ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 80 (2013) 591 – 607

Procedia
Social and Behavioral Sciences

The 20th International Symposium on Transportation and Traffic Theory

A Bayesian approach to traffic estimation in stochastic user equilibrium networks

Chong Wei[a,b*], Yasuo Asakura[b]

[a]MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing, 100044, China,
[b]Department of Civil and Environmental Engineering, Tokyo Institute of Technology,

2-12-1-M1-20, Ookayama, Meguro, Tokyo 152-8552, Japan

Abstract

This study proposes a statistical model to estimate route traffic flows in congested networks. In the study, it is assumed that route traffic flows conform to the stochastic user equilibrium (SUE) principle while being treated as random variables in order to exploit the stochastic nature of traffic. The proposed model formulates the distribution of these random variables as the conditional distribution of route flows and origin–destination (O-D) travel demand, given the observed link flows and the SUE principle. Here, the SUE principle is accounted for through the likelihood of user behaviours rather than by using a bi-level formulation. In this study, the Bayesian theorem is applied to derive the probability density function (PDF) of the conditional distribution. Based on the PDF, characteristics such as the means and variances of route/link traffic flows are estimated using a blocked Metropolis–Hastings (M–H) algorithm. To facilitate the use of prior knowledge, a hierarchical form is designed to provide a straightforward way to integrate prior knowledge into the traffic estimation model. The performance of the proposed method is tested on the Sioux–Falls network through a series of numerical examples.

Selection and peer-review under responsibility of Delft University of Technology

Keywords: Path flow estimation, Bayes' theorem, Likelihood, Stochastic User Equilibrium, Markov chain Monte Carlo;

1. Introduction

Knowing the traffic state of each link across a network is a fundamental requirement of telematics systems. However, contemporary traffic sensors such as inductive loops and Global Positioning System (GPS) devices can only cover parts of a road network, and the traffic volumes on unobserved links are not obtained. In order to improve the availability of traffic data, we need to develop a method that can estimate the traffic flows on

* Corresponding author. Tel.: +86-010-5168-7154; fax: +86-010-5168-7154.
E-mail address: chwei@bjtu.edu.cn (C.Wei)

Selection and peer-review under responsibility of Delft University of Technology
doi:10.1016/j.sbspro.2013.05.032

unobserved links.

This kind of problem, which is usually referred to as a path flow estimation problem in the literature (see [1]), is related to the Origin–Destination (OD) matrix estimation problem that has been widely discussed over the past few decades. Sasaki [2] pointed out the usefulness of an entropy model for estimating travel demands. Van Zuylen & Willumsen [3] and Van Zuylen [4] presented methods to find the most likely O–D matrix using maximum entropy and minimum information estimators. A number of previous studies also indicated that the traffic assignment model should be integrated into the O–D matrix estimators in order to take into account the dependency between route travel times and route flows in congested networks. For example, [5] developed a bi-level model in which route flows are constrained by the User Equilibrium (UE) principle; [6] proposed a method to estimate route flows in Stochastic UE (SUE) networks.

On the other hand, researchers have noted that the stochastic nature of traffic flows should not be ignored. [1] and [7] looked at the problem from a statistical perspective and suggested that a generalised least-squares estimator be used to account for the random terms of O–D demands and route traffic flows. More recently, [8] showed the benefit of combining a maximum entropy model and a least-square estimator for traffic estimation problems. Unlike its use of a generalised least-squares estimator, the approach of [9] employs a class of maximum likelihood estimators to solve the O–D estimation problem, a method that shows the potential effect of the likelihood principle. [10], [11], [12], and [13] also constructed different likelihood-based approaches to the problem. [14], [15], [16], [17], [18], and [19] formulated likelihood-based models from Bayesian perspectives. Nevertheless, almost all of these likelihood-based approaches were originally developed for uncongested networks, so there is still a need to explore the application of the likelihood principle to capture the stochastic nature of traffic flows in a congested network.

