mean and/or variance of OD demands	Mean, variance, and **covariance of OD demands**
OD estimation problem or traffic count location problem	Yang (1995); Yang et al. (2001)	Shao et al. (2014); Shao et al. (2015)
Combination of traffic count location and OD estimation problems	Owais et al., (2019); Zhu et al., (2016)	**This paper**

To clearly review the relationship between the OD estimation and traffic count location models, the previous related studies are summarized as follows.

In literature, some of the existing approaches might only focus on the OD estimation problem while some measures or criteria were proposed for quantifying the traffic count location scheme (Yang et al., 1991; Lo et al., 1996). On the other hand, some previous studies only considered the traffic count location problem in which traffic count locations were directly determined based on the network structure/topology but without explicit use of prior OD information (Ehlert et al., 2006; Yang and Zhou, 1998). Note that in traffic count location problem, the link choice proportions can be adopted as fixed parameters without consideration of covariance of OD demands and link flows (Yang and Zhou, 1998).

Finally, there are some studies considering both traffic count location and OD estimation problems in which the traffic count locations were determined together with the estimation of OD demands. The objectives of these traffic count location models are to minimize the uncertainty of posterior OD demands (Zhou and List, 2010; Simonelli et al., 2012) or posterior link flows (S. Zhu et al., 2014). Note that the commonly used method to update deterministic link choice proportions in these models is the traffic assignment incorporated into OD estimation model as shown in Table 2 .

1.1.2 Assumptions of link choice proportions

In literature, fixed link choice proportions determined exogenously by AVI data or explicitly by traffic assignment methods are always assumed if congestion effects are not considered as shown in Table 2 (Owais et al., 2019; Simonelli et al., 2012; Yang and Zhou, 1998; Zhou and List, 2010; S. Zhu et al., 2014; Zhu et al., 2016). To consider the stochastic effects of congestion and updated OD demand estimates on the traffic count location problem, stochastic link choice proportions should be used and updated because the travelers will change the choices of their routes under different congestion conditions.

Table 2 Assumptions of link choice proportions in different problems

	Traffic count location problem	OD estimation problem	Combination of traffic count location and OD estimation problems
Fixed link choice proportions	Yang and Zhou (1998)	Van Zuylen and Willumsen (1980)	Simonelli et al. (2012)
Deterministic link choice proportions	--	Yang (1995); Yang et al. (2001)	Zhou and List (2010); Owais et al. (2019); Zhu et al. (2016); S. Zhu et al. (2014)
Stochastic link choice proportions	--	Lam and Xu (1999); Lo et al. (1996)	**This paper**

To relax the assumption of fixed link choice proportions, two different methods (1) considering deterministic link choice proportions by traffic assignment or (2) considering stochastic link choice proportions, were commonly used in literature, as shown in the above Table 2.

During the iterations for estimation of OD demands from traffic counts, SUE or UE traffic assignment methods can be used to update the deterministic link choice proportions based on OD demands estimated at each iteration (Shao et al., 2014; Yang, 1995; Yang et al., 2001). On the other hand, when stochastic effects of road traffic are taken into account, link choice proportions should also be considered as stochastic variables and updated using various methods. With explicit consideration of variation of travel times and OD demands, a traffic flow simulator was proposed for assessing the network reliability by Lam and Xu (1999).

To account for the impacts of stochastic OD demands and link flows, the link choice proportions are regarded as stochastic variables and updated by an adapted traffic flow simulator during the iteration process at the second stage of the proposed model. However, there are alternative methods such as the Reliability-based Stochastic User Equilibrium (R-SUE) traffic assignment model (Shao et al., 2006; Shen et al., 2019). The Automatic Vehicle Identification (AVI) data can also be used for direct estimation of the sample link choice proportions together with their variations. In our model, the traffic count location model is related to the quality of the estimated mean and covariance of OD demands. The accuracy of the OD demand estimation in terms of mean and covariance is captured by the concept of WMPREM and WMPREC proposed in this paper. Further elaboration on motivation and contribution of this paper would be given in the next section.

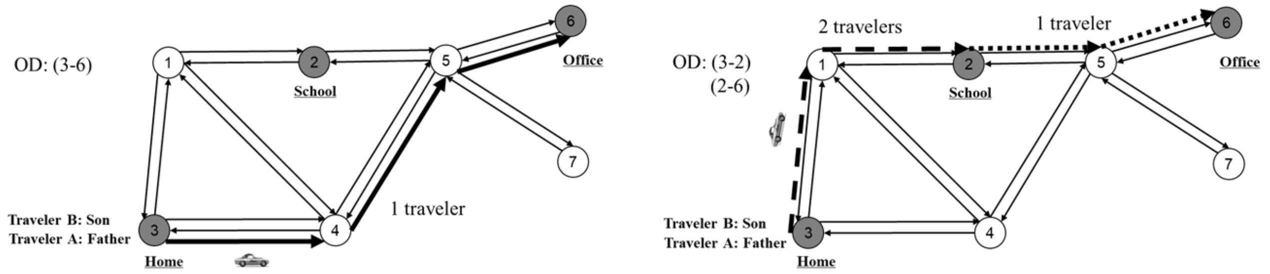
1.2 Motivations and contributions of the paper

1.2.1 Motivating example

As mentioned in the preceding Section 1.1, the demands of different OD pairs are not deterministic and independent. The variance and covariance of OD demands generated from multiple factors do exist in the road network. The relationship among different OD pairs should be considered when allocating the traffic sensors for stochastic OD demand estimation.

It is noted that the stochastic OD demands particularly for OD demand covariance are affected by the trip chaining activities. An illustrative example related to the relevant trip chaining activities is used to demonstrate the existence of covariance between different OD pairs and effects of stochastic OD demands on traffic count locations. Seven nodes and three OD pairs $\{(3,2),(2,6),(3,6)\}$ are included in this illustrative network, given that Node 2 stands for a school, Node 3 stands for a home, and Node 6 stands for an office, respectively.

In this example, suppose that there are two travelers - a father (Traveler A) and his son (Traveler B), Traveler A would like to travel from home to office (3-6) and Traveler B would like to travel from home to school (3-2) respectively. Fig. 1 illustrates the paths of the father, who has a car, in two different scenarios: (a) with no trip chaining, and (b) with trip chaining, respectively. As shown in Fig. 1(a), in the "no trip chaining" scenario, The father will travel from home (Node 3) to office (Node 6) directly. In contrast, as shown in Fig. 1(b), in "trip chaining" scenario, he will first travel to school (Node 2) together with his son, and then travel to office (Node 6) alone. The father and his son will share the same car for the journey from home (Node 3) to school (Node 2). When trip chaining exists, a journey between (3-6) will be split into two: (i) travel between (3-2), and (ii) travel between (2-6). The overall demands between (3-2) and (2-6) will therefore be correlated.



(a) No trip chaining scenario (b) Trip chaining scenario

Fig. 1 Travel paths in (a) no trip chaining and (b) trip chaining scenarios

We use W_1, W_2, and W_3 to denote the OD demands for (3-6), (3-2), and (2-6), respectively, given that W_1, W_2, and W_3 are independent and follow normal distributions, where $W_1 \sim N(200, 30^2)$, $W_2 \sim N(150, 36^2)$, and $W_3 \sim N(100, 25^2)$, respectively. We denote the proportion of trip chaining users as x, and the standard deviation of OD demands W_1, W_2, and W_3 as σ_1, σ_2, and σ_3, respectively. Therefore, the covariance between W_2 and W_3 could be deduced by the following equation.

$$\mathrm{cov}_{2,3} = x^2 \cdot \sigma_1^2 \quad (1)$$

Table 3 illustrates the relationship between proportion of trip chaining, covariance, and average vehicle occupancy.

Table 3 Relationship among the proportion of trip chaining users, covariance of OD demands, and average vehicle occupancy

Proportion of travelers using trip chaining	Covariance (W_2, W_3)	Average vehicle occupancy [a]
0.0%	**0.0**	**1.00**
17.8%	28.5	1.19
50.0%	225.0	1.80
100.0%	**900.0**	**2.25**

[a] average number of people in a vehicle, including the driver

As shown in Table 3, when the proportion of travelers using trip chaining increases, both the covariance between W_2 and W_3 and average vehicle occupancy increase. For instance, when there is no trip chaining (proportion equal to zero), covariance will be zero and vehicle occupancy will be 1 respectively. In contrast, when everyone uses trip chaining (proportion equal to 1), covariance and vehicle occupancy will increase to 900 and 2.25, respectively. This testifies the prevalence of covariance between OD demands when trip chaining becomes popular.

1.2.2 Contributions

Different from the conventional traffic count location optimization models, the main contribution of this paper is to optimize the traffic count locations so as to estimate **both the mean and covariance** OD demands with minimization of the overall estimation error (in terms of the relative errors of both the mean and covariance of OD demands).

In general, the main contributions of this paper could be viewed from two aspects: theoretical and methodological developments.

For the **theoretical** development, a new concept is introduced, together with new model formulation, to explicitly capture the effects of covariance between OD flows on determination of traffic count locations for OD demand estimation from traffic counts. The mathematical properties of the new model are examined and discussed in the paper. As shown in the numerical results, the conventional model is indeed a special case of the newly proposed model. With the new model proposed in the paper, the traffic count location problem can be generalized. Therefore, the effects of covariance between OD flows can be incorporated to optimize the traffic count locations for simultaneous estimation of mean and covariance of OD demands. Stochastic link choice proportions are updated using an adapted traffic flow simulator to consider the effects of traffic congestion and stochastic OD demands.

As for the **methodological** development, we have extended the proposed model by considering covariance among OD pairs and different magnitudes of travel demands by OD pair. The metaheuristic solution algorithm has also been improved for solving the bi-objective optimization problem.

The rest of this paper is organized as follow. The notations and problem statement are given in Section 2. In Section 3, given the model assumptions, the formulation and properties of WMPREM and WMPREC are presented and discussed. With introduction of the new concept of WMPREM and WMPREC, the model proposed in this paper is shown in Section 4 for optimizing traffic count locations. It follows with the solution algorithm in Section 5 for solving the single-objective and bi-objective traffic count location problems. In Section 6, numerical examples are used to illustrate the key findings of the proposed model and solution algorithm. Conclusions and further studies are finally given in Section 7.

2. Notations and problem statement

We first define below all the variables and vectors that have been adopted in this paper.

Sets

$\mathbf{A}$ = set of links in the traffic network

$\tilde{\mathbf{A}}$ = set of links with traffic sensors

$\mathbf{W}$ = set of OD pairs

Size of sets

m = number of links in the traffic network

$\tilde{m}$ = number of links with traffic sensors

n = number of OD pairs

Measurements

h = total number of days

Subscripts

a, b = subscript for link

w = subscript for OD pair

l = day of counting sample

Parameters

q_w^* = "true" mean OD demands

q_w^{prior} = the prior mean OD demands

$p_{a,w}$ = choice proportion of the traffic flow on link $a \in \mathbf{A}$ of OD pair w

$\delta_{w,a}$ = entry of the OD-link incidence matrix Δ. If the trips on OD pair w pass through link a, $\delta_{w,a} = 1$, otherwise $\delta_{w,a} = 0$.

α = weighting parameter on the proposed WMPREC

Dependent variables

Q_w = "estimator" of OD demands on OD pair w

q_w = estimated mean OD demands

$\sigma_{w,w'}^q$ = estimated covariance between different OD pairs w and w'

λ_w^{mean} = relative deviation of the estimated mean OD demands from the "true" one for OD pair $w \in \mathbf{W}$

$\lambda_{w,w'}^{\mathrm{cov}}$ = relative deviation of the estimated covariance variable of OD demands from the "true" one for different OD pairs $w, w' \in \mathbf{W}$

ρ_{w_0} = weight of traffic flows in OD pair w_0.

ρ_{w_0,w'_0} = weight of covariance between traffic flows in OD pair w_0 and OD pair w_0'.

V_a = random traffic flow on observed link $a \in \tilde{\mathbf{A}}$

$v_a^{(l)}$ = observed traffic flow on link $a \in \tilde{\mathbf{A}}$ during the peak hourly period on day l

v_a = mean of the observed traffic flow on link $a \in \tilde{\mathbf{A}}$

$\sigma_{a,b}^v$ = covariance between observed traffic flows on link a and link b

$r_{w,w'}$ = coefficient of correlation of travel demands between OD pair w and w'

Decision variables

z_a = integer variable, $z_a = 0$ or 1, $z_a = 1$ if a traffic sensor is located on link $a \in \mathbf{A}$, $z_a = 0$ otherwise

Vectors

$\mathbf{z}$ = traffic count locations

$\mathbf{v}^{(l)}$ = observed link flows

$\mathbf{v}$ = sample mean of traffic flows on observed links

Σ^{v} = covariance matrix of traffic flows on observed links

$\mathbf{Q}$ = "estimator" matrix of OD demands

$\mathbf{q}$ = estimated mean matrix of OD demands

$\mathbf{q}^{\mathbf{prior}}$ = prior mean matrix of OD demands

Σ^{q} = covariance matrix of estimated OD demands

$\mathbf{q}^*$ = "true" mean OD demand vector

Σ^{q^*} = covariance matrix of "true" OD demands

P = link choice proportion matrix

$\tilde{\mathrm{P}}$ = sub-matrix of link choice proportion matrix whose links are equipped with traffic sensors

$\boldsymbol{\lambda}^{mean}$ = relative deviation of the estimated mean OD demands matrix

$\boldsymbol{\lambda}^{\mathrm{cov}}$ = relative deviation of the estimated covariance matrix of OD demands

In this paper, there are mainly two stages for modeling the traffic count location problem for stochastic OD demand estimation: (i) the traffic count location stage, and (ii) the stochastic OD demand estimation stage as shown in Fig. 2. The connections between these two stages relate to the observed link flows and the values of WMPREM and WMPREC. The observed link flows based on traffic count locations from the first stage are the inputs of the second stage. Conversely, the outputs of the second stage, WMPREM and WMPREC, calculated from resultant OD demand estimates, are the inputs of the first stage.

