IET Intelligent Transport Systems

Special Issue: Selected Papers from the 25th ITS World Congress Copenhagen 2018

Estimation of sparse O–D matrix accounting for demand volatility

ISSN 1751-956X
Received on 26th January 2018
Revised 22nd April 2018
Accepted on 24th May 2018
E-First on 30th July 2018
doi: 10.1049/iet-its.2018.0069
www.ietdl.org

Tao Wen[1] ✉, Chen Cai[1], Lauren Gardner[2], Steven Travis Waller[2], Vinayak Dixit[2], Fang Chen[1]

[1]Data61, CSIRO, Sydney NSW, Australia
[2]Research Centre for Integrated Transport Innovation (RCITI), School of Civil and Environmental Engineering, University of New South Wales, Sydney NSW, Australia
✉ *E-mail: tao.wen@data61.csiro.au*

Abstract: A critical issue in origin–destination (O–D) demand estimation is under-determination: the number of O–D pairs to be estimated is often much greater than the number of monitored links. In real world, some centroids tend to be more popular than others, and only few trips are made for intro-zonal travel. Consequently, a large portion of trips will be made for a small portion of O–D pairs, meaning many O–D pairs have only a few or even zero trips. Mathematically, this implies that the O–D matrix is sparse. Also, the correlation between link flows is often neglected in the O–D estimation problem, which can be obtained from day-to-day loop detector count data. Thus, sparsity regularisation is combined with link flow correlation to provide additional inputs for the O–D estimation process to mitigate the issue of under-determination and thereby improve estimation quality. In addition, a novel strategic user equilibrium model is implemented to provide route choice of users for the O–D estimation problem, which explicitly accounts for demand and link flow volatility. The model is formulated as a convex generalised least squares problem with regularisation, the usefulness of sparsity assumption, and link flow correlation is presented in the numerical analysis.

Nomenclature

$\boldsymbol{N}$	link (index) set
$\boldsymbol{M}$	O–D pair (index) set
$\boldsymbol{K}_m$	path set for O–D pair m
$\boldsymbol{V}$	vector of link flow
$\tilde{\boldsymbol{V}}$	vector of the observed expected link flow
$\boldsymbol{T}$	vector of O–D travel demand
$\tilde{\boldsymbol{T}}$	vector of the prior estimate of O–D travel demand
$\boldsymbol{T^p}$	vector of prior estimated O–D demand
$t_{n\text{f}}$	free flow travel time on link n
C_n	capacity on link n
$t_n()$	travel cost function for link n
d_n^m	users' O–D specific link choice, which represents the proportion of O–D pair demand T_m on link n
$\overline{h_k^m}$	expected flow on path k, connecting O–D pair m
$G()$	probability density function of a variable
T_m	demand variable for O–D pair m
λ_n	parameter of the Poisson distribution for flow on link n. $[\lambda] = [\lambda_1 \quad \dots \quad \lambda_n]^{\text{T}}$
γ_m	parameter of the Poisson distribution for O–D pair demand T^m
$\delta_{n,k}^m$	link-path indicator variable. $\delta_{n,k}^m = \begin{cases} 1 & \text{if link } n \text{ is on path } k \text{ between OD pair } m \\ 0 & \text{otherwise} \end{cases}$
α	parameter of the bureau of public roads (BPR) function
β	parameter of the BPR function
$\boldsymbol{A}$	assignment map matrix, which represents the proportion of O–D pair demand T_m traversing link n
$\boldsymbol{Y}$	covariance matrix of observed link flows
δ_1	weight parameter for the prior O–D estimation
δ_2	weight parameter of the L1 regularisation
p_k^m	proportion of flow on path k, connecting OD pair m, must be non-negative

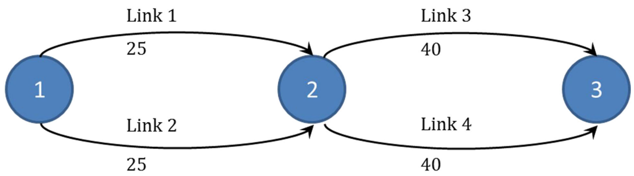
1 Introduction

The development of a country brings changes in the land-use and economic state of affairs, and the number of trips could vary accordingly. It is therefore important to identify the frequency of trips between different centroids in a network for policy makers and transport planners. A common practice is to infer origin–destination (O–D) matrix using observed traffic flows collected by various techniques such as field count, loop detectors or camera. A simple example below illustrates why O–D matrix is necessary information in the transportation planning process.

In Fig. 1, assume people will evenly split themselves among all possible routes. The number below each link indicates the number of people on that link. Based on the assumption, all three O–D matrices can produce the same set of link flows specified in the figure. However, they differ from each other significantly. Now, say an expressway will be constructed which directly connects centroid 1 and 3, but what if the second O–D matrix in the figure is the actual one? In this case, the expressway would just be a waste of money. Therefore, a reliable O–D matrix is fundamental and vital for transportation planning.

Traditionally, the O–D matrix is obtained from plate surveys, household surveys or roadside surveys [1]. Such survey activities may suffer from limited response, financial constraint or sampling coverage. Additionally, by the time the survey data are collected and processed, the O–D data obtained become obsolete. Nowadays, the dissemination and application of some modern technologies such as inductive loop detectors, cellular phones and automatic license plate recognition systems have provided an efficient way to collect up-to-date traffic count data, as an alternative to the traditional approaches, enhanced O–D matrix estimation methodologies could prove useful for transportation planning.

It is a statistical approach for estimating or calibrating an O–D matrix from observed traffic counts and some prior knowledge of the O–D demand (which is represented by O–D matrix in this paper). In the past, many models have been proposed and widely applied for O–D matrix estimation [2, 3]. The accuracy of these calibrated matrices depends on the calibration model used, the input data errors and on which set of link counts is available. Comparing to the traditional approaches, O–D matrix calibration is much more cost-efficient and time-efficient, besides, due to that



O-D pair	Number of trips
1-2	20
1-3	30
2-3	50

O-D pair	Number of trips
1-2	50
1-3	0
2-3	80

O-D pair	Number of trips
1-2	30
1-3	20
2-3	60

Fig. 1 *Simple example*

traffic counts are collected automatically every day, the calibrated results are always up-to-date. However, it is difficult to infer a unique O–D matrix directly using these approaches because the number of O–D pairs is much larger than the number of links, and users' routing mechanism is unknown; thus some assumptions or some prior information of O–D matrix is necessary to guarantee a unique solution.