This study utilizes a Bayesian approach to solve the problem. Unlike previous studies, however, the proposed method takes into congestion effects through the likelihood of route choices on congested networks, rather than using a bi-level formulation.

We assume that traffic flow patterns conform to the SUE principle while treating route flows as random variables and seeking to estimate the characteristics of the random route flow variables. To do this, we need to obtain the probability distribution of the route flow variables. Because we consider that the route traffic flow patterns are constrained by both link traffic counts and the SUE principle, this probability distribution must be precisely represented as the conditional distribution of route traffic flows for given link traffic counts, in conformity to the SUE principle.

To exploit the conditional distribution, Bayes' theorem is adopted in this study in order to decompose the conditional distribution into a likelihood function and prior probability distributions. We first address the likelihood function and the prior probability distributions, and then combine the results to yield a formulation of the conditional distribution.

To obtain the characteristics of link flow variables, we develop a blocked Metropolis–Hastings algorithm to sample the conditional distribution of route flows, and then aggregate the samples to produce the characteristics of link flows or O–D flows.

The highlights of the proposed method are as follows:

(1) The proposed model is a likelihood-based statistical estimation model that can take into account users' contemporaneous interactions on a congested network without using bi-level formulations and can guarantee the estimate is unique; the method does not find the equilibrium solution in each iteration and does not impose specific requirements on the application of user behaviour models.
(2) The model can handle the inconsistencies among observed link traffic counts.
(3) The model can work with a hierarchical form to flexibly integrate prior knowledge.

The remainder of this article is organised as follows. Section 2 formulates the conditional probability distribution of route traffic flow variables. Section 3 outlines the blocked Metropolis–Hastings algorithm that is used to estimate the characteristics of link traffic flow variables. Section 4 provides some numerical examples that demonstrate the effectiveness of the proposed method. Section 5 offers a conclusion to the study.

2. Methodology

Let R be the set of routes; N, the set of O–D pairs; R_n, the set of routes that connect the O–D pair n; L, the set of links across the network; and L^*, the set of observable links. I_n, denotes the set of users who make trip between O–D pair n. $\mathbf{q} = [q_1, \ldots, q_{|N|}]$ is the vector of O–D demands; $\mathbf{c}_n = \{c_i | \forall i \in I_n\}$ is the set of the route choice result variables corresponding to O–D pair n; and $\mathbf{y} = [y_1, \dots, y_{|R|}]$ is the vector of the route flow variables (Note that the traffic flow on route r, y_r, is also the number of users who choose the route.); and $\mathbf{x}^* = \{x_l^* | \forall l \in L^*\}$ denotes the link traffic counts during a given time period. Other notations will be defined when they are first introduced.

2.1. Estimate traffic with pre-specified O–D demand

We begin by solving a primary estimation problem in which we aim to estimate $\mathbf{y}$ based on the following factors:
(1) The O–D demand vector $\mathbf{q} = [q_1, \dots, q_{|N|}]$ is given as a pre-condition that will not be estimated along with the other variables.
(2) A part of the links can be observed, i.e. $\mathbf{x}^*$ is available and is used to estimate **y.**
(3) We presume the same user behaviour principle as in [6] in considering that the route flows in a network conform to the SUE principle.

Clearly, $\mathbf{y}$ is constrained by factors (1) and (2) whether the network is congested or not, while $\mathbf{y}$ will be constrained to satisfy factor (3) if the network is congested. We consider $\mathbf{y} = [y_1, \dots, y_{|R|}]$ as a vector of random variables, and we intend to specify the conditional probability distribution of $\mathbf{y}$ for a given $\mathbf{x}^*$, $\mathbf{q}$, and the SUE principle. The characteristics of $\mathbf{y}$ can be obtained from the conditional distribution (note that the characteristics of the link flows can also be derived from the conditional distribution of $\mathbf{y}$).