The first stage model is to generate the traffic count locations based on prior OD demands, where some existing traffic count location rules are mathematically treated as constraints. However, these traffic count locations may not be the optimal ones concerning the accuracy of OD demand estimation (Larsson et al., 2010). Therefore, the criteria WMPREM and proposed WMPREC for measuring OD demand estimation accuracy are incorporated to optimize the traffic count locations.

For validation purpose, we assume that mean and covariance of observed link flows can be obtained from the "true" stochastic OD demands which are acquired by the adapted traffic flow simulator (Lam and Xu, 1999). It should be noted that only those links equipped with traffic sensors can provide observed link flows.

At the second stage, stochastic OD demands are estimated using Bayes method based on the observed stochastic link flows obtained from the first stage, together with prior OD demands. With the resultant stochastic OD demand estimates, the proposed WMPREM and WMPREC for each traffic count location scheme can be calculated to examine the accuracy of estimated results on mean and covariance of OD demands. By comparing the values of resultant WMPREM and WMPREC among the traffic count location schemes using the Genetic Algorithm (GA), the optimal traffic count location scheme with the minimum OD demand estimation errors can be selected, which is the output of our model.

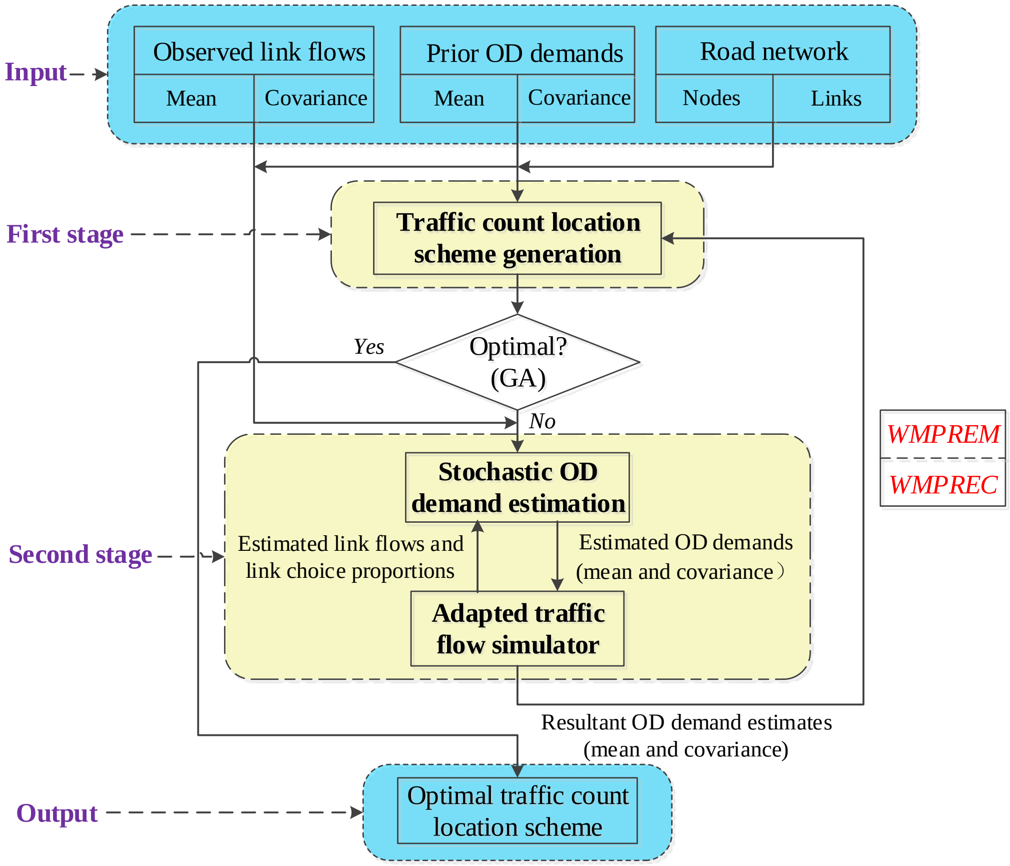


Fig. 2 The relationship between traffic count location and OD demand estimation problems

3. Criteria for measuring OD demand estimation accuracy

3.1 Model assumptions

A1. It is assumed that all the observed link traffic counts are error-free (Yang et al., 1991; Yang and Zhou, 1998).
A2. The covariance of travel demands between any two arbitrary OD pairs is positive. Specifically, it is assumed that all the entries in the OD demand covariance matrix are positive.
It should be pointed out that the covariance of travel demands between any two arbitrary OD pairs could be negative theoretically. However, it is probably hard to observe the negative OD covariance in reality. Taking the trip chaining as shown in Fig. 1 for illustration, once traveler A chooses Destination 6 through node 2 instead of choosing the route 3-4-5-6, the covariance of travel demands between OD (3-2) and OD (2-6) should then be positive. Nonetheless, this assumption deserves to be relaxed in further study to propose a more generalized model.
A3. It is assumed that there is only one traffic sensor allocated at a link.

3.2 Relationship between OD demands and observed day-to-day link flows

The traffic flows on a link are observed during the same peak hourly period (say 8:00 am - 9:00 am) for a typical weekday l over a given number of days h in the year concerned. The link with (without) a traffic sensor is called "observed link" ("unobserved link") throughout this paper. Because of the daily fluctuation in travel demand over the year, the link flows should be stochastic instead of deterministic. The sample mean matrix of hourly traffic flows $\mathbf{v}$ on observed links over the year can be calculated as:

$$\mathbf{v} = (\ldots, v_a, \ldots)^T = \frac{1}{h}\sum_{l=1}^{h}\mathbf{v}^{(l)} \tag{2}$$

The sample covariance matrix of traffic flows on observed links can be calculated as:

$$\Sigma^{\mathbf{v}} = \{\sigma_{a,b}^{v}\}_{\tilde{m}\times\tilde{m}} = \frac{1}{h-1}\sum_{l=1}^{h}\{(\mathbf{v}^{(l)} - \mathbf{v})(\mathbf{v}^{(l)} - \mathbf{v})^T\} \tag{3}$$

Then, the mean of observed link flow can be obtained by the following Equation (4):

$$v_a = E[V_a] = E\left[\sum_{w\in\mathbf{W}} p_{a,w}Q_w\right] = \sum_{w\in\mathbf{W}} p_{a,w}q_w \quad \forall a \in \tilde{\mathbf{A}} \tag{4}$$

According to A2, the covariance between V_a and V_b can be deduced as:

$$\begin{aligned}
\sigma_{a,b}^{v} &= \mathrm{cov}[V_a, V_b] \\
&= \mathrm{cov}[\sum_{w\in\mathbf{W}} p_{a,w}Q_w, \sum_{w'\in\mathbf{W}} p_{b,w'}Q_{w'}] \\
&= \sum_{w\in\mathbf{W}}\sum_{w'\in\mathbf{W}} p_{a,w}p_{b,w'}\,\mathrm{cov}[Q_w, Q_{w'}] \\
&= \sum_{w\in\mathbf{W}}\sum_{w\in\mathbf{W}'} p_{a,w}p_{b,w'}\sigma_{w,w'}^{q} \quad \forall a,b \in \tilde{\mathbf{A}}
\end{aligned} \tag{5}$$

Then Equations (4) and (5) can be rewritten as the following matrix:

$$\mathbf{v} = \tilde{\mathrm{P}}\mathbf{q} \tag{6}$$

And

$$\Sigma^{\mathbf{v}} = \tilde{\mathrm{P}}\Sigma^{q}\tilde{\mathrm{P}}^{T} \tag{7}$$

3.3 Formulations of WMPREM and WMPREC

Yang et al. (1991) proposed the maximum possible relative error (MPRE) to measure the reliability of the estimated OD demands. In this paper, we modify the MPRE as WMPREM. According to assumption A1, the "true" and estimated mean OD demand must satisfy the following relationships.

$$\sum_{w} p_{a,w}q_w = v_a \quad \forall a \in \tilde{\mathbf{A}} \tag{8}$$

$$\sum_{w} p_{a,w}q_w^{*} = v_a \quad \forall a \in \tilde{\mathbf{A}} \tag{9}$$

Subtracting (8) from (9), it follows that

$$\sum_{w} p_{a,w}(q_w^{*} - q_w) = 0 \quad \forall a \in \tilde{\mathbf{A}} \tag{10}$$

Denote $\lambda_w^{mean} = (q_w^{*} - q_w)/q_w$ as the relative deviation of the estimated mean OD demands from the "true" one for OD pair $w \in \mathbf{W}$. It follows from $q_w^* \geq 0$ and $q_w \geq 0$ that $\lambda_w^{mean} \geq -1$. Substituting λ_w^{mean} into Equation (10), it follows that

$$\sum_w p_{a,w} q_w \lambda_w^{mean} = 0 \quad \forall a \in \tilde{\mathbf{A}} \tag{11}$$

Denote ρ_{w_0} as the weight of traffic flows in OD pair w_0.

$$\rho_{w_0} = q_{w_0} / \sum_w q_w \tag{12}$$

Define the average relative deviation of the mean OD demand as

$$G(\boldsymbol{\lambda}^{mean}) = \sqrt{\varphi(\boldsymbol{\lambda}^{mean}) / n} \tag{13}$$

Where

$$\boldsymbol{\lambda}^{mean} = (\ldots, \lambda_w^{mean}, \ldots)^T \tag{14}$$

$$\boldsymbol{\rho}_w = (\ldots, \rho_{w_0}, \ldots)^T \tag{15}$$

$$\varphi(\boldsymbol{\lambda}^{mean}) = \sum_{w \in \mathbf{W}} \rho_w (\lambda_w^{mean})^2 \tag{16}$$

$G(\boldsymbol{\lambda}^{mean})$ is a measure of the estimation error of the mean OD demand. Obviously, the smaller $G(\boldsymbol{\lambda}^{mean})$, the higher the accuracy of the estimation. Therefore, the WMPREM can be defined by the following maximization problem.

$$\text{WMPREM}(\mathbf{z}) = \max_{\lambda^{mean}} G(\boldsymbol{\lambda}^{mean}) \tag{17a}$$

Subject to

$$\sum_w p_{a,w} q_w \lambda_w^{mean} = 0 \quad \forall a \in \tilde{\mathbf{A}} \tag{17b}$$

$$\lambda_w^{mean} \geq -1 \quad \forall w \in \mathbf{W} \tag{17c}$$

Similar to the definition of WMPREM, WMPREC can be defined as below. The covariance matrices of "true" and estimated OD demands must satisfy the following relationships.

$$\sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} p_{a,w} p_{b,w'} \sigma_{w,w'}^q = \sigma_{a,b}^v \quad \forall a, b \in \tilde{\mathbf{A}} \tag{18}$$

$$\sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} p_{a,w} p_{b,w'} \sigma_{w,w'}^{q*} = \sigma_{a,b}^v \quad \forall a, b \in \tilde{\mathbf{A}} \tag{19}$$

Subtracting (18) from (19), it follows that

$$\sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} p_{a,w} p_{b,w'} (\sigma_{w,w'}^{q*} - \sigma_{w,w'}^q) = 0 \quad \forall a, b \in \tilde{\mathbf{A}} \tag{20}$$

Similar to the definition of λ_w^{mean}, let $\lambda_{w,w'}^{\text{cov}} = (\sigma_{w,w'}^{q*} - \sigma_{w,w'}^q) / \sigma_{w,w'}^q$ denote the relative deviation of the estimated covariance matrix of OD demands from the "true" one between OD pairs *w* and *w'*. According to assumption A2, $\sigma_{w,w'}^{q*}$ and $\sigma_{w,w'}^q$ have the same sign, i.e. $\sigma_{w,w'}^{q*} / \sigma_{w,w'}^q \geq 0$, it follows that

$$\lambda_{w,w'}^{\text{cov}} = \frac{\sigma_{w,w'}^{q*}}{\sigma_{w,w'}^q} - 1 \geq -1 \quad \forall w, w' \in \mathbf{W} \tag{21}$$

Substituting $\lambda_{w,w'}^{\text{cov}} = (\sigma_{w,w'}^{q*} - \sigma_{w,w'}^q) / \sigma_{w,w'}^q$ into Equation (20), it follows that

$$\sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} p_{a,w} p_{b,w'} \sigma_{w,w'}^q \lambda_{w,w'}^{\text{cov}} = 0 \quad \forall a, b \in \tilde{\mathbf{A}} \tag{22}$$

Denote ρ_{w_0,w'_0} as the weight of covariance between traffic flows in OD pair w_0 and OD pair w_0'.

$$\rho_{w_0,w'_0} = \sigma^q_{w_0,w'_0} / \sum_{w,w' \in \mathbf{W}} \sigma^q_{w,w'} \tag{23}$$

Define the average relative deviation of the OD demand covariance matrix as

$$H(\boldsymbol{\lambda}^{\text{cov}}) = \sqrt{\psi(\boldsymbol{\lambda}^{\text{cov}}) / n^2} \tag{24}$$

Where

$$\boldsymbol{\lambda}^{\text{cov}} = \{\lambda^{\text{cov}}_{w,w'}\}_{n \times n} \tag{25}$$

$$\psi(\boldsymbol{\lambda}^{\text{cov}}) = \sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} \rho_{w,w'} (\lambda^{\text{cov}}_{w,w'})^2 \tag{26}$$

$H(\boldsymbol{\lambda}^{\text{cov}})$ is a measure of the estimation error of the OD demand covariance matrix. Obviously, the smaller $H(\boldsymbol{\lambda}^{\text{cov}})$, the higher the accuracy of the estimation. Therefore, the WMPREC can be defined by the following maximization problem, which is very similar to the definition of WMPREM.

$$\text{WMPREC}(z) = \max_{\lambda^{\text{cov}}} H(\boldsymbol{\lambda}^{\text{cov}}) \tag{27a}$$

Subject to

$$\sum_{w \in \mathbf{W}} \sum_{w' \in \mathbf{W}} p_{a,w} p_{b,w'} \sigma^q_{w,w'} \lambda^{\text{cov}}_{w,w'} = 0 \quad \forall a, b \in \tilde{\mathbf{A}} \tag{27b}$$

$$\lambda^{\text{cov}}_{w,w'} \geq -1 \quad \forall w, w' \in \mathbf{W} \tag{27c}$$

In Equations (17b) and (27b), both the mean OD demand q_w ($\forall w \in \mathbf{W}$) and covariance variable of OD demand $\sigma^q_{w,w'}$ ($\forall w, w' \in \mathbf{W}$) can be estimated by traditional techniques, such as a weighted least squares method proposed by Shao et al. (2014) or the entropy maximizing method proposed by Van Zuylen and Willumsen (1980). In this paper, the Bayes method is used to estimate the mean and covariance of OD demands because of the consideration of variance and covariance of OD demands.