The issue that the number of O–D pair to be estimated is often much greater than the number of monitored links is denoted as ill-posedness or under-determination. The assumption that O–D matrix tends to be sparse can be used to mitigate such an under-determination issue. The assumption of a sparse matrix originates from the commonly observed phenomenon that some centroids tend to be more popular than others, and only few trips are made for intro-zonal travel. Consequently, a large portion of trips will be made for a small portion of O–D pairs, that is, there are a lot of O–D pairs with only a few or even zero trips. Mathematically, this implies that the O–D matrix is a sparse matrix. The sparsity of O–D matrix is represented as the L1 regularisation in the model [4].

As a fundamental element of the transportation planning process, O–D trip matrix plays a principal role and can have a significant impact on the prediction results. O–D demand is inherently volatile and may vary day-to-day due to various factors. This paper explicitly treats demand as a causal variable: the correlation and variation of link flow are caused by the demand volatility. This is a different interpretation of link flow variation, which was often explained as a measurement error in the generalised least squares method [5, 6]. Such a notion allows the utilisation of link flow correlation, which can be obtained from loop detectors on multiple days, and could hence improve the estimation quality [7–9].

Additionally, users' route choice information are often assumed known, or obtained with an elaborate computation. Besides, the inherent volatility in demand and the resulted link flow correlation are often neglected, whilst both could influence users' route choice significantly. Therefore, in addition to the sparsity regularisation, the strategic user equilibrium for independently distributed O–D demands (I-STRUE) [10] is also implemented to account for demand uncertainty. I-STRUE is a user equilibrium assignment model which assumes that travellers choose a route to minimise their expected travel cost, where their decision is based on knowledge of a demand distribution, rather than a deterministic demand value. I-STRUE is defined such that at equilibrium, all used paths have equal and minimal expected travel costs. I-STRUE can account for demand volatility while maintaining computation efficiency. The consistent incorporation of demand volatility in both route choice assignment and O–D demand estimation helps provide a robust estimation of O–D matrix.

In general, sparsity regularisation is combined with link flow correlation to provide additional inputs for the O–D estimation process. The usefulness of sparsity assumption and link flow correlation is presented in the numerical analysis, and the uniqueness of the model is proved. The highlights of the proposed model are:

(i) The non-negativity of O–D demands is included in the optimisation problem.
(ii) The sparsity of O–D matrix is accounted for and is used as a regularisation to enhance estimation quality.
(iii) Link flow correlation is incorporated to improve more information for the under-determined O–D estimation problem.
(iv) A novel assignment model is implemented to provide users' route choice information, which accounts for demand volatility while maintaining the computation efficiency.

2 Background

From the past studies, O–D matrix estimation can be categorised as gravity model [11], growth factor model [12, 13] and traffic count data-based O–D matrix estimation [3]. This paper focuses on the traffic count data-based O–D matrix estimation problem, which mainly relies on statistical approaches using traffic count data. However, the problem is often challenging due to that the number of observable links in a traffic network is often much smaller than the number of O–D pairs to be estimated. Therefore, it may not be possible to obtain a unique solution from a single set of link counts alone. As a consequence, various forms of additional assumptions and a priori knowledge are required to obtain a unique solution.

A wide range of statistical O–D estimation methods have been proposed according to their assumptions, including the generalised least square method [5, 6], the maximum likelihood method [14], bi-level programming approach [15], Bayesian approaches [16, 17] and maximum entropy [18]. Integration of the methods mentioned above was also of recent interest [1, 19].

Basically, the objective of traffic count data-based O–D matrix estimation is to optimise an objective function (which may vary based on model requirements) subject to a set of constraints (typically the assignment of O–D flows, such as (3); and positiveness of O–D trips and link flows, such as (2) and (4)). Mathematically, the problem is to find an optimal O–D matrix $\boldsymbol{T}^* = [T_1, \ldots, T_m]^{\text{trans}}$ such that

$$\boldsymbol{T}^* = \arg\min : f_1\left(\boldsymbol{T}, \tilde{\boldsymbol{T}}\right) + f_2\left(\boldsymbol{V}, \tilde{\boldsymbol{V}}\right) \tag{1}$$

Subject to

$$\boldsymbol{T} \geq 0 \tag{2}$$

$$\boldsymbol{V} = \boldsymbol{A}\boldsymbol{T} \tag{3}$$

$$\boldsymbol{V} \geq 0 \tag{4}$$

where $\boldsymbol{T}$ is the target O–D matrix to be estimated; $\boldsymbol{T}^*$ is the optimal/estimated O–D matrix; $\tilde{\boldsymbol{T}}$ is a prior O–D matrix; $\boldsymbol{A}$ is the assignment matrix, which represents the proportion of O–D trips on a link; $\boldsymbol{V}$ is a vector of link flow produced by the target O–D matrix; $\tilde{\boldsymbol{V}}$ is a vector of observed link flow, normally obtained from traffic count data.

Note that $()^{\text{trans}}$ represents the transpose of a vector/matrix. As aforementioned, additional assumptions are required to ensure solution uniqueness, which is reflected in the measurement functions $f_1(), f_2()$. Among all the statistical O–D estimation methods, one advantage of the generalised least squares method is that no distributional assumptions on the data are required, which increases the method's flexibility. Also, the method associates survey data directly with traffic count data, while considering the relative accuracy of these data [6]. The method is also proved to be useful in exploiting O–D matrix structure [20].