The conditional probability distribution of $\mathbf{y}$ can be represented as $P(\mathbf{y} \mid sue, \mathbf{x}^*, \mathbf{q})$, in which sue indicates that the network is a SUE network, and $\mathbf{x}^*$ and $\mathbf{q}$ correspond to the observed link traffic counts and the pre-defined O–D demand vector, respectively.

Applying Bayes' theorem, we decompose $P(\mathbf{y} \mid sue, \mathbf{x}^*, \mathbf{q})$ as follows:

$$P(\mathbf{y} \mid sue, \mathbf{x}^*, \mathbf{q}) = P(sue, \mathbf{x}^* \mid \mathbf{y}) P(\mathbf{y}|\mathbf{q}) P(\mathbf{q}) / P(sue, \mathbf{x}^*, \mathbf{q}) \tag{1}$$

In Eq. (1), since both $P(sue, \mathbf{x}^*, \mathbf{q})$ and $P(\mathbf{q})$ are *not* dependent on the value of $\mathbf{y}$, they can be regarded as constant terms. This leads to

$$P(\mathbf{y} \mid sue, \mathbf{x}^*, \mathbf{q}) \propto P(sue, \mathbf{x}^* \mid \mathbf{y}) P(\mathbf{y}|\mathbf{q}) \tag{2}$$

On the right-hand side of Eq. (2), $P(\mathbf{y}|\mathbf{q})$ represents the prior probability distribution of $\mathbf{y}$, whose given condition $\mathbf{q}$ simply means that $\mathbf{y}$ should satisfy $\sum_{\forall r \in \mathrm{R}_n} y_r = q_n$, but does not impose any constraint on the assignment of traffic flow among the alternative routes. We will show how to specify a prior probability distribution in section 2.2.

Next, we focus on the first term in the right-hand side of Eq. (2), $P(sue, \mathbf{x}^* \mid \mathbf{y})$, which denotes the likelihood of $\mathbf{y}$ in a SUE network with the observed data $\mathbf{x}^*$, and which can be further decomposed as

$$P(sue, \mathbf{x}^* \mid \mathbf{y}) = P(\mathbf{x}^*|\mathbf{y}) P(sue \mid \mathbf{y}) \tag{3}$$

$P(\mathbf{x}^* \mid \mathbf{y})$ denotes the probability of the link traffic being observed as $\mathbf{x}^*$ given the route flow pattern $\mathbf{y}$. Given the route flow pattern $\mathbf{y}$, the corresponding traffic flow, $x_l(\mathbf{y})$, on link l can be uniquely determined. The link traffic count x_l^* can be considered as the sum between $x_l(\mathbf{y})$ and the random observation error. If we form the observation error as a normal distribution with zero mean and variance σ^2 and assume that the covariance between the links is 0, then $P(\mathbf{x}^* \mid \mathbf{y})$ will be given as

$$P(\mathbf{x}^* \mid \mathbf{y}) = \prod_{\forall l \in L^*} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[x_l(\mathbf{y}) - x_l^*]^2}{2\sigma^2}} \tag{4}$$

In this paper, we consider the variance is constant over all links. On the other hand, the variance parameter may vary with traffic [20]. A feasible way to account this issue is to estimate the variance along with other variables in the model. We leave this issue in feature works. The right-hand side of Eq. (4) implies that $P(\mathbf{x}^* \mid \mathbf{y})$ does not require consistency among the observed link traffic counts. The second term of Eq. (3), $P(sue \mid \mathbf{y})$, denotes the likelihood of route flows, which plays a role in constraining $\mathbf{y}$ to the SUE principle in the estimation model. This likelihood was partly discussed by Wei et al in [21]. Below, we provide a more comprehensive description of Eqs. (5) to (11).