3.4 Properties of WMPREM and WMPREC

As known in the literature, if the number of traffic sensors is less than the number of OD pairs, the mean OD demand cannot be uniquely identified. This property indicates that if the number of traffic sensors is insufficient. The traffic count location scheme needs to be optimized to improve the estimation error. Such a property needs to be extended for the case of OD demand covariance estimation. To address this issue, the following property explains the relationship between the number of traffic sensors and the uniqueness of the estimated OD demand covariance matrix.

Property 1: If $\tilde{\text{P}}$ is a matrix with full column rank, i.e., $rank(\tilde{\text{P}}) =$ the number of column of $\tilde{\text{P}}$, $\Sigma^{\mathbf{q}}$ must be uniquely identified according to Equation (7).

Remark: It should be noted that the number of columns in $\tilde{\text{P}}$ is equal to the number of traffic sensors. If the number of traffic sensors is less than the number of OD pairs, $\tilde{\text{P}}$ must not be a matrix with full column rank. In other words, only under the condition that the number of traffic sensors is less than the number of OD pairs, the solution for the estimation of OD demand covariance matrix is not unique and should be optimized. Property 1 is consistent with the relationship between the number of traffic sensors and uniqueness of estimated mean OD demand.

In addition, the definition of WMPREC is different from that of WMPREM. The key difference is that WMPREC relates to a pair of traffic count locations ($a, b \in \tilde{\mathbf{A}}$ in Equation (27b)), while WMPREM only depends on a single traffic count location ($a \in \tilde{\mathbf{A}}$ in Equation (17b)). Even if the formulations are different, the mathematical properties are the same under some assumptions.

Before the investigation of the mathematical properties for WMPREC, the definition of OD Covering Rule is introduced as follow.

OD Covering Rule: the traffic sensors on the road network must be located in order to ensure that the traffic flows (or vehicular trips) between each OD pair can be observed.

The OD covering rule was proposed by Toi (1986). Its mathematical formulation with a clear explanation was presented by Yang et al. (1991). It is a common rule widely used for the traffic count location problem in the literature but is extended below in this paper for consideration of the covariance between OD flows.

Property 2: The WMPREM ($G(\boldsymbol{\lambda}^{mean})$) and WMPREC ($H(\boldsymbol{\lambda}^{\mathrm{cov}})$) are both finite under the following two conditions

(i) the OD Covering Rule is satisfied (i.e., the traffic flows between any OD pair are observed by at least one traffic sensor).

(ii) $\sigma_{w,w'} = r_{w,w'}\sigma_w\sigma_{w'}$ and $\sigma_w = c_w q_w$, where $r_{w,w'}$ is the coefficient of correlation between the travel demands of OD pairs *w* and *w'* and c_w is the coefficient of variation of OD pair *w*. In this condition, the covariance between the travel demands of any two OD pairs can be either **positive or negative**.

In other words, the WMPREC and WMPREM are both bounded provided that the OD covering rule and the linear relationship between the mean and covariance of the OD demand are both satisfied. However, if the condition (ii) does not hold, we need the following property to guarantee that WMPREC is finite.

Property 3: The WMPREM ($G(\boldsymbol{\lambda}^{mean})$) and WMPREC ($H(\boldsymbol{\lambda}^{\mathrm{cov}})$) are both finite under the following two conditions

(i) the OD Covering Rule is satisfied (i.e., the traffic flows between any OD pair are observed by at least one traffic sensor).

(ii) the covariance of travel demands between any two arbitrary OD pairs is **positive**. Specifically, it is assumed that all the entries in the OD demand covariance matrix are positive.

As discussed above, either the condition (ii) in Property 2 or the one in Property 3 is satisfied, the WMPREM and WMPREC are bounded. Both Properties 2 and 3 aim to ensure finite WMPREM and WMPREC so that traffic count locations can be optimized in the following models. The proofs of the above three properties can refer to the Appendix.

4. Traffic count location optimization and stochastic OD demand estimation

In this section, we firstly discuss the equivalent optimization model for optimizing the traffic count locations with consideration of both the WMPREM and WMPREC at the first stage. At the second stage, the extension of the formulation based on Bayes method is presented to estimate both the mean and covariance matrix of OD demands from traffic counts and stochastic link choice proportions are updated by an adapted traffic flow simulator.

4.1 Model formulation for traffic count location optimization

The aim of this paper is to enhance the reliability of the estimated OD demands by locating traffic counts in the road network. As the WMPREM and WMPREC can be regarded as two measures of the estimation errors for the estimated mean and covariance OD demands, the smaller the WMPREM and WMPREC are, the higher the accuracy of the estimation is. Then, minimizing WMPREM and WMPREC can enhance the reliabilities of the estimated mean and covariance matrices of OD demands. Thus, a bi-objective model is proposed for improving the estimation reliability of the OD mean and covariance matrices. As discussed above, two objective functions can be used. The first objective function is expressed as

$$\min O_1 = \mathrm{WMPREM}(\mathbf{z}) \tag{28a}$$

The second objective function is expressed as

$$\min O_2 = \mathrm{WMPREC}(\mathbf{z}) \tag{28b}$$

Subject to

$$z_a = 0 \text{ or } 1 \quad \forall a \in \mathbf{A} \tag{28c}$$

It is well known for the traffic count location problem that the more traffic sensors are used in the road network, the smaller the estimation errors are. Thus, an obvious solution for programming (28) is $z_a = 1$ for all $a \in \mathbf{A}$. This solution may not be feasible in practice. In reality, there is usually a budget constraint for locating traffic counts in the road network. Thus, the number of traffic sensors should be within a given threshold based on the budget constraint, which can be expressed as below:

$$\sum_{a \in \mathbf{A}} c_a z_a \leq B \tag{29}$$

Where c_a is the cost for installing and maintaining one traffic sensor on link *a*, and *B* is the total budget.

However, the number of possible traffic sensors under consideration could be very large if a traffic network contains a large number of road links. For example, in a network with 100 links, there are totally $2^{100} - 1$ possible schemes. If the number of traffic sensors used for OD estimation is fixed, the number of schemes is still too large for any methodology to cope with in practice. One approach to solve this problem is to use some traffic count location constraints to remove the redundant schemes so as to reduce the size of possible schemes.

Yang and Zhou (1998) have proposed four traffic count location rules: **OD covering rule**, maximal flow fraction rule, maximal flow intercepting rule and **link independence rule**. OD covering rule is considered as a fundamental rule, so in this paper. Note that it should always be satisfied and be treated as a constraint. The link independence rule should also be satisfied because it could exclude the redundant links. So, in this paper, we also transform the link independence rule into an equivalent constraint for solving the optimization problem concerned. The following two Equations (30) and (31) are the mathematical expressions of OD covering rule and link independence rule, respectively.

$$\sum_{a=1}^{m} \delta_{w,a} z_a \geq 1 \quad \forall w \in \mathbf{W} \tag{30}$$

$$rank(\tilde{\mathrm{P}}) = \tilde{m} \tag{31}$$

For every moderate road network, the above two constraints could only remove a small portion of "not too good" schemes, and there still be a large number of possible schemes for consideration. According to the maximal flow intercepting rule proposed by Yang and Zhou (1998), traffic counts should be located on links so that the observed flows are as many as possible. In this paper, because we also consider the fluctuation and covariance of link flows, a variant maximal flow intercepting rule is proposed as a **maximal probability of flow intercepting rule**. That is, traffic counts should be located on road links so that the probability of the observed link flows larger than a given value (e.g. mean link flow) is as much as possible. The proposed probability-based rule could be mathematically expressed as a constraint as follow:

$$\Pr(\mathbf{V} > v) > p^0 \text{ (the value of } p^0 \text{ varies with the number of traffic sensors)} \quad (32)$$

Finally, the mathematical model for optimization of traffic count locations can be formulated as the following constrained bi-objective problem:

$$\min \begin{cases} O_1 = \text{WMPREM}(\mathbf{z}) \\ O_2 = \text{WMPREC}(\mathbf{z}) \end{cases} \quad (33a)$$

Subject to

(28c) and (29) -(32) (33b)

The optimization problem (33) aims to find the traffic count locations that can minimize the estimation errors within a given budget constraint. To better understand the proposed model, a proposition should be noticed as follow:

Proposition 1: the value of WMPREM (or WMPREC) will be positive if the number of traffic sensors is less than the number of OD pairs; the value of WMPREM (or WMPREC) will be zero if the number of traffic sensors is no less than the number of OD pairs.

This could be proved directly from that when the number of traffic sensors is no less than the number of OD pairs, the number of equations is equal to the number of unknown variables according to the conservation law (Equation (6)). Thus, q will be a matrix with full column rank, then the OD flows could be calculated accurately.

4.2 Estimation of stochastic OD demands and update of stochastic link choice proportions

In this study, statistical methods such as the maximum likelihood method and Bayes method may be more suitable for this problem than some optimization methods, like Generalized Least Square method and Entropy Maximization method. The statistical methods may perform better because these methods can explicitly consider the variation of OD demands in our case. As such, after determining the traffic count locations at the first stage of the proposed model, the second-stage problem is to estimate the mean and covariance of OD demands using Bayes method and then update stochastic link choice proportions using an adapted traffic flow simulator based on the estimated stochastic OD demands.

Due to that the OD demands will change during the iteration, the stochastic link choice proportions and stochastic link flows are also updated during determining traffic count locations as shown in Fig. 2. This information obtained from the adapted traffic flow simulator at the second stage should be consistent with the observed ones. Thus, the stochastic link flows and stochastic link choice proportions are updated iteratively until the difference between the estimated link flows from the adapted traffic flow simulator and observed link flows from traffic counts less than the predetermined tolerance.

The "estimated link flows" including their mean and covariance are obtained from the adapted traffic flow simulator based on the estimated stochastic OD demands at the second stage. In this paper, the "true" OD demands are assumed so that the mean and covariance of "observed link flows" can be obtained by assigning the "true" mean and covariance of OD demands using the adapted traffic flow simulator. The "true" OD demands only serve as a reference for validation in the numerical examples.

Based on the stochastic link flows observed from traffic counts of which the locations have been determined at the first stage, the Bayes method for estimating the stochastic OD demands is formulated as follows:

Suppose that observed link flows follow a multivariate normal distribution, $V \mid Q \sim MVN(PQ, \Sigma^v)$, where V is an $m \times 1$ vector of observed link flow, Q is an $n \times 1$ parameter vector of estimated OD demands, P is an $m \times n$ given matrix of link choice proportion, and Σ^v is an $m \times m$ observed covariance matrix of link flow. In addition, suppose the prior distribution of OD demand is also multivariate normal, $Q \sim MVN(Q^{prior}, \Sigma_q^{prior})$, where Q^{prior} is an $n \times 1$ parameter vector of prior OD demands, Σ_q^{prior} is an $n \times n$ covariance matrix of prior OD demands. Please note that both Q^{prior} and Σ_q^{prior} are known. From Bayes method (Carlin et al., 2000), we can get the marginal distribution of observed link flows V as follows:

$$V \sim MVN(PQ^{prior}, \Sigma^{v} + P\Sigma_q^{prior} P^T) \tag{34}$$

The posterior distribution of OD demands Q can then be obtained as follows:

$$Q \mid V \sim MVN(Dd, D) \tag{35a}$$

Where

$$D^{-1} = P^T \Sigma^{v-1} P + (\Sigma_q^{prior})^{-1} \tag{35b}$$

$$d = P^T \Sigma^{v-1} V + (\Sigma_q^{prior})^{-1} Q^{prior} \tag{35c}$$

Therefore, the mean OD demands can be estimated.

$$\mathbf{q} = E(Q \mid V) = Dd \tag{36}$$

The posterior covariance matrix of OD demand can be expressed as follows:

$$\Sigma^{\mathbf{q}} = \mathrm{cov}(Q \mid V) = D \tag{37}$$

Based on the estimation of mean and covariance OD demands, stochastic link choice proportions together with the stochastic link flows are then updated by an adapted traffic flow simulator. The initial method can be referred to Lam and Xu (1999). In this study, based on the estimation of mean and covariance of OD demands from the Bayes method, we assume that OD demands follow a multivariate normal distribution. By sampling the OD demand from the overall population, SUE assignment is used to obtain mean and covariance of link flows and link choice proportions. Stochastic OD demands will be updated according to the proposed Bayes method.

5. Solution algorithm

In this section, the firefly algorithm will be adapted to solve the proposed bi-objective optimization problem.