The aforementioned statistical approaches are mainly applied in static networks. However, sometimes the temporal impact and congestion effect should not be neglected; in this case, a time-dependent O–D matrix is required [21–24]. Additionally, the O–D estimation problem has been extended to account for the stochastic nature of observed flows [25, 26], Some computer-aided heuristic algorithms have also been applied to this problem such as the genetic algorithm [27, 28]. The main advantage of the genetic algorithm is its capability of solving non-convex, complex optimisation, while the drawback is that the solution is not guaranteed to be optimal. Some other methods have also been proposed by researchers to enhance the model applicability, such as multi-class O–D estimation [28, 29], fuzzy-based approach [30, 31] and neutral network based approach [32]. However, issues regarding computation complexity and the application to large-scale networks still remain a challenge. The model proposed in this paper can provide a unique estimation of O–D matrix, while maintaining the computation simplicity.

On the other hand, higher order information of a network, such as the variance and covariance of observed link flows, can potentially provide more constraints to the traffic count data-based model. This is considered as network tomography problems in statistics and computer science literature [2, 33, 34], but its application in transportation models is yet to be fully explored. Cremer and Keller [35] demonstrated that aggregating or averaging link count data collected over a sequence of time period may result in the loss of important information. Hazelton [9] proposed a weighted least squares method to account for the covariance of links and assumed a parameter to explain the circumstances when the variance exceeds the mean if a Poisson distribution is used. Bell [36] proposed a maximum likelihood method and found the analytical solution to the covariance of O–D matrix by using a Taylor approximation. However, measurement error is the main source of uncertainty in literature. This paper consistently accounts for demand volatility and the resulted link flow correlation, the latter one can be easily provided by loop detectors data on a daily basis.

Estimation of the O–D trip matrix also requires a robust assignment model. Logit-based stochastic assignment model was incorporated in a linear programming model, such models are called path flow estimation based models [37], these models either needs path enumeration [38] or information on the set of shortest paths [39]. However, when applying the assignment model to a large network, realism and computational complexity are both critical in determining a model's practical applicability. Further, a major complication in transportation modelling is the ability to properly account for the inherent uncertainties regarding demand [40, 41] and capacity levels [42, 43]. Additionally, as has been noted, uncertainty regarding these variables directly affects route choice behaviour [44] and traffic predictions [45]. It is, therefore, necessary to incorporate these stochastic elements into models to ensure robust planning capabilities, but to do so in a manner that maintains computational tractability. The strategic user equilibrium [46, 47] effectively accounts for the impact of demand uncertainty subject to Wardrop's UE conditions, and under the static user equilibrium framework, the computation tractability and simplicity are preserved. The model was extended to the dynamic traffic assignment [48], independently distributed O–D demands [10] and road pricing scheme [49].

In the proposed model, the assumption of sparse O–D matrix is represented by the L1 regularisation, because minimisation of the L1 regularisation term induces a sparse solution. Previously researchers have used L1 regularisation to account for network anomalies [50–52], the impact of path flow sparsity in the O–D estimation problem was also explored [53]. However, few researches have considered that the O–D demands should be non-negative in the regularisation problem, which is an important constraint in the O–D estimation process. Menon *et al.* [4] provide an in-depth discussion of the importance of non-negativity in O–D demands in ill-posed problems and showed that it could be useful in providing a potentially unique solution. Some researchers have focused on the Lagrangian dual of the for L1 regularisation, which regards the problem as an example of the basis pursuit principle. The advantage is to avoid tuning weight parameters for the regularisation term, with a compromise of spending more time on the optimisation procedure [53–56].

In Section 3, the model is formulated as a convex generalised least squares problem with regularisation. The usefulness of the sparsity assumption and link flow correlation are presented in the Sioux Falls network in Section 4. Section 5 provides a conclusion and possible future research direction.

3 Model formulation

This section defines the mathematical formulation of our proposed model; a summary of the notations used in the section is listed as follows.

Traditionally, the application of the generalised least squares method in transport O–D estimation problem is to find an O–D matrix that minimises the squared Mahalanobis distance of two residual vectors: the vector of link flow and the vector of O–D demands. The method can be formulated as the following optimisation program:

$$\begin{aligned} \boldsymbol{T}^{\mathrm{GLS}} = \arg\min : & (\boldsymbol{T} - \tilde{\boldsymbol{T}})^{\mathrm{trans}} \boldsymbol{Z}^{-1} (\boldsymbol{T} - \tilde{\boldsymbol{T}}) \\ & + (\boldsymbol{V} - \tilde{\boldsymbol{V}})^{\mathrm{trans}} \boldsymbol{Y}^{-1} (\boldsymbol{V} - \tilde{\boldsymbol{V}}) \end{aligned} \tag{5}$$

Subject to

$$\boldsymbol{T} \geq 0 \tag{6}$$

$$\boldsymbol{V} = \boldsymbol{A}\boldsymbol{T} \tag{7}$$

$$\boldsymbol{V} \geq 0 \tag{8}$$

where Y^{-1} indicates the inverse of the variance–covariance matrix of the 'errors' in observed link flows, Z^{-1} indicates the inverse of the variance–covariance matrix of O–D demands, $()^{\mathrm{trans}}$ means the transpose of a matrix. Since the O–D estimation problem is generally an ill-posed problem if no prior information is available, it is necessary to utilise as much information as possible. However, a majority of the previous literature has often neglected the correlation between link flows, that is, Y is considered as a diagonal matrix whose diagonal elements represent the variance of the observation error. In addition, the non-negativity constraint is often ignored during the optimisation process. To mitigate the under-determination of such a problem, these two issues will be discussed and resolved below, to provide a robust estimation of O–D matrix.

3.1 Correlation of link flows

Incorporating the correlation of link flows can potentially utilise more information from observed link flows. Given a network of n links, n pieces of information are input to the optimisation problem if only the variance of link flow is considered, but n^2 pieces of information are utilised if the covariance matrix is incorporated, because the dimension of expected link flow vector is n by 1 while the dimension of covariance of link flows is n by n. A key assumption is made here to account for the covariance of link flows.

A1: Each O–D pair demand follows the Poisson distribution and is independent of each other.