Researchers have long recognized that the route flows in a SUE network should have a probability distribution rather than an exact flow pattern. This probability distribution can be stated as $P(\mathbf{y} \mid sue, \mathbf{q})$; here, the given condition sue indicates that the network is a SUE network, and $\mathbf{q}$ implies that the value of $\sum_{\forall r \in \mathrm{R}_n} y_r$ must equal q_n. Correspondingly, the likelihood of $\mathbf{y}$ in a SUE network can be represented as $P(sue \mid \mathbf{y})$. This likelihood can be further expressed as follows:

$$P(sue \mid \mathbf{y}) = \sum_{\forall \mathbf{c}|\mathbf{y}} P(sue|\mathbf{c}) P(\mathbf{c}|\mathbf{y}) \tag{5}$$

where $\mathbf{c} = [\mathbf{c}_1, \dots, \mathbf{c}_{|N|}]$ and $\mathbf{c}|\mathbf{y}$ denotes a feasible pattern of $\mathbf{c}$ that satisfies:

$$y_r = \sum_{\forall i \in I_n} \delta(r, c_i) \quad \forall r \in R_n, \forall n \in N \tag{6}$$

where $\delta(r, c_i)$ is 1 if user i chooses route r; otherwise $\delta(r, c_i)$ is 0. The second term in the right hand side of Eq. (5), $P(\mathbf{c}|\mathbf{y})$ is the probability of $\mathbf{c}$ for a given $\mathbf{y}$. We have $\sum_{\forall \mathbf{c}|\mathbf{y}} P(\mathbf{c}|\mathbf{y}) = 1$.

$P(sue|\mathbf{c})$ in Eq. (5) refers to the likelihood of $\mathbf{c}$ in a SUE network. Previous studies ([21], [22]) have shown that sue is equivalent to $sub_i \ \forall i$, i.e.

$$sue \Leftrightarrow sub_i \ \forall i \tag{7}$$

where sue, again, indicates that the network is in SUE, sub_i denotes that user i displays stochastic user behaviour (SUB), i.e. user i selects the route that is perceived to have the maximum utility, and $sub_i \ \forall i$ denotes that all the users in the network display SUB. [21] indicated that sub_i can also be equivalently expressed as

$$sub_i \Leftrightarrow V(c_i) + \theta(c_i) > V(h) + \theta(h) \ \forall h \neq c_i, h \sim c_i \tag{8}$$

where $V(\cdot)$ is the deterministic utility of a route, $h \sim c_i$ denotes that route h connects the same O–D pair as route c_i (c.f. [22]), and $\theta(\cdot)$ is a random term; here, we consider the probability distribution of the random variable to be the researcher's subjective probability (see [23]). Thus, [21] formulated the likelihood of $\mathbf{c}$ in a SUE network for homogeneous users as follows:

$$P(sue \mid \mathbf{c}) = P(sub_i \ \forall i \mid \mathbf{c}) = \prod_{\forall i \in I} P(sub_i|\mathbf{c},) = \prod_{\forall r \in R} p(r|\mathbf{y}(\mathbf{c}))^{y_r(\mathbf{c})} \tag{9}$$

where $y_r(\mathbf{c})$ denotes the traffic flow on route r that is determined from $\mathbf{c}$ using Eq. (6), $\mathbf{y}(\mathbf{c})$ is the vector of $y_r(\mathbf{c})$, and $p(r|\mathbf{y}(\mathbf{c}))$ is defined as

$$p(r|\mathbf{y}(\mathbf{c})) = P(V(r) + \theta(r) > V(h) + \theta(h) \ \forall h \neq r, h \sim r|\mathbf{y}(\mathbf{c})) \tag{10}$$

In essence, the value of $p(r|\mathbf{y}(\mathbf{c}))$ is equal to the choice probability of route r, whose value can be calculated using any route choice model that adheres to the random utility theory. We consider that the cost of a route appearing in the utility function of the route is dependent on $\mathbf{y}(\mathbf{c})$.