5.1 Solution formulation

For each traffic count location scheme, we can solve the problem (35) to estimate the mean OD demands $\mathbf{q}$ and the covariance matrix of OD demands $\Sigma^{\mathbf{q}}$. Similar to Equations (6) and (7), Equations (17b) and (27b) can be rewritten in the following matrix form:

$$\tilde{\mathrm{P}} \mathbf{q}^d \boldsymbol{\lambda}^{mean} = \mathbf{0} \tag{38}$$

$$\tilde{\mathrm{P}} \Sigma^{\mathbf{q}} \boldsymbol{\lambda}^{\mathrm{cov}} \tilde{\mathrm{P}}^T = \mathbf{0} \tag{39}$$

Where $\mathbf{q}^d$ is a diagonal matrix whose diagonal elements are q_w.

$$\mathbf{q}^d = \begin{bmatrix} q_1 & 0 & \dots & 0 \\ 0 & q_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & q_n \end{bmatrix}$$

Then the matrix form of the optimization problem (17) for finding WMPREM is shown as below:

$$\text{WMPREM}(\mathbf{z}) = \max_{\lambda^{mean}} G(\boldsymbol{\lambda}^{mean}) \tag{40a}$$

Subject to

$$\tilde{\text{P}}\mathbf{q}^d\boldsymbol{\lambda}^{mean} = \mathbf{0} \tag{40b}$$

$$\boldsymbol{\lambda}^{mean} \geq -1 \tag{40c}$$

To facilitate efficient solutions, the matrix equation constraint (39) can be transformed into a linear equation constraint as follow.

$$\text{M}^{\lambda}\text{vec}(\boldsymbol{\lambda}^{\text{cov}}) = \mathbf{0} \tag{41}$$

The method to obtain matrix M^{λ} is shown in Appendix A. Subsequently, the optimization problem (27) should be rewritten as below:

$$\text{WMPREC}(\mathbf{z}) = \max_{\lambda^{\text{cov}}} H(\text{vec}(\boldsymbol{\lambda}^{\text{cov}})) \tag{42a}$$

Subject to

$$\text{M}^{\lambda}\text{vec}(\boldsymbol{\lambda}^{\text{cov}}) = \mathbf{0} \tag{42b}$$

$$\text{vec}(\boldsymbol{\lambda}^{\text{cov}}) \geq -1 \tag{42c}$$

Then, the bi-objective model could be formulated as below:

$$\min \begin{cases} O_1 = \text{WMPREM}(\mathbf{z}) \\ O_2 = \text{WMPREC}(\mathbf{z}) \end{cases} \tag{43a}$$

Subject to

$$\begin{bmatrix} \mathbf{q}^d & \\ & \text{M}^{\lambda} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}^{mean} \\ \text{vec}(\boldsymbol{\lambda}^{\text{cov}}) \end{bmatrix} = \mathbf{0} \tag{43b}$$

$$\begin{bmatrix} \boldsymbol{\lambda}^{mean} \\ \text{vec}(\boldsymbol{\lambda}^{\text{cov}}) \end{bmatrix} \geq -1 \tag{43c}$$

To solve this constrained bi-objective optimization problem, two main approaches are used: weighted-sum approach and bi-objective approach.

The weighted-sum approach converts the bi-objective problem into a single objective problem as Equation (44) by varying the weights of the two objectives. However, this weighted-sum approach requires a good background knowledge of the problem so as to determine the weighting parameter α and then to obtain a reliable solution.

$$\text{WMPRE}(\mathbf{z}) = \min_{\mathbf{z}}\{(1-\alpha)\text{WMPREM}(\mathbf{z}) + \alpha\text{WMPREC}(\mathbf{z})\} \tag{44}$$

More generally, the bi-objective approach is adopted to acquire the Pareto front of this bi-objective problem without the need of determining the weighting parameter α . A set of non-dominated solutions called Pareto optimal solutions will be generated, which represents the relationship between the two objectives. Hence, no absolute unique solution could be obtained from the bi-objective approach.

To better understand the pattern of solutions from these two approaches, the results obtained from the weighted-sum approach and bi-objective approach will be compared below. The proposed traffic count location optimization is a non-convex problem with binary decision variables that represent the traffic count locations. This problem is NP-hard such that no global optimization algorithm can be used to solve it. For the two approaches concerned, the main difference between the two solution algorithms is the fitness function described in the following subsection.

5.2 Firefly algorithm

Firefly algorithm is a novel and powerful nature-inspired algorithm inspired by social behavior of fireflies (Yang, 2008). The original firefly algorithm is to optimize continues problems. Some scholars have further developed the algorithm to apply it for different areas such as mixed integer programming (Jati and Suyanto, 2011; Sayadi et al., 2010) and multi-output support vector regression (Xiong et al., 2014). This study has adapted the firefly algorithm to solve the traffic count location problem that is regarded as mixed integer programming with binary decision variables and constraints. The numerical example as will be shown in the following section demonstrates its efficiency with comparison to the classical genetic algorithm, which is widely used in literature (see Yin, 2000).

5.2.1 Firefly representation

It is very convenient to represent a traffic count location scheme by a firefly in the platform of firefly algorithm. **z** is a vector with values 0 and 1 only. The purpose of the proposed FA is to determine the value of **z**. Thus one firefly represents one traffic count location scheme **z** and each binary variable indicates the existence of a traffic sensor on link *a*, i.e., 1 if a traffic sensor is located on link *a* and 0 otherwise.

5.2.2 Fitness function

Weighted-sum approach

Define $I(\mathbf{z}) = \text{WMPRE}(\mathbf{z})$ as the light intensity of fireflies. $I(\mathbf{z})$ is selected as the fitness function, here. The smaller the $I(\mathbf{z})$ is, the more likely the scheme **z** will be selected.

Bi-objective approach

As for our bi-objective problem, the solution is non-dominated (Pareto optimal) if there were no other feasible solutions that could improve one objective without worsening another objective. In one iteration, the non-dominated solutions but not a unique optimal solution will be determined. The population in the next iteration will be generated based on the non-dominated solutions.

5.2.3 Algorithm steps

Step 1 (Initialization)

After determining the number of traffic sensors *l* by considering the budget constraint, generate initial population of fireflies $\mathbf{Z}^0$ randomly, the size of $\mathbf{Z}^0$ is *k*, and the number of variables in each firefly equals to the number of links in the road network. Set the maximum iterations (or the maximum generations) to *T*, and set the iteration number *t* to zero. (Bielli et al., 2002; Nayeem et al., 2014)

Step 2 (Conditions for the termination judgment)

If the number of iterations is larger than the threshold for maximum iteration (if $t > T$), terminate. Otherwise, go to next step. Two different approaches are considered in this paper. If weighted-sum approach is used, go to Step 3a. Otherwise, go to Step 3b.

(Selection operation)

Step 3a: For each firefly $\mathbf{z}^{(i)}$ ($i = 1, 2, ..., k$) of population pool $\mathbf{Z}^{(t-1)}$, estimating the mean and covariance matrix of OD demands by using the Bayes method (shown as formulation (35)), updating stochastic link choice proportions using the adapted traffic flow simulator, and then solving the bi-objective optimization problem (43). After calculating the value of its light intensity $I(\mathbf{z}^{(i)})$ according to the fitness function, rank the fireflies based on light intensity and find the current best. Then go to step 4a.
Step 3b: After estimation of the mean and covariance matrix of OD demands and update of stochastic link choice proportions using the adapted traffic flow simulator for each individual of the population pool, calculate WMPREM and WMPREC according to the proposed criteria (formulations (40) and (42)). By comparing the values of WMPREM and WMPREC among all feasible solutions, the non-dominated solutions will be determined.

(Variation operation)
Step 4a: Vary attractiveness of fireflies $\mathbf{Z}^{(t)}$ according to the distance and light intensity, respectively, update $\mathbf{Z}^{(t)}$ consisted of *k* individuals based on the roulette wheel. Set $t = t + 1$, go back to step 2.
Step 4b: Based on the selected non-dominated solutions, update the population pool according to the distance and light intensity, respectively. Set $t = t + 1$, then go back to Step 2.

6. Numerical examples

In this section, two numerical examples are used to illustrate the applicability of the proposed model and solution algorithm for estimating the mean and covariance of OD demands. Example 1 is a small transportation network to examine: (a) effects of OD demand covariance on the optimal traffic count locations; (b) effects of the number and location of traffic counts on estimation reliability; (c) effects of traffic congestion on the estimation results; and (d) sensitivity of weighting parameter α in the weighted-sum approach. Example 2 employed a medium-size transportation network to demonstrate the applicability of proposed method and convergence of the solution.

6.1 A small transportation network

As shown in Fig. 3, a small transportation network that consists of 7 nodes, 16 links and 12 OD pairs is deployed. The mean and covariance of OD demands during morning peak hours for 300 days were estimated for this Example 1 network.

In this example, the prior mean and covariance of OD demands are set following $\mathbf{q}^{prior} = 0.5\mathbf{q}^*$, $\Sigma_{\mathbf{q}}^{prior} = 0.5^2 \Sigma^{\mathbf{q}^*}$, which are given in Table 4 and Table 6, respectively. For example, the travel demands from origin 2 to destination 6 is numbered OD 2, in which the prior mean OD demands in OD 2 are assumed **240**; the demand from origin 3 to destination 2 is numbered OD 4, in which the prior mean OD demands are assumed **208**. The covariance between OD 2 and OD 4, the maximum covariance among all OD pairs, is **1996.1**.

With respect to the "true" covariance matrix of OD demands, the re-sampling method, depicted as follow, is used (Lo et al., 1996). First, the "true" traffic flows for each OD pair during the morning peak hours during a sequence of days (e.g. 300 days) are generated from a normal distribution with mean q and variance $(0.2q)^2$. We then re-sample the OD demands with sampling fraction 10% to enlarge the sample size. From the enlarged samples, covariance of different OD demands which are considered as the "true" covariance OD demands could then be calculated for the numerical examples in this paper. However, with the recent advancement of automatic vehicle identification (AVI) technologies such as Bluetooth, WiFi, RFID and automatic license plate recognition technologies etc., the sample prior OD mean and covariance matrices can now be estimated based on these AVI data. In addition, the databases of online trip chaining platforms, e.g. Uber and DiDi, are also useful for generating the prior mean and covariance of OD demands.

The parameters of the adapted traffic flow simulator model are set to be the same as that in Lam and Xu (1999). A Monte Carlo based algorithm is used for the adapted traffic flow simulator in this study. It should be noted that to save the computational time for searching the feasible paths by OD pair in our numerical examples, the path choice set for each OD pair is assumed to be known and fixed.

In addition, the initial mean link choice proportion matrix obtained from prior stochastic OD demands using the adapted traffic flow simulator is given in Table 5. Table 4 shows the other network parameters for Example 1 network.

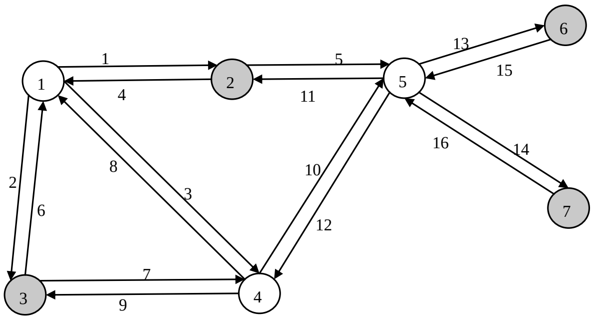


Fig. 3 Example 1 network

Table 4 The network parameters

OD number	Origin-Destination	Routes	Prior mean OD demands
1	2-3	4-2; 4-3-9	168
2	**2-6**	4-3-10-13; 5-13	**240**
3	2-7	4-3-10-14; 5-14	96
4	**3-2**	6-1; 7-8-1; 7-10-11	**208**
5	3-6	6-1-5-13; 6-3-10-13; 7-10-13	223
6	3-7	6-1-5-14; 6-3-10-14; 7-10-14	240
7	6-2	15-11; 15-12-8-1	144
8	6-3	15-12-9; 15-12-8-2	168
9	6-7	15-14	184
10	7-2	16-11; 16-12-8-1	120
11	7-3	16-12-8-2; 16-12-9	136
12	7-6	16-13	208

Table 5 Initial Link choice proportions by OD pair

Link \ OD No.	1	2	3	4	5	6	7	8	9	10	11	12	Link Flow
1				0.9	0.2	0.3	0.4			0.4			485.6
2	0.8							0.4			0.5		350.0
3	0.2	0.2	0.2	0.1	0.4	0.3							358.1
4	1.0	0.2	0.2										284.0
5		0.8	0.8		0.2	0.3							477.0
6				0.9	0.6	0.6							565.9
7				0.1	0.4	0.4							274.1

8				0.1			0.4	0.4		0.4	0.5		343.9
9	0.2							0.6			0.5		240.0
10		0.2	0.2	0.1	0.8	0.75							560.1
11				0.1			0.6			0.6			235.4
12							0.4	1.0		0.4	1.0		511.8
13		1.0			1.0							1.0	840.0
14			1.0			1.0			1.0				650.0
15							1.0	1.0	1.0				620.0
16										1.0	1.0	1.0	580.0

Table 6 The prior covariance matrix of OD demands

OD No.	1	2	3	4	5	6	7	8	9	10	11	12
1	1129.0											
2	1240.2	2304.0										
3	366.6	450.1	368.6									
4	820.6	**1996.1**	354.9	1730.6								
5	831.5	996.1	333.1	648.2	1989.0							
6	1526.3	1877.5	596.7	1461.7	1384.5	2304.0						
7	824.5	954.7	354.9	678.6	709.0	1089.7	829.4					
8	973.4	1151.3	429.8	841.6	960.2	1565.1	758.9	1129.0				
9	1049.1	1287.0	407.9	884.5	1282.3	1745.6	943.0	1077.9	1354.2			
10	490.6	514.0	123.2	325.3	471.1	756.6	274.6	447.7	544.4	576.0		
11	782.3	765.2	322.1	706.7	599.0	1205.9	570.2	769.1	780.8	279.2	739.8	
12	825.2	1047.5	295.6	819.0	726.2	1522.6	618.5	861.9	959.4	397.8	556.9	1730.6

6.1.1 Effects of OD demand covariance on the optimal traffic count locations

In this subsection, the effects of OD demand covariance on the optimal traffic count locations are examined by considering different scenarios on how to use the WMPREC. Different values of weighting parameter α reflect the different levels of importance considered for estimating the mean OD demand and its covariance. Therefore, three scenarios are proposed by setting three values of the α in the weighted-sum approach (e.g., Equation (44)). The three scenarios are list as below:

- $\alpha = 1$ represent the scenario that only WMPREC is considered;
- $\alpha = 0$ represent the scenario that only WMPREM is considered;
- $\alpha = 0.5$ represent the scenario that both WMPREC and WMPREM are considered.

In this example, the network as shown in Fig. 3 is adopted. We use the “real” relative error of mean and covariance of OD demands to compare the estimated mean and covariance of OD demands with the assumed “true” OD demands. The “real” relative errors of mean and covariance OD demand are calculated by Equations (13) and (24), respectively, with the assumption that the “true” mean and covariance OD demands are known (Zhou and List, 2010).