Note that a Poisson variable is non-negative, which is consistent with real-world demand. As a univariate distribution defined only for integers it has been adopted in many previous literature [3, 6, 9, 12, 57]. However, actual demand distribution may vary depending on network type, time frame and many other factors, which distribution best fit the actual demand is an open question. In this paper, such an assumption allows us to interpret the error term differently from other literature: here the ‘error’ is actually the volatility of link flow which is caused by the demand uncertainty, instead of measurement errors. In other words, each O–D pair demand is a ‘causal’ variable which is projected to each link by the assignment map matrix $\boldsymbol{A}$. More importantly, such an interpretation of the covariance matrix $\boldsymbol{Y}$ allows us to scrutinise the correlation between each link. Thus, given the observed link flows on multiple days, we will be able to find the covariance matrix of link flows:

$$\text{Obs} = \begin{bmatrix} l_1^1 & \dots & l_1^d \\ \vdots & \ddots & \vdots \\ l_n^1 & \dots & l_n^d \end{bmatrix} \tag{9}$$

$$\boldsymbol{Y} = \text{Covariance(Obs)} \tag{10}$$

As each O–D pair demand follows a Poisson distribution, they can be defined by their corresponding parameters (for a Poisson distribution, these parameters are each variable's expectation)

$$\boldsymbol{T} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{bmatrix}, \quad m \in M \tag{11}$$

Therefore, the objective is to find a vector of demand parameter for the generalised least squares method. Before proceeding further in solving the objective function, we introduce another regularisation term for our optimisation problem.

3.2 Regularisation inducing sparsity and non-negativity of demand

It is a common phenomenon that only a small portion of O–D pairs will have a large number of trips, especially for commuter trips, which implies that the O–D matrix tends to be a sparse matrix. Hence, it is assumed here that the O–D matrix is sparse to some extent. In our optimisation problem, the sparsity of an O–D matrix can be obtained by adding the L1 regularisation to the objective function

$$\begin{aligned} \boldsymbol{T}^{\text{GLS}} = \arg\min :& (\boldsymbol{A}\boldsymbol{T} - \tilde{\boldsymbol{V}})^{\text{trans}}\boldsymbol{Y}^{-1}(\boldsymbol{A}\boldsymbol{T} - \tilde{\boldsymbol{V}}) \\ &+ \delta_1(\boldsymbol{T} - \tilde{\boldsymbol{T}})^{\text{trans}}(\boldsymbol{T} - \tilde{\boldsymbol{T}}) + \delta_2\boldsymbol{T}_1 \end{aligned} \tag{12}$$

where $\boldsymbol{T}_1$ represents the L1 norm of a vector, $\boldsymbol{Y}^{-1}$ indicates the inverse of the variance–covariance matrix of link flows, δ_1 and δ_2 are weight parameters which are contingent on our belief of the prior estimates and sparsity of O–D matrix, respectively. The L1 norm here can be expressed as

$$\boldsymbol{T}_1 = \sum_{m \in M} |\gamma_m| \tag{13}$$

To explain why minimisation of the L1 norm induces sparsity, we start from the L0 norm of a vector, which is

$$\boldsymbol{T}_0 = \sum_{m \in M} \widehat{\gamma_m} \tag{14}$$

$$\widehat{\gamma_m} = \begin{cases} 1, & \gamma_m \neq 0 \\ 0, & \gamma_m = 0 \end{cases} \tag{15}$$

Clearly, the L0 norm is the number of non-zero elements in the O–D matrix. Previous literature on the Lasso algorithm and on the compressed sensing has suggested that under some assumptions, minimising L1 norm can approximate the minimisation of L0 norm [58, 59], that is, minimisation of L1 norm induces a sparse O–D matrix.

The function of L1 norm is not differentiable everywhere, however, one may note that the parameter for each O–D pair is non-negative, which allows us to take the absolute-value sign out of the objective function. This allows us to write the objective function as the following form:

$$\begin{aligned} z(\boldsymbol{T}) = \arg\min_{\boldsymbol{T} \geq 0}\, & (\boldsymbol{A}\boldsymbol{T} - \tilde{\boldsymbol{V}})^{\text{trans}}\boldsymbol{Y}^{-1}(\boldsymbol{A}\boldsymbol{T} - \tilde{\boldsymbol{V}}) \\ &+ \delta_1(\boldsymbol{T} - \boldsymbol{T}^p)^{\text{trans}}(\boldsymbol{T} - \boldsymbol{T}^p) + \delta_2 \sum_{m \in M} \gamma_m \end{aligned} \tag{16}$$

Hence, the objective function becomes differentiable everywhere. When δ_2 equals 0, the problem is exactly the classical generalised least squares O–D estimation problem. The differentiability enables us to prove the convexity of the objective function we take the second partial derivatives with respect to the demand vector, which provides us the Hessian matrix of the objective function

$$\frac{\partial z(\boldsymbol{T})}{\partial \boldsymbol{T} \partial \boldsymbol{T}} = 2\boldsymbol{A}^{\text{trans}}\boldsymbol{Y}^{-1}\boldsymbol{A} + 2\delta_1\boldsymbol{I} \tag{17}$$

where I is an identity matrix of dimension m by m. Clearly, the Hessian matrix is positive definite, hence the proposed model has a unique optimal solution. It is vital that the model can provide a unique estimation of O–D matrix, because it guarantees that the proposed model is applicable to a variety of transportation planning process.

The use of non-negative O–D parameters is consistent with our intuition – the number of trips made between each O–D pair should always be greater than or equal to zero. In addition, such a constraint allows the assumption of Poisson distribution. If the non-negativity constraint is not considered, we may obtain an analytical optimal solution by taking the first partial derivative. Many researchers have solved the generalised least squares O–D estimation by such a closed-form update technique. Notwithstanding this, it is hard to interpret a negative O–D demand, so we avoid a negative solution in our formulation, the optimisation problem with non-negative constraints can be solved by various methods such as gradient descent method.