Next, we provide a toy example to illustrate Eqs. (9) and (10). Fig. 1 describes a network with a single O–D pair and consists of two routes (routes A and B). Five users ($i = 1, \dots, 5$) move from node O to node D, in which three users ($i = 1,2,3$) choose route A, and two users choose route B ($i = 4,5$). According to the notation defined earlier, the route choice results are represented as $c_1 = \mathrm{A}, c_2 = \mathrm{A}, c_3 = \mathrm{A}, c_4 = \mathrm{B}, \text{and } c_5 = \mathrm{B}$. From the route choice results, the traffic flow on each route is determined by Eq. (6) as $y_{\mathrm{A}}(\mathbf{c}) = 3, \text{and } y_{\mathrm{B}}(\mathbf{c}) = 2$ (i.e. $\mathbf{y}(\mathbf{c}) = [3,2]$). Note that the route choice pattern $\mathbf{c}$ can determine a unique route flow pattern; on the contrary a route flow pattern may correspond to different route choice patterns. If we use $d_r = 1 + y_r^2$ as the travel time function, then the travel times of routes A and B would be $d_{\mathrm{A}} = 10$ and $d_{\mathrm{B}} = 5$, respectively. If we let $V(r) = -0.1 \cdot d_r$ and let $\theta(r)$ in Eq. (10) have an extreme value type I distribution, then the value of $p(r|\mathbf{y}(\mathbf{c}))$ for routes A and B can be calculated using the Logit model as $p(\mathrm{A}|\mathbf{y}(\mathbf{c})) = 0.38$ and $p(\mathrm{B}|\mathbf{y}(\mathbf{c})) = 0.62$. The value of the likelihood of $\mathbf{c}$ can be calculated using Eq. (9) as $P(sue \mid \mathbf{c}) = 0.02$.

Combining Eqs. (5) and (9), we can obtain $P(sue|\mathbf{y})$ as

$$P(sue\ |\mathbf{y}) = \sum_{\forall \mathbf{c}|\mathbf{y}} \prod_{\forall r \in R} p(r|\mathbf{y}(\mathbf{c}))^{y_r(\mathbf{c})}\, P(\mathbf{c}|\mathbf{y}) = \prod_{\forall r \in R} p(r|\mathbf{y})^{y_r} \tag{11}$$

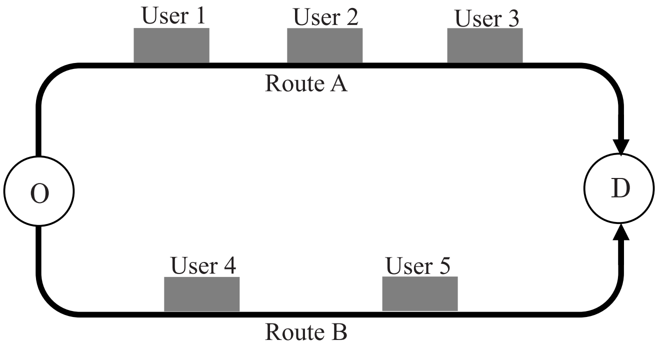

Fig. 1. A toy example for Eq. (9) and (10)

2.2. Estimation of route flows with endogenous O–D demand

In this section we estimate $\mathbf{y}$ based on the following factors:

(1) Each O-D pair corresponds to a set of potential travellers. We assume the set is unchanged. Correspondingly, I_n can be considered as a subset of the potential travellers of O-D pair n, and q_n can be considered as the size of I_n.

(2) $\mathbf{x}^*$ is available and is used to estimate $\mathbf{y}$.

(3) The route flows in a network conform to the SUE principle.

Now $\mathbf{q}$ is no longer pre-specified. This leads us to represent the traffic estimation problem as $P(\mathbf{y}\ |sue, \mathbf{x}^*)$, i.e. the conditional probability distribution of $\mathbf{y}$ for a given $\mathbf{x}^*$ and sue. Applying the Bayesian theorem, we can decompose $P(\mathbf{y}|sue, \mathbf{x}^*)$ as follows:

$$P(\mathbf{y}|sue, \mathbf{x}^*) = P(sue, \mathbf{x}^*\ |\ \mathbf{y})P(\mathbf{y})/P(sue, \mathbf{x}^*) \tag{12}$$