As shown in Tables 7 -9, links traversed by OD pairs with larger covariance should be covered by traffic sensors if WMPREC is considered in the objective function. For instance, as depicted in Table 6 the covariance between OD 3-2 and OD 2-6 is the greatest. If WMPREC is considered in the objective function, more traffic sensors should be located on the set of links {6,1,5,13}. It could be seen from Tables 7 -9 that when the number of traffic sensors is 7, one link (link 13) in the set {6,1,5,13} is covered by a traffic sensor if only WMPREM is considered ($\alpha = 0$); three links (link 1, 5, and 13) in this set are covered by traffic sensors if both WMPREM and WMPREC are considered ($\alpha = 0.5$); and four links in this set are covered by traffic sensors if only WMPREC is considered ($\alpha = 1$) respectively.

In addition, it is noted that if only WMPREM is considered ($\alpha = 0$), WMPRE is **1.71** when the number of traffic sensors is 8. However, if estimation accuracy of both mean and covariance are considered, to achieve comparable estimation accuracy of WMPRE (**1.79**), only 7 traffic sensors are needed on the network. Furthermore, even only 5 traffic sensors are needed when $\alpha = 1$. Therefore, we could conclude that number of traffic sensors required can be reduced when the estimation accuracy of covariance is considered in the objective function, given that the change in overall estimation accuracy is marginal.

As shown in Table 9, when there are 7 traffic sensors, the "Real" relative error of covariance of OD demands (**0.07**) is the least when $\alpha = 1$, as compared to that (e.g., **0.39** and **0.21**) using other weighting parameters. Hence, it can be concluded that the estimation error of the covariance of OD demands could be reduced when WMPREC is adopted. In other words, WMPREC could be more applicable to the situation that both the mean and covariance of OD demands are needed to be estimated.

Table 7 Results of the optimal traffic count location scheme selected in accordance to WMPREM ($\alpha = 0$)

Number of traffic sensors	WMPRE	**WMPREM**	WMPREC	"Real" relative error of mean OD demands	"Real" relative error of covariance of OD demands	The optimal traffic count location scheme selected by WMPREM
5	3.84	3.84	4.21	0.31	0.55	3,10,13,15,16
7	2.59	2.59	**1.82**	0.27	**0.39**	***3,4,7,10,13,14,16***
8	**1.71**	1.71	1.33	0.20	0.31	1,4,7,10,12,13,14,15
11	0.17	0.17	0.33	0.11	0.19	2,3,4,7,8,9,10,11,12,13,15,16

Table 8 Results of the optimal traffic count location scheme selected in accordance to WMPRE ($\alpha = 0.5$)

Number of traffic sensors	**WMPRE**	WMPREM	WMPREC	"Real" relative error of mean OD demands	"Real" relative error of covariance of OD demands	The optimal traffic count location scheme selected by WMPRE
5	2.99	3.92	2.05	0.35	0.29	3,5,10,15,16
7	**1.79**	2.65	**0.92**	0.31	**0.21**	***1,2,4,5,7,13,16***
8	1.36	1.80	0.91	0.24	0.17	1,3,6,9,11,12,13,14
11	0.15	0.21	0.08	0.12	0.09	1,2,3,4,5,7,9,11,12,13,15

Table 9 Results of the optimal traffic count location scheme selected in accordance to WMPREC ($\alpha = 1$)

Number of traffic sensors	WMPRE	WMPREM	**WMPREC**	"Real" relative error of mean OD demands	"Real" relative error of covariance of OD demands	The optimal traffic count location scheme selected by WMPREC
5	**1.78**	4.03	1.78	0.40	0.11	3,5,10,15,16
7	0.84	2.75	**0.84**	0.39	**0.07**	***1,5,6,9,12,13,14***
8	0.77	1.86	0.77	0.25	0.06	2,5,6,9,11,13,14,16
11	0.05	0.23	0.05	0.15	0.02	1,5,6,9,10,11,12,13,14,15,16

6.1.2 Effects of the number and location of traffic counts on estimation reliability

In this subsection, the weighted-sum approach is used to obtain a unique solution, and to investigate the effects of the number and location of traffic counts on estimation reliability. For this sensitivity test, the value of weighting parameter α is set to be 0.5. This implies that in the process of optimizing the traffic count locations, the estimation reliability of the mean and covariance OD demands have equal importance.

As shown in Fig. 4, it could be observed that when the number of traffic sensors increases, both the Pareto optimal solutions for WMPREM and WMPREC decrease, because more traffic sensors deployed contribute to more information acquired. The estimation accuracy for both the mean and covariance of OD demands could then be enhanced.

It is however noted in Fig. 4 that the reduction range of WMPREM is remarkably greater than that of WMPREC when the number of traffic sensors increases. In other words, WMPREM is more sensitive to the number of traffic sensors deployed, compared to WMPREC. For instances, when the number of traffic sensors is 5, WMPREM (**1.85**) is almost 50% greater than WMPREC (**1.16**). In contrast, when the number of traffic sensors increases to 11, WMPREM reduces to **0.08**, while WMPREC only reduces to **0.21**.

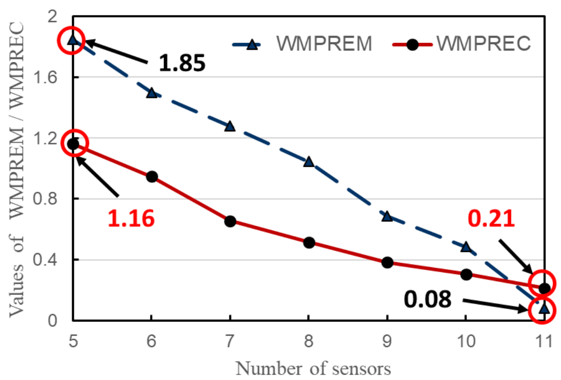


Fig. 4 Effects of number of traffic sensors on WMPREM and WMPREC

Fig. 5 illustrates the ranges of WMPREM and WMPREC for different traffic count locations, stratified into groups of 500 feasible schemes in the descending order of WMPRE for different numbers of traffic sensors (i.e. 7, 9, and 11, respectively).

As shown in Fig. 5, the ranges of both WMPREM and WMPREC vary with both the number and location of traffic counts. When the number of traffic sensors is 7, the values of WMPREM are, in general, much greater than that of WMPREC. In addition, the ranges of both WMPREM and WMPREC are large. When the number of traffic sensors increases, values of WMPREC and WMPREM both reduce in general. Particularly, the reductions of WMPREM value are more remarkable than that of WMPREC. When the number of traffic sensors is 11, it is noted that the values of WMPREM are even less than that of WMPREC in general.

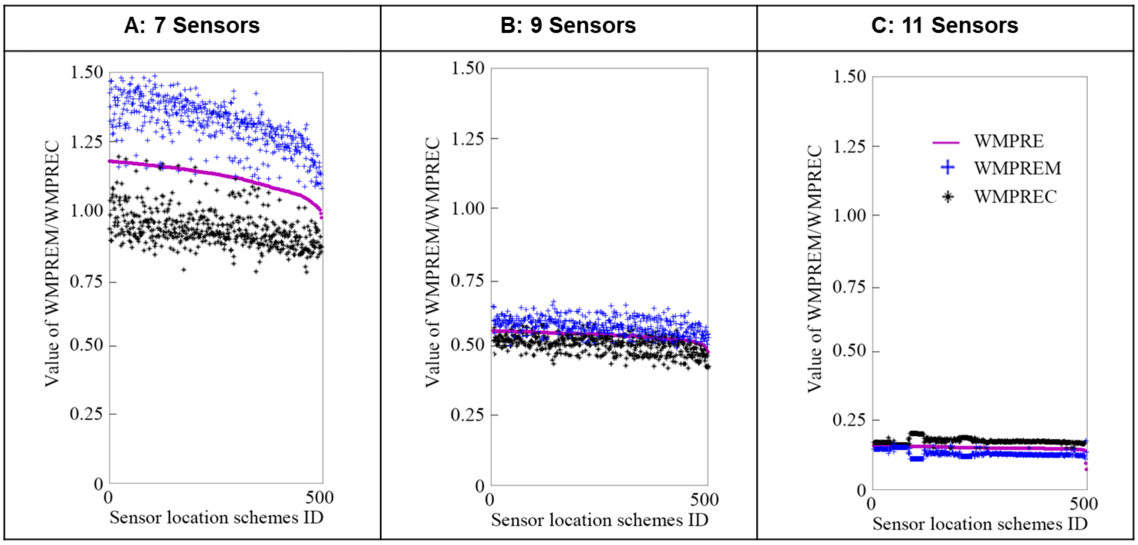


Fig. 5 Effects of location of traffic counts on WMPREM and WMPREC

In addition, the newly proposed criteria (WMPREM and WMPREC) in this paper are intended to weight the conventional criteria by considering different magnitudes of travel demands by OD pair during the morning peak hours. We would like to validate the proposed model by comparing with the previous non-weighted method, MPREM in Yang et al. (1991) as well as the MPREC in this paper.

As shown in Table 10, when the number of traffic sensors in the Example 1 network is greater than or equal to 12 (number of OD pairs), the estimates of both the mean and covariance of OD demands can be absolutely accurate (with zero MPREM and MPREC).. It could be seen from Table 10 that with the same number of traffic sensors, the proposed weighted maximum possible relative errors are much smaller than that based on the non-weighted criterion. For instance, the relative increase for WMPREM is **-91%** when the number of traffic sensors is 11; the Relative increase for WMPREC is **-82%** with 5 traffic sensors. Thus, we could conclude that considering different magnitudes of travel demands by OD pair is effective in reducing the estimation errors for both mean and covariance of OD demands, compared to the original method (Maximum Possible Relative Error, MPRE).

Table 10 Comparison between the methods with and without weight on MPREM or MPREC

Number of traffic sensors	MPREM	WMPREM	Relative increase [a]	MPREC	WMPREC	Relative increase
5	3.38	1.85	-45%	6.50	1.16	**-82%**
7	2.68	1.28	-52%	3.44	0.66	-81%
8	2.02	1.04	-48%	1.92	0.52	-73%
11	0.88	0.08	**-91%**	0.56	0.21	-62%
12,13,14,15,16	0	0	**--**	0	0	--

[a] Relative increase = (WMPRE-MPRE)/MPRE

6.1.3 Effects of traffic congestion on the proposed model

Intuitively, when the growth of number of private cars is faster than that of road capacity, the traffic network is more likely to be congested. Under such circumstances, different OD demands are more likely to be correlated.

To illustrate the importance of considering WMPREC, two scenarios: (i) uncongested condition (Scenario A) and (ii) congested condition (Scenario B) are set out below. Under the uncongested condition, the mean and covariance of OD demands are set to be half of that as given in Table 4 and Table 6. In contrast, under the congested condition, the mean and covariance of "true" OD demands are doubled. To have a unique Pareto optimal solution, and then to compare different results conveniently, the weighted-sum approach is adopted again in this subsection. The weighting parameter α is set to be 0, which implies that only WMPREM is considered in the objective function.

It is noticed from the Table 11 that the estimation error of covariance of OD demand could be large if only the WMPREM is used to determine the traffic count locations for both the uncongested and congested conditions. For instance, under the congested condition, when there are 5 traffic sensors in the Example 1 network, the value of WMPREM is **2.40**, and the value of WMPREC is **6.62**, much larger than that of WMPREM. This phenomenon is also observed when 7 traffic sensors are deployed. Hence, even though the mean OD demands could be accurately estimated, it is difficult to estimate the covariance of OD demands accurately when only the WMPREM is chosen as the criterion. Such phenomenon is much more obvious under the congested condition. It could be seen in Table 11 that when there are only 5 traffic sensors under the congested condition, the resultant WMPREC is **15.13** which is more than double than the result (**6.62**) under the uncongested condition.

Table 11 Results of the model under different traffic conditions

Number of traffic sensors	Uncongested condition (**half** actual OD demands)			Congested condition (**double** actual OD demands)		
	WMPRE	WMPREM	WMPREC	WMPRE	WMPREM	WMPREC
5	2.40	**2.40**	**6.62**	2.28	2.28	**15.13**
7	1.67	1.67	2.52	1.62	1.62	3.18
8	1.38	1.38	1.66	1.32	1.32	2.52
11	0.48	0.48	0.60	0.54	0.54	0.96

Intuitively, the larger the value of WMPREC implies that the covariance between OD flows could not be captured properly by the optimal traffic count location scheme. As for the illustrative example given in subsection 1.2.1, it will be difficult to assess the effects of the carpooling, ridesharing, and other trip chaining strategies in the traffic network on the basis of OD estimation from traffic counts. Therefore, it is shown that WMPREC is essential for stochastic OD demand estimation, particularly under the congested condition.

To testify the performance of the newly proposed index, WMPREC, especially under the congested condition, a sensitivity test has been carried out for both WMPREM and WMPREC in the following subsection. In this comparison, the weighting parameter α is set to be **0.5**. Scenario **A** and **B** stand for the uncongested and congested condition, respectively. Table 12 shows the resultant relative increases in WMPREM and WMPREC, between congested condition and uncongested conditions [i.e. Relative increase = ($WMPRE_B$ – $WMPRE_A$)/ $WMPRE_A$], for different numbers of traffic sensors.

Table 12 Relative increase for WMPREM and WMPREC under congested condition compared to that under uncongested condition

Number of traffic sensors	WMPREM			**WMPREC**		
	Scenario A	Scenario B	Relative increase ($WMPREM_B$ -$WMPREM_A$)/ $WMPREM_A$	Scenario A	Scenario B	Relative increase ($WMPREC_B$ -$WMPREC_A$)/ $WMPREC_A$
5	3.64	4.13	13.5%	4.01	6.88	71.5%
7	2.56	3.26	27.4%	2.72	2.97	9.1%
8	2.04	2.26	11.0%	1.99	1.77	**-10.9%**
11	0.95	0.97	1.9%	0.87	0.88	1.0%

Negative relative increase implies that estimation under congested condition is more accurate than that under uncongested condition. It can be seen from Table 12 that estimation of mean OD demand is less accurate under congested condition (using the Example 1 network), as compared to that under uncongested condition. However, it is possible for the estimation of covariance OD demands to be more accurate when the traffic network becomes congested, as implied by the negative relative increase of WMPREC. For example, when there are 8 traffic sensors, accuracy of OD covariance estimation under congested condition improve by **10.9%**. It could be concluded that WMPREC should be taken into consideration especially when the traffic network is congested. Under congested condition, the observed data might provide more information about the covariance of OD demands as more travelers may use the carpooling and trip chaining strategies. Therefore, the estimation error could be minimized with the proposed model.