3.3 Regarding the assignment map matrix and demand volatility

Due to the above interpretation that the link flow variation is caused by demand volatility, we need to account for these uncertainties in the assignment map matrix $\boldsymbol{A}$. That is, users should consider demand volatility when making their route choice decision. Such a goal can be achieved by applying the I-STRUE. I-STRUE is defined such that the expected travel costs are equal on all used paths, and this commonly expected travel time is less than the actual expected travel time on any unused path. In other words, given user equilibrium expected path cost, any deviation from the existing expected path flows cannot reduce the expected path cost. This notion can be formulated as the following mathematical program:

$$\min z([\lambda]) = \sum_{n \in N} \int_0^{\lambda_n} \int_{-\infty}^{+\infty} t_n(l_n)G(l_n)\,\mathrm{d}l_n\,\mathrm{d}\lambda_n \tag{18}$$

Subject to

$$\sum_{k \in K_m} \overline{h_k^m} = \gamma_m, \quad \forall m \in M \tag{19}$$

$$\overline{h_k^m} \geq 0, \quad \forall m \in M, \quad k \in K_m \tag{20}$$

$$\lambda_n = \sum_{m \in M} \sum_{k \in K_m} \overline{h_k^m}\delta_{n,k}^m, \quad \forall n \in N \tag{21}$$

where l_n is the link flow variable which will be integrated from negative infinity to positive infinity to represent expected travel cost on a link: $\int_{-\infty}^{+\infty} t_n(l_n)G(l_n)\,\mathrm{d}l_n$. It is a function of λ_n. The mathematical program has been proved to be equivalent to our

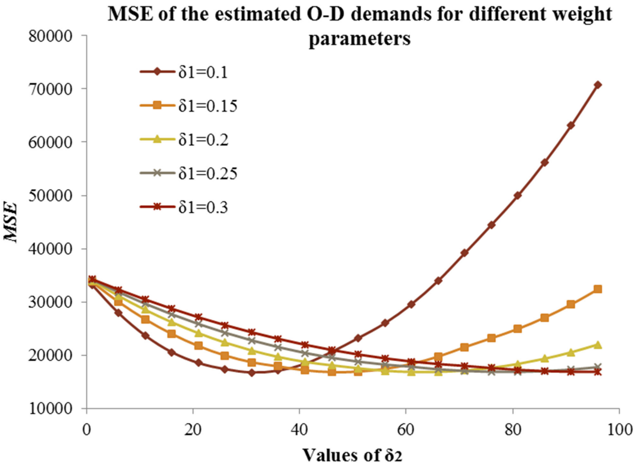


Fig. 2 *Impact of weight parameters on the estimated results*

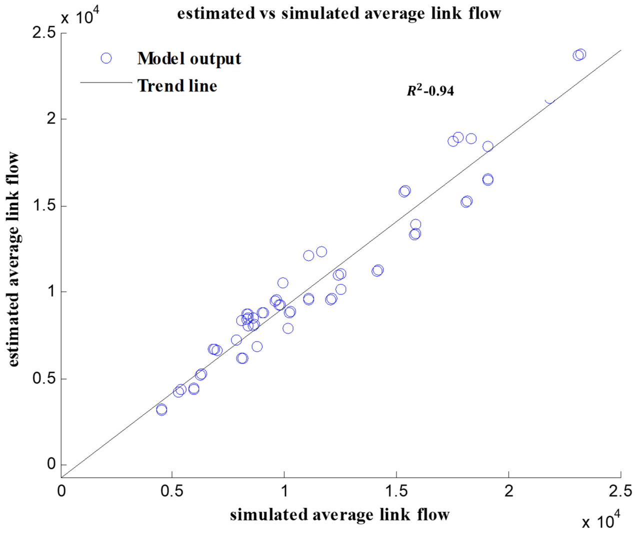


Fig. 3 *Regression analysis of estimated and simulated expected link flow*

definition of I-STRUE. If the expected path flow is expressed as a proportion multiplied by the expected demand, that is

$$\lambda_n = \sum_{m \in M} \sum_{k \in K_m} \overline{h_k^m} \delta_{n,k}^m = \sum_{m \in M} \sum_{k \in K_m} p_k^m \delta_{n,k}^m s_m \tag{22}$$

Then, we can obtain the O–D specific link proportions, also known as the assignment map matrix, by the following equations:

$$d_n^m = \sum_{k \in K_m} p_k^m \delta_{n,k}^m, \quad \forall m \in M, \ \forall n \in N \tag{23}$$

$$\boldsymbol{A} = \begin{bmatrix} d_1^1 & \dots & d_1^m \\ \vdots & \ddots & \vdots \\ d_n^1 & \dots & d_n^m \end{bmatrix} \tag{24}$$

The assignment map matrix represents the proportion of O–D pair demand T_m traversing link n, it indicates the proportion of link flow disaggregated by different O–D pairs, which is extremely important in many transportation applications such as O–D matrix estimation, emission analysis and network design problem. To summarise, the implementation of I-STRUE accounts for the demand uncertainty in users' routing mechanism while maintaining the computation simplicity under the classical user equilibrium formulation.

4 Numerical demonstration

The objective of the analysis is to test if the proposed model can effectively estimate the O–D trips from day-to-day observed link flows. This analysis is conducted on the Sioux Falls network (which has 24 nodes and 76 links). The network properties are pre-defined in [60]. Each O–D demand is assumed to follow a Poisson distribution and is independent of each other. The BPR function parameters α and β are set to 0.15 and 4.0, respectively. The prior O–D matrix is assumed to be a 10% overestimation of the simulated one, that is, the simulated O–D matrix will be inflated by 10% to represent the prior O–D matrix, to demonstrate the fact that a prior O–D matrix is not perfectly accurate.

To collate observed link flow data, a Monte-Carlo simulation is conducted. It consists of running the strategic user equilibrium model and generating random link flow samples accordingly. First, we run the I-STRUE based on the demand matrix pre-defined in Bar-Gera (referred to as simulated O–D matrix in this section) and obtain the assignment map matrix. Then 10,000 O–D matrices are sampled independently. Finally, these sampled O–D matrices are assigned to each link according to the assignment map matrix. The resulted simulated link flows are used to represent observed day-to-day link flow discussed in (9) and (10). The impact of the weight parameters δ_1 and δ_2 are explored, in conjunction with how the sparsity regularisation will facilitate improving the estimation. The estimated O–D matrix should closely approximate the simulated one; the link flow distributions reproduced by the I-STRUE model based on estimated O–D matrix should also closely match the simulated link flows.