Eq. (12) is different from Eq. (1) because $\mathbf{q}$ is no longer a given condition. The denominator of the right-hand side of Eq. (12) can be derived as

$$P(sue, \mathbf{x}^*) = \sum_{\forall \mathbf{y}} P(sue, \mathbf{x}^*|\ \mathbf{y})P(\mathbf{y}) \tag{13}$$

Eq. (13) implies that the value of $P(sue, \mathbf{x}^*)$ is not independent of $\mathbf{y}$, and this yields

$$P(\mathbf{y}|sue, \mathbf{x}^*) \propto P(sue, \mathbf{x}^*\ |\ \mathbf{y})P(\mathbf{y}) \tag{14}$$

where $P(sue, \mathbf{x}^*\ |\ \mathbf{y})$ have been specified in Section 2.1, and $P(\mathbf{y})$ is the prior distribution of $\mathbf{y}$.

As the name implies, the prior distribution will be specified from researchers' subjective perspectives. In this study we specify this prior distribution based on the following considerations: we consider that the conditional distribution of $\mathbf{q}$ given sue should approximate to a uniform distribution. This is because the given condition sue just constrains the assignment of traffic flow among the alternative routes but does not impose any constrain on $\mathbf{q}$.

On the other hand, we recognized that the general form of $P(\mathbf{q}|sue)$ is rather complex. To avoid adding too much complexity into the model, we prefer to specify $P(\mathbf{y})$ through looking at $P(\mathbf{q}|sue)$ in a special case in which link capacity is unlimited. To account for the consideration, we specify $P(\mathbf{y})$ as follows:

$$\begin{aligned} P(\mathbf{y}) &= \prod_{\forall n\in N} \sum_{\forall \mathbf{c}_n|\mathbf{y}_n} P(\mathbf{c}_n) \\ &= \prod_{\forall n\in N} \sum_{\forall \mathbf{c}_n|\mathbf{y}_n} \left[\left(\frac{w_n!}{(w_n-q_n)!q_n!}\right)^{-1}\cdot\eta\right] \\ &= \prod_{\forall n\in N} \left[\left(\frac{w_n!}{(w_n-q_n)!q_n!}\right)^{-1}\cdot\eta\right]\cdot\left(\sum_{\forall \mathbf{c}_n|\mathbf{y}_n} 1\right) \\ &= \prod_{\forall n\in N} \left[\left(\frac{w_n!}{(w_n-q_n)!q_n!}\right)^{-1}\cdot\eta\right]\cdot\left(\frac{q_n!}{\prod_{\forall r\in R_n} y_r!}\cdot\frac{w_n!}{(w_n-q_n)!q_n!}\right) \\ &= \prod_{\forall n\in N} \left[\frac{q_n!}{\prod_{\forall r\in R_n} y_r!}\cdot\eta\right] \end{aligned} \tag{15}$$

where $\mathbf{c}_n|\mathbf{y}_n$ is used to denote a feasible pattern of $\mathbf{c}_n$ given $\mathbf{y}_n$, η is used to denote a constant that is independent on $\mathbf{y}$, q_n is equal to $\sum_{\forall r\in R_n}\sum_{\forall i\in I_n}\delta(r,c_i)$, and w_n is used to denote the size of the set of the potential travellers corresponding to O-D pair n.

In Eq. (15), $P(\mathbf{c}_n)$ denotes the prior distribution of $\mathbf{c}_n$. Since we consider the conditional distribution of $\mathbf{q}$ given sue approximates to a uniform distribution and has no other information regarding $\mathbf{c}_n$, it is reasonable to consider each route choice pattern that satisfies $\sum_{\forall r\in R_n}\sum_{\forall i\in I_n}\delta(r,c_i)=q_n$ has the same probability of occurring on the basis of the principle of indifference.