6.1.4 Sensitivity of weighting parameter α in the weighted-sum approach

The proportion of joint travel activities (e.g., carpooling, ridesharing, and other trip chaining strategies) could affect the weighting parameter α, intuitively. In addition, as depicted in the illustrative example in section 1.2.1, the covariance between OD demands increased with the proportion of joint travel activities during the typical period (e.g. peak hour) concerned. Therefore, the weighting parameter α can be estimated given the covariance. Fig. 6 illustrates the relationship between covariance and optimal values of weighting parameter α that minimizes WMPRE with 8 traffic sensors using the OD 2 and OD 4 as an example (with the maximum covariance among all OD pairs).

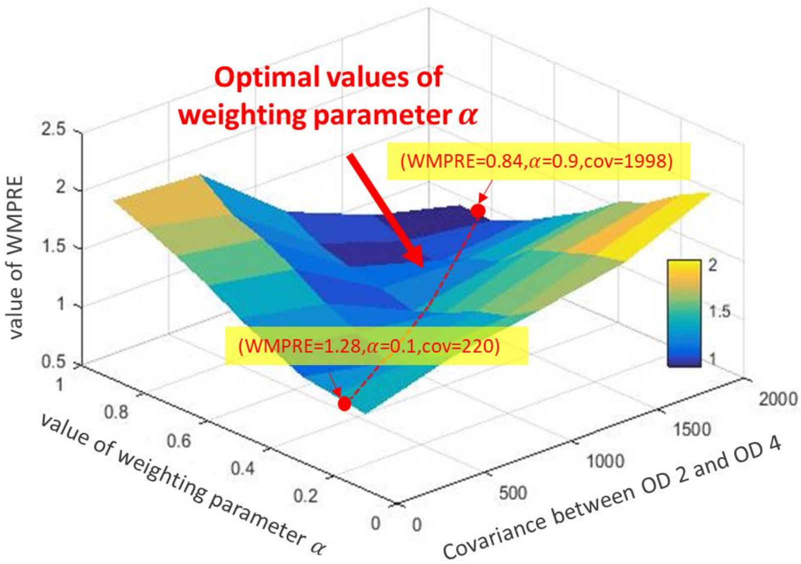


Fig. 6 Effects of covariance between OD 2 and OD 4
on the value of weighting parameter α

[The red line in the graph actually depicts the relationship between optimal weight parameter α and covariance when WMPRE is minimized]

As could be seen from Fig. 6, when the covariance between OD 2 and OD 4 is small, the optimal value of weighting parameter α is also small, and vice versa. For instance, when the covariance between OD 2 and OD 4 is **220**, WMPRE is minimum at **1.28**, and the corresponding weighting parameter is **0.1**. On the other hand, when the covariance between OD 2 and OD 4 is **1998**, WMPRE is minimum at **0.84**, and the corresponding weighting parameter α is **0.9**. This finding could be intuitively interpreted as that when covariance increases, the covariance of OD demands should be more important than the mean OD demands when evaluating the estimation accuracy, and therefore weighting parameter α also increases. Thus, a larger weighting parameter α should be chosen to improve the estimation accuracy of the OD covariance under this circumstance.

From the perspective of formulation, the weighting parameter α just quantifies the trade-offs between mean and covariance. If the road is more congested, more emphasis should be placed on the covariance, therefore the weighting parameter α should be larger. As depicted in subsection 6.1.1, different values of weighting parameter α reflect different levels of importance for estimating the mean and covariance of OD demands. For instance, $\alpha = 1$ implies that only the estimation accuracy of OD demand covariance matrix is taken into consideration for optimization of the traffic count locations. The larger the value of α, the more accurate the estimated covariance of traffic flows, and vice versa. The choice of value of α depends on the degree of covariance between OD flows. If one focuses on estimating the covariance matrix of OD demands, a larger value of α should be used and vice versa.

In view of the above discussion, how to assess the Pareto efficiency (Tan et al., 2014) becomes another interesting research question. Another method with use of the bi-objective approach is also adopted to further investigate this traffic count location problem with consideration of the effects of both the mean and covariance OD demands. The FA has also been adapted accordingly, as shown in Section 5. With the use of the Example 1 network, we compare the results from the weighted-sum and the bi-objective approaches so as to examine their relationship as depicted in Fig. 7.

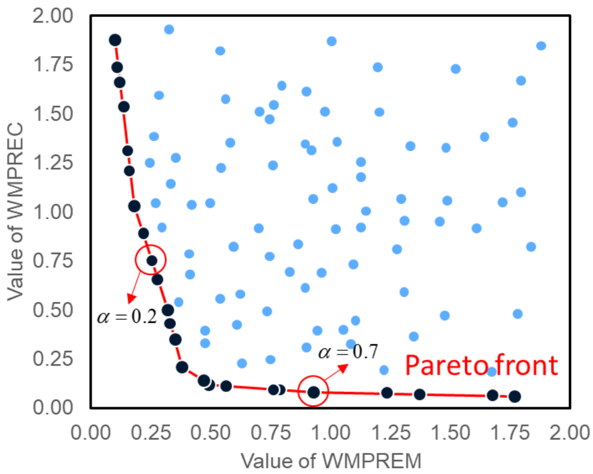


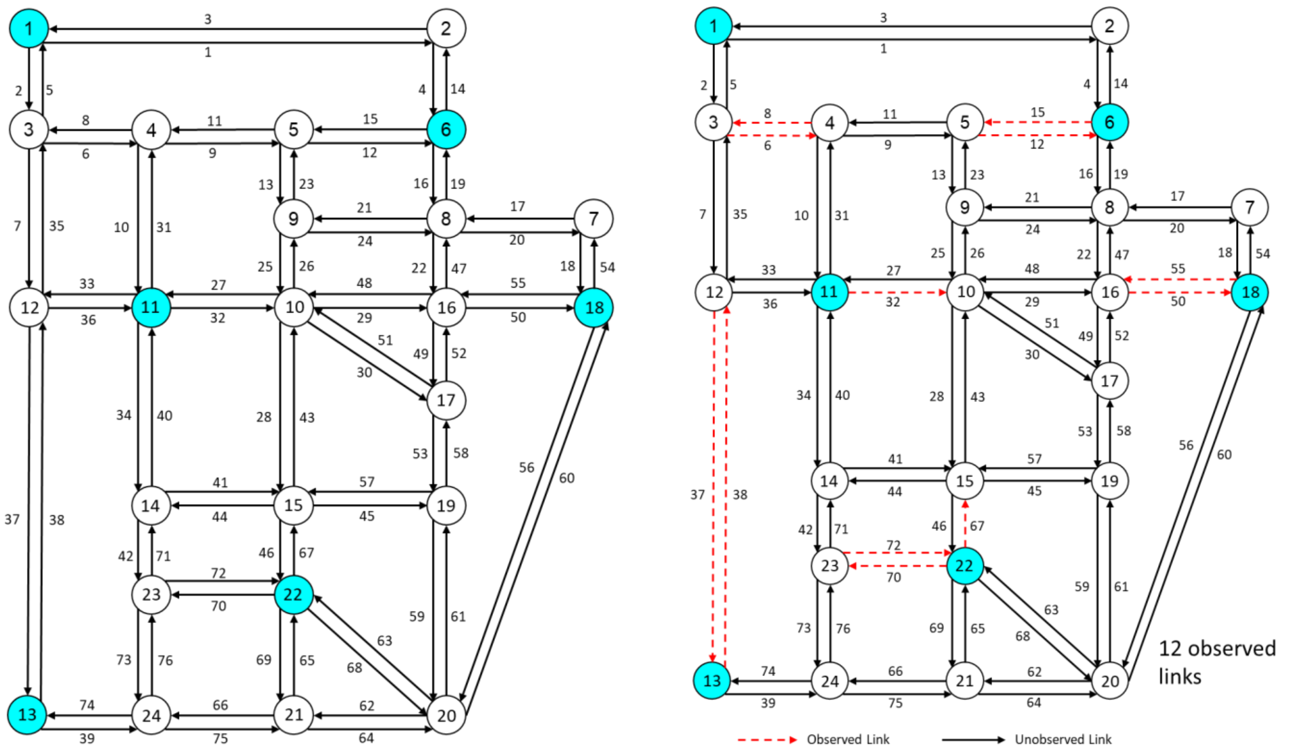
Fig. 7 Pareto optimal solutions obtained from bi-objective approach

In Fig. 7, if the value of WMPREM (or WMPREC) is equal to 1, it means that the relative error of the mean OD demand estimation (or covariance OD demand estimation) is 100%.

As shown in Fig. 7, for instance, the Pareto optimal solution when the weighting parameter α is set to be **0.2** falls on the Pareto Front. Another example when α is set to be **0.7** could also demonstrate this finding. It could be concluded that the solutions obtained from the weighted-sum approach are the sub-set of that from the bi-objective approach. A unique optimal solution could be acquired from the weighted-sum approach, with given the specified value of weighting parameter α. However, from the bi-objective approach, much more than one Pareto optimal solutions would be acquired. As such, it is more difficult to make the final decision on the traffic count location scheme by using the bi-objective approach. Thus, in practice, the weighted-sum approach is more effective for finding the optimal traffic count location scheme.

6.2 A medium-size transportation network

In this subsection, the well-known medium-size Sioux Falls network as shown in Fig. 8(a) is used to examine the performance of the proposed model and solution algorithm. This Example 2 network consists of 24 nodes, 76 links, and 30 OD pairs. Based on the budget constraint, the maximum number of traffic sensors is determined to be 12. The effects of OD covariance on the overall estimation accuracy are shown in Table 13.



(a) Origin Sioux Falls network (b) Optimal location scheme of Sioux Falls network

Fig. 8 The Sioux Falls network – Example 2 network

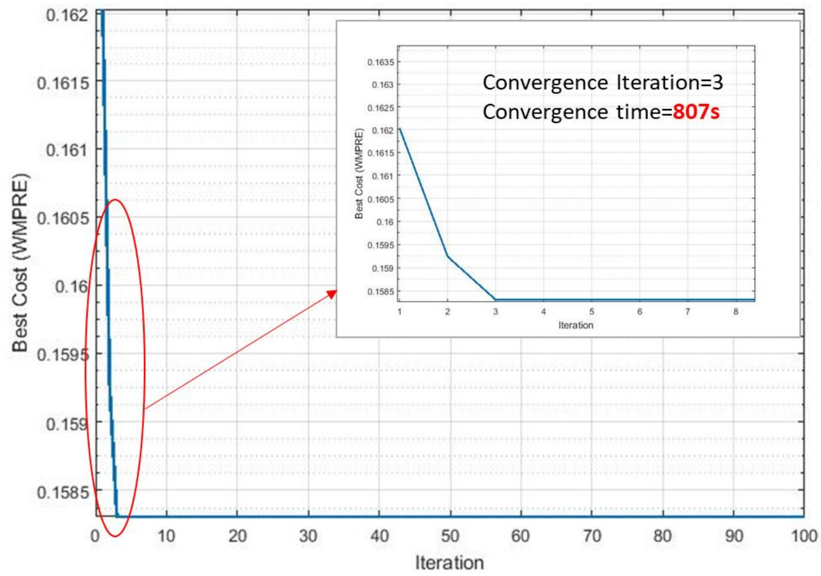
Table 13 Effects of OD covariance on objective function in the Sioux Falls network

Consider OD covariance in the objective or not	WMPRE	Relative increase [a]	WMPREM	WMPREC
No ($\alpha = 0$)	**12.32**	--	12.32	7.72
Yes ($\alpha = 0.5$)	8.26	**-33.0%**	13.01	3.51
Yes ($\alpha = 1$)	2.17	**-82.4%**	13.98	2.17

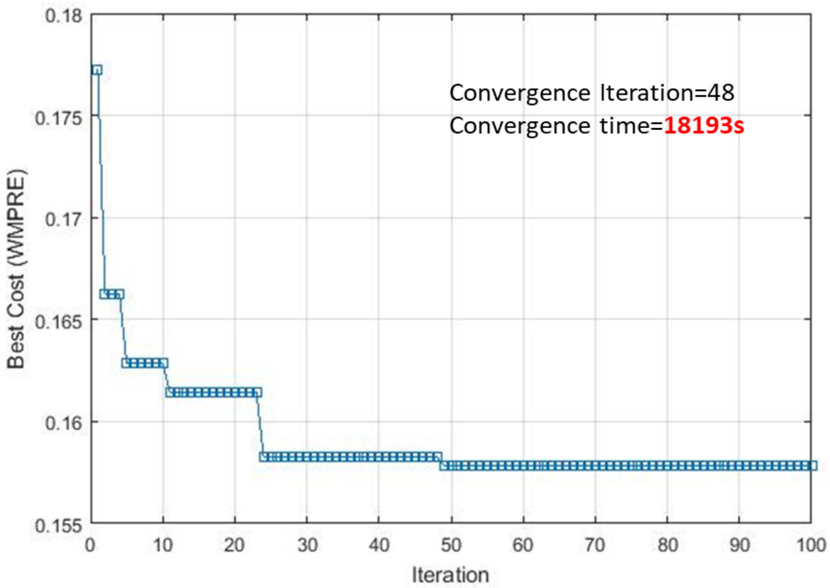
[a] Relative increase=(WMPRE($\alpha = 0.5$ or $\alpha = 1$)− WMPRE($\alpha = 0$))/WMPRE($\alpha = 0$)

From Table 13, it could be noticed that WMPREC is the largest (i.e. **12.32**) when the covariance is not taken into consideration ($\alpha = 0$). As the OD covariance ($\alpha = 1$ and $\alpha = 0.5$) is considered in the objective, the Relative increase in error% could be reduced to **-82.4%** and **-33.0%**, respectively, as compared to that without considering covariance ($\alpha = 0$). This phenomenon demonstrates again that considering covariance in the objective function could reduce the estimation error of covariance between OD demands. We could also conclude that the proposed model is particularly effective under three circumstances: (i) the number of traffic sensors is relatively small; (ii) the covariance between different OD pairs are very large; and (iii) the prior OD demand is not very close to the actual OD demands.