To evaluate the performance of the proposed model, the mean square error (MSE) is introduced here

$$\text{MSE} = \sum_{m \in M} \frac{(\gamma_m - \gamma_m^*)^2}{K} \tag{25}$$

where γ_m^* and γ_m denote the simulated O–D demand and estimated O–D demand, respectively, K is the number of O–D pairs. MSE indicates how the estimation deviates from the simulated O–D matrix. In Fig. 2, the impact of the weight parameters is demonstrated. When δ_1 is fixed to 0.25 or 0.3, as illustrated in the two series in the figure, the MSE of the estimated O–D demands drops with the increase of δ_2. On the contrary, If δ_1 is fixed to 0.1, 0.15 or 0.2, the curve is similar to parabola, and the MSE will climb up after the minimum is reached. The figure shows that the incorporation of sparsity regularisation improves the estimated results in general; however, if we excessively amplify the importance of sparsity (that is, if we believe the O–D matrix is very sparse, which is not the case for the Sioux Falls network), the regularisation may have detrimental impact on the estimation. Therefore, the choice of the weight parameters should be scrutinised according to different cases rather than apply a set of uniform values.

In Fig. 3, the simulated expected link flow and the corresponding estimated mean link flow are plotted from the smallest to the largest. The estimated expected link flows are produced by the I-STRUE model based on the estimated O–D matrix (when $\delta_1 = 0.1, \delta_2 = 31$). The R-squared value of 0.94 illustrates that the estimated expected link flow closely approximates the simulated expected link flow. Hence, the proposed model is capable of finding an estimated O–D matrix that produces a set of link flow similar to the simulated one while being as sparse as possible.

Fig. 4 demonstrates the sparsity level for different δ_2. According to [61, 62], the Hoyer's formula provides a scalable, normalised and generalised sparsity measure. It is therefore adopted to illustrate the sparsity level of the estimated O–D matrix

$$H = \left(\sqrt{K} - \frac{\sum_{m \in M} \gamma_m}{\sqrt{\sum_{m \in M} \gamma_m^2}}\right) / (\sqrt{K} - 1) \tag{26}$$

The weight parameter δ_1 is fixed to 0.1, and δ_2 will vary from 1 to 96 with an increment of 5.

It is presented that the estimated O–D matrix becomes sparser with the increase of δ_2. Therefore, as discussed in the previous section, minimisation of the L1 norm regularisation can induce a

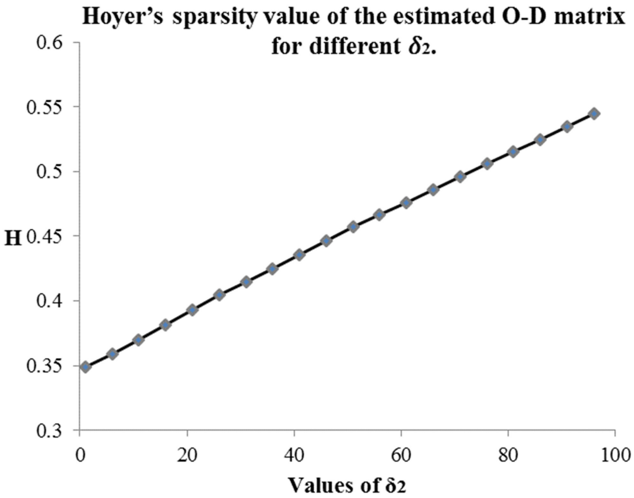


Fig. 4 *Hoyer's sparsity value of the estimated O–D matrix for different δ_2*

sparse solution. Additionally, if it is believed that the true O–D matrix tends to be sparse, a higher weight should be imposed on δ_2.

5 Conclusion

Under-determination is a common issue in O–D matrix estimation problem, to mitigate such an issue, this paper incorporates sparsity regularisation and link flow correlation in the generalised least squares method. In addition, a specific assignment model, I-STRUE, is implemented to provide users' route information while accounting for demand volatility. The solution to the proposed mathematical formulation has been proved to be unique, which is critical to have stable modelling outputs. The application of I-STRUE in O–D estimation is also novel, which can account for demand volatility and users' strategic route choice mechanism in the O–D estimation framework. The numerical analysis suggests that sparsity regularisation can improve estimation quality if treated properly, and the link flows produced based on the estimated O–D matrix can closely approximate the observed link flows. Therefore, by utilising the sparsity feature of O–D matrix, as well as the link flow correlation, the model is capable of providing a more robust estimation of O–D matrix. Moreover, if the sparsity level of O–D matrix is known, one can tune the parameters in the model to further improve output accuracy.

However, every model has its limitation. The weight parameter for the sparsity regularisation term needs to be tuned according to different cases. As a consequence, additional investigation on O–D matrix sparsity may be required to determine the optimal weight parameters before implementation of the model. Another issue is the assumption of Poisson distributed O–D demand, which is not always the case in real world. Hence, extending the model to other distributions can enhance the model's applicability.