When using a desktop computer with the system of Intel Core i7-2600 CPU, 3.40GHz, and 8 GB RAM, the convergence time required for solving the traffic count location problem is 807 seconds by the weighted-sum approach, given that the weighting parameter α is set to be 0.5. The Pareto optimal traffic count location scheme is shown in Fig. 8(b). To demonstrate the efficiency of the proposed adapted firefly algorithm, the convergence of the adapted firefly algorithm with 12 traffic sensors is shown in Fig. 9(a) as compared to that of the classical genetic algorithm as shown in Fig. 9(b). Fig. 9 depicts that even though both of these two algorithms are able to obtain the same optimal target value, it converges at 3rd iteration by the adapted firefly algorithm, while at 48th iteration by the genetic algorithm. In addition, the convergence time of the firefly algorithm (807s) is much less than that of the conventional genetic algorithm (18193s). The above descriptions have demonstrated the efficiency of the adapted firefly algorithm.



(a) Firefly algorithm



(b) Genetic algorithm

Fig. 9 Convergence of the solution algorithm

7. Conclusions and further studies

The covariance effects are increasingly important as traffic flows between different OD pairs in a typical period (such as morning peak hour) from day to day could be statistically correlated with each other in reality. The covariance is mainly generated from the daily variation of travel patterns, network topology, and trip chaining activities of household members. For OD estimation from traffic counts, the OD estimation bias will increase, as the covariance between different OD pairs is not considered in the conventional approach. When the traffic count locations are determined without considering the covariance effect, the OD estimation accuracy will be reduced dramatically and hence the traffic counts may be located inefficiently.

In fact, different travel activity purposes (and different OD pairs) have various magnitudes of travel demands, especially in the morning peak hours. For example, the number of trips from home to CBD should be much larger than the one from home to market, in the morning peak hours during weekdays. To address this issue, a score was assigned to the OD pair based on the magnitude of the traffic flow. The criterion of measuring the accuracy of estimated mean OD demands (MPREM) has been extended as Weighted Maximum Possible Relative Error for Mean OD demands (WMPREM) and for OD demand covariance (WMPREC). The results with and without weight on the MPREM (and MPREC) have been compared to validate the proposed criteria. It could be seen that with the same number of traffic sensors, the proposed weighted maximum possible relative errors are much smaller than the criterion without weight.

It was found in the illustrative numerical examples that using the weighted-sum approach, WMPRE ((1-α)WMPREM+ α WMPREC) is a **more generalized** criterion for choosing the optimal traffic count location scheme in practice, in contrast to use of only the MPREM as the criterion. For a given traffic count location scheme and weighting parameter α, the values of WMPREM and WMPREC could not both reach to the optimum at the same time. When the number of traffic sensors increases, the variances of WMPREM and WMPREC would both change. The values of WMPREM are less than that of WMPREC for most of the traffic count location schemes when the number of traffic sensors is relatively small, and vice versa. In general, the bi-objective approach could be used to get the Pareto optimization solutions by solving the bi-objective problem. However, no unique optimal solutions could be obtained from the bi-objective approach. Therefore, in practice, the weighted-sum approach is preferable particularly when information on the proportion of trip chaining users in the study network can now be available from the AVI data.

The performance and convergence of the proposed model and solution algorithm have also been testified using the Sioux Falls network. To solve the bi-objective optimization problem, the firefly algorithm has been adapted for the weighted-sum approach and the bi-objective approach, respectively. To better understand the efficiency of the adapted firefly algorithm, the widely used classical genetic algorithm is also applied for comparison. It can be concluded from the numerical example in Section 6.2 that the adapted firefly algorithm can dramatically reduce the convergence time.

Further research should be carried out to determine the optimal weighting parameter α between WMPREM and WMPREC by the time of day (Lam and Yin, 2001). Therefore, a consistent set of the optimal traffic count location scheme across time can be identified. It has been pointed out that optimization of traffic count locations for travel time estimation should also be further investigated due to the significance of estimated travel time in practice (Fu et al., 2017; Yu et al., 2015; Zhu et al., 2018, 2017). The efficiency of the adapted firefly algorithm should be improved in further studies for solving the traffic count location problem in large-scale traffic networks. More advanced metaheuristic algorithm could be further investigated to improve the computational efficiency for realistic road networks (Xiang et al., 2015; Yu et al., 2019; N. Zhu et al., 2014). In addition, it would be worth exploring to extend the proposed model to consider the effects of multi-user classes and their covariance on the optimal traffic count location problem in a multi-modal traffic networks, when the information on vehicle composition and occupancy are available (Fu and Lam, 2018; Munizaga and Palma, 2012; Sumalee et al., 2011; Zangui et al., 2015). Further study will also be carried out with the use of the R-SUE model or AVI data for assessing the impacts of the updated stochastic link choice proportions on the traffic count location problem.

Acknowledgements

The authors would like to express their sincere thanks and sympathy to the co-author of this paper, Dr. H. P. Lo, who passed away during the drafting period on 9th August 2018. The authors are filled with the deepest sorrow and regret concerning his loss. Dr. H.P. Lo has given his continuous support and insightful advices to the authors for completion of the previous version of this paper. The authors shall sorely miss his scholarship and friendship.

The work described in this paper was financially supported by grants from National Natural Science Foundation of China (Project No. 71671184), together with grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project Nos. PolyU 152628/16E and R5029-18).

References

Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., Ciuffo, B., Djukic, T., Hoogendoorn, S., Marzano, V., Montero, L., Nigro, M., Perarnau, J., Punzo, V., Toledo, T., van Lint, H., 2016. Towards a generic benchmarking platform for origin-destination flows estimation/updating algorithms: Design, demonstration and validation. Transportation Research Part C: Emerging Technologies 66, 79–98. https://doi.org/10.1016/j.trc.2015.08.009

Bianco, L., Confessore, G., Reverberi, P., 2001. A network based model for traffic sensor location with implications on O/D matrix estimates. Transportation Science 35, 50–60. https://doi.org/10.1287/trsc.35.1.50.10140

Bielli, M., Caramia, M., Carotenuto, P., 2002. Genetic algorithms in bus network optimization. Transportation Research Part C: Emerging Technologies 10, 19–34. https://doi.org/10.1016/S0968-090X(00)00048-6

Caggiani, L., Ottomanelli, M., Sassanelli, D., 2013. A fixed point approach to origin-destination matrices estimation using uncertain data and fuzzy programming on congested networks. Transportation Research Part C: Emerging Technologies 28, 130–141. https://doi.org/10.1016/j.trc.2010.12.005

Cantelmo, G., Viti, F., Tampère, C.M.J., Cipriani, E., Nigro, M., 2014. Two-step approach for correction of seed matrix in dynamic demand estimation. Transportation Research Record: Journal of the Transportation Research Board 2466, 125–133. https://doi.org/10.3141/2466-14

Carlin, B., Louis, T., Louis, T.A., 2000. Bayes and empirical bayes methods for data analysis, second edition, Chapman & Hall/CRC Texts in Statistical Science. Chapman and Hall/CRC. https://doi.org/10.1201/9781420057669

Chootinan, P., Chen, A., Yang, H., 2005. A bi-objective traffic counting location problem for origin-destination trip table estimation. Transportmetrica 1, 65–80. https://doi.org/10.1080/18128600508685639

Clark, S., Watling, D., 2005. Modelling network travel time reliability under stochastic demand. Transportation Research Part B: Methodological 39, 119–140. https://doi.org/10.1016/J.TRB.2003.10.006

Ehlert, A., Bell, M.G.H., Grosso, S., 2006. The optimisation of traffic count locations in road networks. Transportation Research Part B: Methodological 40, 460–479. https://doi.org/10.1016/J.TRB.2005.06.001

Fei, X., Mahmassani, H.S., Murray-Tuite, P., 2013. Vehicular network sensor placement optimization under uncertainty. Transportation Research Part C: Emerging Technologies 29, 14–31. https://doi.org/10.1016/j.trc.2013.01.004

Fu, C., Zhu, N., Ma, S., 2017. A stochastic program approach for path reconstruction oriented sensor location model. Transportation Research Part B: Methodological 102, 210–237.

https://doi.org/10.1016/J.TRB.2017.05.013
Fu, X., Lam, W.H.K., 2018. Modelling joint activity-travel pattern scheduling problem in multi-modal transit networks. Transportation 45, 23–49. https://doi.org/10.1007/s11116-016-9720-8
Gentili, M., Mirchandani, P.B., 2012. Locating sensors on traffic networks: Models, challenges and research opportunities. Transportation Research Part C: Emerging Technologies 24, 227–255. https://doi.org/10.1016/j.trc.2012.01.004
Haas, C.N., 1999. On modeling correlated random variables in risk assessment. Risk Analysis 19, 1205–1214. https://doi.org/10.1023/A:1007047014741
Hazelton, M.L., 2008. Statistical inference for time varying origin–destination matrices. Transportation Research Part B: Methodological 42, 542–552. https://doi.org/10.1016/J.TRB.2007.11.003
Hu, S.R., Liou, H.T., 2014. A generalized sensor location model for the estimation of network origin-destination matrices. Transportation Research Part C: Emerging Technologies 40, 93–110. https://doi.org/10.1016/j.trc.2014.01.004
Jati, G.K., Suyanto, 2011. Evolutionary discrete firefly algorithm for travelling salesman problem. Adaptive and Intelligent Systems. ICAIS 2011. 393–403. https://doi.org/10.1007/978-3-642-23857-4_38
Lam, W.H.K., Xu, G., 1999. A traffic flow simulator for network reliability assessment. Journal of Advanced Transportation 33, 159–182. https://doi.org/10.1002/atr.5670330206
Lam, W.H.K., Yin, Y., 2001. An activity-based time-dependent traffic assignment model. Transportation Research Part B: Methodological 35, 549–574. https://doi.org/10.1016/S0191-2615(00)00010-2
Larsson, T., Lundgren, J.T., Peterson, A., 2010. Allocation of link flow detectors for Origin-Destination matrix estimation-a comparative study. Computer-Aided Civil and Infrastructure Engineering 25, 116–131. https://doi.org/10.1111/j.1467-8667.2009.00625.x
Lo, H.P., Zhang, N., Lam, W.H.K., 1996. Estimation of an origin-destination matrix with random link choice proportions: A statistical approach. Transportation Research Part B: Methodological 30, 309–324. https://doi.org/10.1016/0191-2615(95)00036-4
Munizaga, M., Palma, C., 2012. Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile. Transportation Research Part C: Emerging Technologies 24, 9–18. https://doi.org/10.1016/j.trc.2012.01.007
Nayeem, M.A., Rahman, M.K., Rahman, M.S., 2014. Transit network design by genetic algorithm with elitism. Transportation Research Part C: Emerging Technologies 46, 30–45. https://doi.org/10.1016/J.TRC.2014.05.002
Owais, M., Moussa, G.S., Hussain, K.F., 2019. Sensor location model for O/D estimation: Multi-criteria meta-heuristics approach. Operations Research Perspectives 6, 100100. https://doi.org/10.1016/j.orp.2019.100100
Sayadi, M.K., Ramezanian, R., Ghaffari-Nasab, N., 2010. A discrete firefly meta-heuristic with local search for makespan minimization in permutation flow shop scheduling problems. International Journal of Industrial Engineering Computations 1, 1–10. https://doi.org/10.5267/j.ijiec.2010.01.001
Shao, H., Lam, W.H.K., Sumalee, A., Chen, A., Hazelton, M.L., 2014. Estimation of mean and covariance of peak hour origin–destination demands from day-to-day traffic counts. Transportation Research Part B: Methodological 68, 52–75. https://doi.org/10.1016/J.TRB.2014.06.002
Shao, H., Lam, W.H.K., Sumalee, A., Hazelton, M.L., 2015. Estimation of mean and covariance of stochastic multi-class OD demands from classified traffic counts. Transportation Research Procedia 7, 192–211. https://doi.org/10.1016/J.TRPRO.2015.06.011
Shao, H., Lam, W.H.K., Tam, M.L., 2006. A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand. Networks and Spatial Economics 6, 173–204. https://doi.org/10.1007/s11067-006-9279-6

Shen, L., Shao, H., Li, C., Sun, W., Shao, F., 2019. Modeling stochastic overload delay in a reliability-based transit assignment model. IEEE Access 7, 3525–3533. https://doi.org/10.1109/ACCESS.2018.2888890

Simonelli, F., Marzano, V., Papola, A., Vitiello, I., 2012. A network sensor location procedure accounting for o–d matrix estimate variability. Transportation Research Part B: Methodological 46, 1624–1638. https://doi.org/10.1016/J.TRB.2012.08.007

Sumalee, A., Uchida, K., Lam, W.H.K., 2011. Stochastic multi-modal transport network under demand uncertainties and adverse weather condition. Transportation Research Part C: Emerging Technologies 19, 338–350. https://doi.org/10.1016/j.trc.2010.05.018

Tan, Z., Yang, H., Guo, R., 2014. Pareto efficiency of reliability-based traffic equilibria and risk-taking behavior of travelers. Transportation Research Part B: Methodological 66, 16–31. https://doi.org/10.1016/J.TRB.2013.12.003

Toi S., 1986. A study on the disposing method of traffic counting point, in: Proceeding of the 43th Annual Conference of The Japan Society of Civil Engineering. pp. 480–481.