6 References

[1] Castillo, E., Menéndez, J.M., Sánchez-Cambronero, S., *et al.*: 'A hierarchical optimization problem: estimating traffic flow using gamma random variables in a Bayesian context', *Comput. Oper. Res.*, 2014, **41**, pp. 240–251

[2] Hazelton, M.L.: 'Network tomography for integer-valued traffic', *Ann. Appl. Stat.*, 2015, **9**, (1), pp. 474–506

[3] Bera, S., Rao, K.: 'Estimation of origin-destination matrix from traffic counts: the state of the art', *Eur. Transp.*, 2011, **49**, pp. 2–23

[4] Menon, A.K., Cai, C., Wang, W., *et al.*: 'Fine-grained OD estimation with automated zoning and sparsity regularisation', *Transp. Res. B, Methodol.*, 2015, **80**, pp. 150–172

[5] Cascetta, E.: 'Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator', *Transp. Res. B, Methodol.*, 1984, **18**, (4–5), pp. 289–299

[6] Bell, M.G.H.: 'The estimation of origin-destination matrices by constrained generalised least squares', *Transp. Res. B, Methodol.*, 1991, **25**, (1), pp. 13–22

[7] Goel, P., McCord, M., Park, C.: 'Exploiting correlations between link flows to improve estimation of average annual daily traffic on coverage count segments: methodology and numerical study', *Transp. Res. Rec., J. Transp. Res. Board*, 2005, **1917**, pp. 100–107

[8] Hazelton, M.L.: 'Inference for origin–destination matrices: estimation, prediction and reconstruction', *Transp. Res. B, Methodol.*, 2001, **35**, (7), pp. 667–676

[9] Hazelton, M.L.: 'Some comments on origin–destination matrix estimation', *Transp. Res. A, Policy Pract.*, 2003, **37**, (10), pp. 811–822

[10] Wen, T., Cai, C., Gardner, L., *et al.*: 'A strategic user equilibrium for independently distributed origin-destination demands', *Comput.-Aided Civ. Infrastruct. Eng.*, 2018, **33**, (4), pp. 316–332

[11] Willumsen, L.: 'Simplified transport models based on traffic counts', *Transportation*, 1981, **10**, (3), pp. 257–278

[12] Appiah, J.: 'Quantifying uncertainties in synthetic origin-destination trip matrix estimates', The University of Nebraska-Lincoln, 2009

[13] Bierlaire, M.: 'Mathematical models for transportation demand analysis', *Transp. Res. A, Policy Pract.*, 1997, **31**, (1), pp. 86–86

[14] Spiess, H.: 'A maximum likelihood model for estimating origin-destination matrices', *Transp. Res. B, Methodol.*, 1987, **21**, (5), pp. 395–412

[15] Kim, H., Baek, S., Lim, Y.: 'Origin-destination matrices estimated with a genetic algorithm from link traffic counts', *Transp. Res. Rec., J. Transp. Res. Board*, 2001, **1771**, pp. 156–163

[16] Tebaldi, C., West, M.: 'Bayesian inference on network traffic using link count data', *J. Am. Stat. Assoc.*, 1998, **93**, (442), pp. 557–573

[17] Li, B.: 'Bayesian inference for origin-destination matrices of transport networks using the EM algorithm', *Technometrics*, 2005, **47**, (4), pp. 399–408

[18] Fisk, C.S.: 'On combining maximum entropy trip matrix estimation with user optimal assignment', *Transp. Res. B, Methodol.*, 1988, **22**, (1), pp. 69–73

[19] Aerde, M., Rakha, H., Paramahamsan, H.: 'Estimation of origin-destination matrices: relationship between practical and theoretical considerations', *Transp. Res. Rec., J. Transp. Res. Board*, 2003, **1831**, pp. 122–130

[20] Bierlaire, M., Toint, P.L.: 'Meuse: an origin-destination matrix estimator that exploits structure', *Transp. Res. B, Methodol.*, 1995, **29**, (1), pp. 47–60

[21] Frederix, R., Viti, F., Tampère, C.M.J.: 'Dynamic origin–destination estimation in congested networks: theoretical findings and implications in practice', *Transportmetrica A, Transp. Sci.*, 2011, **9**, (6), pp. 494–513

[22] Cipriani, E., Nigro, M., Fusco, G., *et al.*: 'Effectiveness of link and path information on simultaneous adjustment of dynamic od demand matrix', *Eur. Transp. Res. Rev.*, 2014, **6**, (2), pp. 139–148

[23] Nie, Y.M., Zhang, H.M.: 'A variational inequality formulation for inferring dynamic origin–destination travel demands', *Transp. Res. B, Methodol.*, 2008, **42**, (7), pp. 635–662

[24] Sherali, H.D., Park, T.: 'Estimation of dynamic origin–destination trip tables for a general network', *Transp. Res. B, Methodol.*, 2001, **35**, (3), pp. 217–235

[25] Lo, H., Zhang, N., Lam, W.: 'Decomposition algorithm for statistical estimation of od matrix with random link choice proportions from traffic counts', *Transp. Res. B, Methodol.*, 1999, **33**, (5), pp. 369–385

[26] Wen, T., Gardner, L., Dixit, V., *et al.*: 'Two methods to calibrate the total travel demand and variability for a regional traffic network', *Comput.-Aided Civ. Infrastruct. Eng.*, 2018, **33**, (4), pp. 282–299

[27] Kattan, L., Abdulhai, B.: 'Noniterative approach to dynamic traffic origin-destination estimation with parallel evolutionary algorithms', *Transp. Res. Rec., J. Transp. Res. Board*, 2006, **1964**, pp. 201–210

[28] Baek, S., Kim, H., Lim, Y.: 'Multiple-vehicle origin-destination matrix estimation from traffic counts using genetic algorithm', *J. Transp. Eng.*, 2004, **130**, (3), pp. 339–347

[29] Wong, S., Tong, C., Wong, K., *et al.*: 'Estimation of multiclass origin-destination matrices from traffic counts', *J. Urban Plan. Dev.*, 2005, **131**, (1), pp. 19–29

[30] Foulds, L.R., do Nascimento, H.A.D., Calixto, I.C.A.C., *et al.*: 'A fuzzy set-based approach to origin–destination matrix estimation in urban traffic networks with imprecise data', *Eur. J. Oper. Res.*, 2013, **231**, (1), pp. 190–201

[31] Reddy, K.H., Chakroborty, P.: 'A fuzzy inference based assignment algorithm to estimate OD matrix from link volume counts', *Comput. Environ. Urban Syst.*, 1998, **22**, (5), pp. 409–423

[32] Gong, Z.: 'Estimating the urban OD matrix: a neural network approach', *Eur. J. Oper. Res.*, 1998, **106**, (1), pp. 108–115