Van Zuylen, H.J., Willumsen, L.G., 1980. The most likely trip matrix estimated from traffic counts. Transportation Research Part B: Methodological 14, 281–293. https://doi.org/10.1016/0191-2615(80)90008-9

Vardi, Y., 1996. Network tomography: estimating source-destination traffic intensities from link data. Journal of the American Statistical Association 91, 365–377. https://doi.org/10.1080/01621459.1996.10476697

Xiang, W., Zhu, N., Ma, S., Meng, X., An, M., 2015. A dynamic shuffled differential evolution algorithm for data clustering. Neurocomputing 158, 144–154. https://doi.org/10.1016/j.neucom.2015.01.058

Xiong, T., Bao, Y., Hu, Z., 2014. Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting. Knowledge-Based Systems 55, 87–100. https://doi.org/10.1016/J.KNOSYS.2013.10.012

Yang, H., 1995. Heuristic algorithms for the bilevel origin-destination matrix estimation problem. Transportation Research Part B: Methodological 29, 231–242. https://doi.org/10.1016/0191-2615(95)00003-V

Yang, H., Iida, Y., Sasaki, T., 1991. An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts. Transportation Research Part B: Methodological 25, 351–363. https://doi.org/10.1016/0191-2615(91)90028-H

Yang, H., Meng, Q., Bell, M.G.H., 2001. Simultaneous estimation of the Origin-Destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. Transportation Science 35, 107–123. https://doi.org/10.1287/trsc.35.2.107.10133

Yang, H., Zhou, J., 1998. Optimal traffic counting locations for origin–destination matrix estimation. Transportation Research Part B: Methodological 32, 109–126. https://doi.org/10.1016/S0191-2615(97)00016-7

Yang, X.-S., 2008. Nature-inspired metaheuristic algorithms.

Yang, Y., Fan, Y., Wets, R.J.B., 2018. Stochastic travel demand estimation: Improving network identifiability using multi-day observation sets. Transportation Research Part B: Methodological 107, 192–211. https://doi.org/10.1016/J.TRB.2017.10.007

Yin, Y., 2000. Genetic-algorithms-based approach for bilevel programming models. Journal of Transportation Engineering 126, 115–120. https://doi.org/10.1061/(ASCE)0733-947X(2000)126:2(115)

Yin, Y., Madanat, S.M., Lu, X.-Y., 2009. Robust improvement schemes for road networks under demand uncertainty. European Journal of Operational Research 198, 470–479. https://doi.org/10.1016/J.EJOR.2008.09.008

Yu, Q., Fang, K., Zhu, N., Ma, S., 2019. A matheuristic approach to the orienteering problem with service time dependent profits. European Journal of Operational Research 273, 488–503. https://doi.org/10.1016/J.EJOR.2018.08.007

Yu, Q., Zhu, N., Li, G., Ma, S., 2015. Simulation-based sensor location model for arterial street. Discrete Dynamics in Nature and Society 2015, 1–13. https://doi.org/10.1155/2015/854089

Zangui, M., Yin, Y., Lawphongpanich, S., 2015. Sensor location problems in path-differentiated congestion pricing. Transportation Research Part C: Emerging Technologies 55, 217–230. https://doi.org/10.1016/j.trc.2015.01.012

Zhang, L., Yin, Y., Lou, Y., 2010. Robust signal timing for arterials under day-to-day demand variations. Transportation Research Record: Journal of the Transportation Research Board 2192, 156–166. https://doi.org/10.3141/2192-15

Zhao, Y., Kockelman, K.M., 2002. The propagation of uncertainty through travel demand models: An exploratory analysis. The Annals of Regional Science 36, 145–163. https://doi.org/10.1007/s001680200072

Zhou, X., List, G.F., 2010. An information-theoretic sensor location model for traffic origin-destination demand estimation applications. Transportation Science 44, 254–273. https://doi.org/10.1287/trsc.1100.0319

Zhu, N., Fu, C., Ma, S., 2018. Data-driven distributionally robust optimization approach for reliable travel-time-information-gain-oriented traffic sensor location model. Transportation Research Part B: Methodological 113, 91–120. https://doi.org/10.1016/j.trb.2018.05.009

Zhu, N., Liu, Y., Ma, S., He, Z., 2014. Mobile traffic sensor routing in dynamic transportation systems. IEEE Transactions on Intelligent Transportation Systems 15, 2273–2285. https://doi.org/10.1109/TITS.2014.2314732

Zhu, N., Ma, S., Zheng, L., 2017. Travel time estimation oriented freeway sensor placement problem considering sensor failure. Journal of Intelligent Transportation Systems 21, 26–40. https://doi.org/10.1080/15472450.2016.1194206

Zhu, S., Cheng, L., Chu, Z., Chen, A., Chen, J., 2014. Identification of network sensor locations for estimation of traffic flow. Transportation Research Record: Journal of the Transportation Research Board 2443, 32–39. https://doi.org/10.3141/2443-04

Zhu, S., Zheng, H., Peeta, S., Guo, Y., Cheng, L., Sun, W., 2016. Optimal heterogeneous sensor deployment strategy for dynamic Origin–Destination demand estimation. Transportation Research Record: Journal of the Transportation Research Board 2567, 18–27. https://doi.org/10.3141/2567-03

Appendix A

$\mathbf{M}^{\lambda}$

The matrix $\mathbf{M}^{\lambda}$ could be obtained as below:

$$\mathbf{M}^{\lambda}\text{vec}(\boldsymbol{\lambda}^{\text{cov}}) = \mathbf{0} \tag{A.1}$$

$$\tilde{\mathrm{P}}\Sigma^{\mathbf{q}}\boldsymbol{\lambda}^{\text{cov}}\tilde{\mathrm{P}}^{T} = \mathbf{0} \tag{A.2}$$

Set $\mathbf{A} = \tilde{\mathrm{P}}\Sigma^{\mathbf{q}}$

$$\tilde{\mathrm{P}}\Sigma^{\mathbf{q}}\boldsymbol{\lambda}^{\text{cov}}\tilde{\mathrm{P}}^{T} = \mathbf{A}\boldsymbol{\lambda}^{\text{cov}}\tilde{\mathrm{P}}^{T} \tag{A.3}$$

$$\mathbf{M}^{\lambda} = [a_1 \otimes p_1 \;\; a_2 \otimes p_1 \;\; \ldots \;\; a_{\tilde{m}} \otimes p_1 \;\; a_2 \otimes p_2 \;\; \ldots \;\; a_{\tilde{m}} \otimes p_2 \;\; \ldots \;\; a_{\tilde{m}} \otimes p_{\tilde{m}}]^{T} \tag{A.4}$$

Appendix B

Proof of property 1.

Define $Y = \Sigma^{\mathrm{q}}\tilde{\mathrm{P}}^{T}$, the Equation (7) can be rewritten as:

$$\Sigma^{\mathbf{v}} = \tilde{\mathrm{P}}Y \tag{B.1}$$

For convenience, matrices $\Sigma^{\mathbf{v}}$ and Y is partitioned by column vectors as follow.

$$\Sigma^{\mathbf{v}} = [\sigma_1, \sigma_2, ..., \sigma_m] \tag{B.2}$$

$$Y = [y_1, y_2, ...y_m] \tag{B.3}$$

According to Equations (B.2) and (B.3), Equation (B.1) can be rewritten as:

$$[\sigma_1, \sigma_2, ..., \sigma_m] = \tilde{\mathrm{P}}[y_1, y_2, ...y_m] \tag{B.4}$$

If $\tilde{\mathrm{P}}$ is a matrix with full column rank, the system of linear equations $\sigma_i = \tilde{\mathrm{P}}y_i$ $(i = 1, 2, ..., m)$ has a unique solution. In other words, the matrix $Y = [y_1, y_2, ...y_m]$ would also have a unique solution.

Transpose both sides of the equation $Y = \Sigma^{\mathrm{q}}\tilde{\mathrm{P}}^{T}$, it follows that

$$Y^{T} = \tilde{\mathrm{P}}\Sigma^{\mathrm{q}T} \tag{B.5}$$

Using the similar matrix partition method, it can be proved that when $\tilde{\mathrm{P}}$ is a matrix with full column rank, $\Sigma^{\mathrm{q}T}$ can be uniquely identified according to Equation (B.5) provided Y is unique. Then, the OD demand covariance Σ^{q} can be uniquely identified if $\tilde{\mathrm{P}}$ is a matrix with full column rank. This is the end of the proof.

Appendix C

Proof of property 2.

Yang et al. (1991) proved that the WMPREM ($G(\boldsymbol{\lambda}^{mean})$) is finite if and only if the traffic flows between any OD pair are observed by at least one traffic count location. Thus, it only needs to prove the case where finite WMPREC ($H(\boldsymbol{\lambda}^{\text{cov}})$) is the necessary condition of OD Covering Rule to complete the proof of Property 2.

According to condition (ii) in the text, the relationship between covariance and mean OD demands can be obtained as follow

$$\sigma_{w,w'} = r_{w,w'}(c_w q_w)(c_{w'} q_{w'}) = r_{w,w'}(c_w c_{w'})(q_w q_{w'}) \quad \text{(C.1)}$$

According to the definitions of λ_w^{mean}, $\lambda_{w,w'}^{\text{cov}}$ and Equation (C.1)), the following relationships can be obtained,

$$\lambda_w^{mean} = (q_w^* - q_w)/q_w = q_w^*/q_w - 1 \quad \text{(C.2)}$$

$$\lambda_{w'}^{mean} = (q_{w'}^* - q_w)/q_{w'} = q_{w'}^*/q_{w'} - 1 \quad \text{(C.3)}$$

$$\lambda_{w,w'}^{\text{cov}} = (\sigma_{w,w'}^* - \sigma_{w,w'})/\sigma_{w,w'}$$

$$= (r_{w,w'} c_w c_{w'} q_w^* q_{w'}^* - r_{w,w'} c_w c_{w'} q_w q_{w'})/r_{w,w'} c_w c_{w'} q_w q_{w'}$$

$$= q_w^* q_{w'}^*/q_w q_{w'} - 1 \quad \text{(C.4)}$$

For convenience, denote $x = q_w^*/q_w$ and $y = q_{w'}^*/q_{w'}$, then it follows

$\lambda_w^{mean} = x - 1$, $\lambda_{w'}^{mean} = y - 1$, and $\lambda_{w,w'}^{\text{cov}} = xy - 1$

According to the identity relation $xy - 1 = (x-1) + (y-1) + (x-1)(y-1)$, it follows that

$$\lambda_{w,w'}^{\text{cov}} = \lambda_w^{mean} + \lambda_{w'}^{mean} + \lambda_w^{mean}\lambda_{w'}^{mean} \quad \text{(C.5)}$$

Thus, $\lambda_{w,w'}^{\text{cov}}$ is finite because λ_w^{mean} is of finite. Then, it can be seen that WMPREM ($G(\boldsymbol{\lambda}^{mean})$) and WMPREC ($H(\boldsymbol{\lambda}^{\text{cov}})$) are both finite according to Equations (13) and (24), This is the end of the proof.

Appendix D

Proof of property 3.

Similar to the proof of Property 2, we only need to prove the WMPREC ($H(\boldsymbol{\lambda}^{\text{cov}})$) is finite if the OD Covering Rule is satisfied. The method of reduction to absurdity is used. It is assumed that WMPREC ($H(\boldsymbol{\lambda}^{\text{cov}})$) is infinite and the OD Covering Rule is satisfied. On the one hand, it follows from the infinity of $H(\boldsymbol{\lambda}^{\text{cov}})$ that there exists at least one $\lambda_{w_0,w'_0}^{\text{cov}}$ ($w_0, w'_0 \in \mathbf{W}$), which is infinite (say take any real value greater than -1). On the other hand, as the OD Covering Rule is satisfied, for OD pairs $w_0, w'_0 \in \mathbf{W}$, there exist at least two (not necessarily different) traffic count locations to collect the traffic flows of these two OD pairs. Mathematically, there exist $a, b \in \mathbf{A}$ such that $p_{a,w_0} \neq 0$ and $p_{b,w'_0} \neq 0$. It then follows that $p_{a,w_0} p_{b,w'_0} \neq 0$. According to Equation (27b), it follows that

$$\lambda_{w_0,w'_0}^{\text{cov}} = \frac{-\sum_{w \in \mathbf{W}, w \neq w_0} \sum_{w' \in \mathbf{W}, w' \neq w'_0} p_{a,w} p_{b,w'} \sigma_{w,w'}^q \lambda_{w,w'}^{\text{cov}}}{p_{a,w_0} p_{b,w'_0} \sigma_{w_0,w'_0}^q} \quad \text{(D.1)}$$

$\Rightarrow$

$$\lambda_{w_0,w'_0}^{\text{cov}} = \frac{\sum_{w \in \mathbf{W}, w \neq w_0} \sum_{w' \in \mathbf{W}, w' \neq w'_0} p_{a,w} p_{b,w'} - (\sigma_{w,w'}^q \lambda_{w,w'}^{\text{cov}})}{p_{a,w_0} p_{b,w'_0} \sigma_{w_0,w'_0}^q} \quad \text{(D.2)}$$

$$\lambda_{w,w'}^{\text{cov}} \geq -1 \Rightarrow -\lambda_{w,w'}^{\text{cov}} \leq 1 \quad \text{(D.3)}$$

According to condition (ii), it follows that

$$\sigma^{q}_{w,w'} > 0 \quad \forall w, w' \in \mathbf{W} \tag{D.4}$$

Then, it follows from Inequations (D.3) and (D.4) that

$$-\sigma^{q}_{w,w'}\lambda^{\text{cov}}_{w,w'} \le \sigma^{q}_{w,w'} \quad \forall w, w' \in \mathbf{W} \tag{D.5}$$

It follows from Equation (D.2) and Inequation (D.5) that

$$\lambda^{\text{cov}}_{w_0,w'_0} \le \frac{\sum\limits_{w\in\mathbf{W},w\neq w_0}\ \sum\limits_{w'\in\mathbf{W},w'\neq w'_0} p_{a,w}p_{b,w'}\sigma^{q}_{w,w'}}{p_{a,w_0}p_{b,w'_0}\sigma^{q}_{w_0,w'_0}} \overset{define}{=} c^{\text{cov}}_{w_0,w'_0} \tag{D.6}$$

where $c^{\text{cov}}_{w_0,w'_0}$ is a positive constant, which is the upper bound of $\lambda^{\text{cov}}_{w_0,w'_0}$. Thus, $\lambda^{\text{cov}}_{w_0,w'_0}$ is bounded, which contradicts the infinity assumption of $\lambda^{\text{cov}}_{w_0,w'_0}$. Therefore, $H(\boldsymbol{\lambda}^{\text{cov}})$ is bounded if the OD Covering Rule is satisfied. The proof of necessary condition is completed.