[33] Cao, J., Davis, D., Vander Wiel, S., *et al.*: 'Time-varying network tomography: router link data', *J. Am. Stat. Assoc.*, 2000, **95**, (452), pp. 1063–1075

[34] Airoldi, E.M., Blocker, A.W.: 'Estimating latent processes on a network from indirect measurements', *J. Am. Stat. Assoc.*, 2013, **108**, (501), pp. 149–164

[35] Cremer, M., Keller, H.: 'A new class of dynamic methods for the identification of origin-destination flows', *Transp. Res. B, Methodol.*, 1987, **21**, (2), pp. 117–132

[36] Bell, M.G.H.: 'The estimation of an origin-destination matrix from traffic counts', *Transp. Sci.*, 1983, **17**, (2), pp. 198–217

[37] Chen, A., Chootinan, P., Recker, W.: 'Norm approximation method for handling traffic count inconsistencies in path flow estimator', *Transp. Res. B, Methodol.*, 2009, **43**, (8), pp. 852–872

[38] Sherali, H.D., Narayanan, A., Sivanandan, R.: 'Estimation of origin–destination trip-tables based on a partial set of traffic link volumes', *Transp. Res. B, Methodol.*, 2003, **37**, (9), pp. 815–836

[39] Nie, Y., Zhang, H., Recker, W.: 'Inferring origin–destination trip matrices with a decoupled GLS path flow estimator', *Transp. Res. B, Methodol.*, 2005, **39**, (6), pp. 497–518

[40] Kim, J., Kurauchi, F., Uno, N.: 'Analysis of variation in demand and performance of urban expressways using dynamic path flow estimation', *Transportmetrica*, 2009, **7**, (1), pp. 63–84

[41] Szeto, W., Jiang, Y., Sumalee, A.: 'A cell-based model for multi-class doubly stochastic dynamic traffic assignment', *Comput.-Aided Civ. Infrastruct. Eng.*, 2011, **26**, (8), pp. 595–611

[42] Brilon, W., Geistefeldt, J., Regler, M.: 'Reliability of freeway traffic flow: a stochastic concept of capacity'. Proc. of the 16th Int. Symp. on Transportation and Traffic Theory, Maryland, USA, 2005

[43] Wu, X., Michalopoulos, P., Liu, H.X.: 'Stochasticity of freeway operational capacity and chance-constrained ramp metering', *Transp. Res. C, Emerg. Technol.*, 2010, **18**, (5), pp. 741–756

[44] Uchida, T., Iida, Y.: 'Risk assignment. A new traffic assignment model considering the risk of travel time variation'. Proc. of the 12th Int. Symp. on the Theory of Traffic Flow and Transportation, Berkeley, CA, USA, 21–23 July 1993

[45] Duthie, J.C., Unnikrishnan, A., Waller, S.T.: 'Influence of demand uncertainty and correlations on traffic predictions and decisions', *Comput.-Aided Civ. Infrastruct. Eng.*, 2011, **26**, (1), pp. 16–29

[46] Dixit, V., Gardner, L.M., Waller, S.T.: 'Strategic user equilibrium assignment under trip variability'. Transportation Research Board 92nd Annual Meeting, Washington, DC, USA, 2013

[47] Wen, T., Cai, C., Gardner, L., *et al.*: 'A strategic user equilibrium for independently distributed origin-destination demands', *Comput.-Aided Civ. Infrastruct. Eng.*, 2018, **33**, (4), pp. 316–332

[48] Waller, S.T., Fajardo, D., Duell, M., *et al.*: 'Linear programming formulation for strategic dynamic traffic assignment', *Netw. Spat. Econ.*, 2013, **13**, (4), pp. 1–17

[49] Duell, M., Gardner, L., Dixit, V., *et al.*: 'Evaluation of a strategic road pricing scheme accounting for day-to-day and long term demand uncertainty'. Transportation Research Board 93rd Annual Meeting, Washington, DC, USA, 2014

[50] Mardani, M., Giannakis, G.: 'Robust network traffic estimation via sparsity and low rank'. 2013 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, 2013

[51] Chawla, S., Zheng, Y., Hu, J.: 'Inferring the root cause in road traffic anomalies'. 2012 IEEE 12th Int. Conf. on Data Mining, Brussels, Belgium, 2012

[52] Zhang, Y., Ge, Z., Greenberg, A., *et al.*: 'Network anomography'. Proc. of the 5th ACM SIGCOMM Conf. on Internet Measurement, Berkeley, CA, USA, 2005

[53] Sanandaji, B.M., Varaiya, P.P.: 'Compressive origin-destination matrix estimation', arXiv preprint arXiv:1404.3263, 2014

[54] Chen, S.S., Donoho, D.L., Saunders, M.A.: 'Atomic decomposition by basis pursuit', *SIAM Rev.*, 2001, **43**, (1), pp. 129–159

[55] Cheman, K.M.: 'Optimization techniques for solving basis pursuit problems', North Carolina State University, 2006

[56] Zou, H., Hastie, T.: 'Regularization and variable selection via the elastic net', *J. R. Stat. Soc. B, Stat. Methodol.*, 2005, **67**, (2), pp. 301–320

[57] Clark, S., Watling, D.: 'Modelling network travel time reliability under stochastic demand', *Transp. Res. B, Methodol.*, 2005, **39**, (2), pp. 119–140

[58] Tibshirani, R.: 'Regression shrinkage and selection via the Lasso', *J. R. Stat. Soc. B, Methodol.*, 1996, **58**, (1), pp. 267–288

[59] Donoho, D.L.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306

[60] Bar-Gera, H.: 'Transportation test problems of Sioux falls network', 2012. Available at http://www.bgu.ac.il/~bargera/tntp/

[61] Hurley, N., Rickard, S.: 'Comparing measures of sparsity', *IEEE Trans. Inf. Theory*, 2009, **55**, (10), pp. 4723–4741

[62] Hoyer, P.O.: 'Non-negative matrix factorization with sparseness constraints', *J. Mach. Learn. Res.*, 2004, **5**, pp. 1457–1469