As shown in the second line of Eq. (15), $P(\mathbf{c}_n)$ is specified as $\left[w_n!/\left((w_n-q_n)!\,q_n!\right)\right]^{-1}\cdot\eta$. In the third line of Eq. (15), $\sum_{\forall \mathbf{c}_n|\mathbf{y}_n} 1$ is equal to the number of feasible patterns of $\mathbf{c}_n$ given $\mathbf{y}_n$. In the same manner as Eq. (14), we can derive $P(\mathbf{y}|sue)$ from Eqs. (11) and (15) as:

$$P(\mathbf{y}|sue) \propto P(sue\mid\mathbf{y})P(\mathbf{y}) = \prod_{\forall n\in N}\left[\frac{q_n!}{\prod_{\forall r\in R_n} y_r!}\cdot\prod_{\forall r\in R_n} p(r|\mathbf{y})^{y_r}\cdot\eta\right] \tag{16}$$

In the special case considered above, $P(\mathbf{y}|sue)$ will be proportional to a multinomial density (see [24]); this leads $P(\mathbf{q}|sue)$ to a uniform distribution. Eq. (16) also implies that w_n is not involved in $P(\mathbf{q}|sue)$ or the final result of the model. The numerical examples provided in section 4 confirm this specification of the prior distribution can work well in a congested network with limited link capacity.

2.3. Prior knowledge

In practice, one may have prior knowledge about the travel demand from sources such as an outdated O–D matrix. As an example of this, [6] and [25] showed that the relative magnitudes of the elements of an O–D matrix is a kind of useful prior knowledge to improve the estimation performance. In this section we present a flexible way to integrate such prior knowledge by applying a hierarchical form.

Let $\mathbf{b}=[b_1,\dots,b_{|N|}]$, where b_n are the relative magnitudes of the demand of O–D pair n in the total demand across the network. Without any loss of generality, we next demonstrate how to introduce $\mathbf{b}$ into the estimator.

Because $\mathbf{b}$ served as prior knowledge, the route flow vector $\mathbf{y}$ may not absolutely satisfy the relative magnitudes given by $\mathbf{b}$. It is therefore reasonable to establish the stochastic relationship between $\mathbf{b}$ and $\mathbf{y}$ as a Dirichlet distribution:

$$P(\mathbf{b}|\mathbf{y}) = \frac{\Gamma(\sum_{\forall n\in N} q_n)}{\prod_{\forall n\in N}\Gamma(q_n)}\prod_{\forall n\in N} {b_n}^{q_n-1} \tag{17}$$

where $\Gamma(\cdot)$ denotes a gamma function. $P(\mathbf{b}|\mathbf{y})$ has a hierarchical structure, in which the elements of the given condition $\mathbf{y}$ are not constants, but rather, follow the probability distribution provided by Eq. (14). Next we introduce $\mathbf{b}$ into $P(\mathbf{y}|sue,\mathbf{x}^*)$ as a given condition using $P(\mathbf{b}|\mathbf{y})$, and obtain

$$\begin{aligned} P(\mathbf{y}|sue,\mathbf{x}^*,\mathbf{b}) &= P(\mathbf{b}|\mathbf{y})P(sue,\mathbf{x}^*\mid\mathbf{y})P(\mathbf{y})/P(sue,\mathbf{x}^*,\mathbf{b}) \\ &\propto P(\mathbf{b}|\mathbf{y})P(sue,\mathbf{x}^*\mid\mathbf{y})P(\mathbf{y}) \end{aligned} \tag{18}$$

2.4. Summary

Sections 2.1 through 2.3 have shown how to derive the conditional distribution of $\mathbf{y}$ step by step. In this section, we provide a summary of the results obtained.

Combining Eqs. (3), (14), and (15), we can state the conditional distribution of $\mathbf{y}$ without prior knowledge as

$$P(\mathbf{y}|sue,\mathbf{x}^*) \propto P(sue,\mathbf{x}^* \mid \mathbf{y})P(\mathbf{y}) \propto \prod_{\forall l\in L^*}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{[x_l(\mathbf{y})-x_l^*]^2}{2\sigma^2}} \cdot \prod_{\forall r\in R} p(r|\mathbf{y})^{y_r} \cdot \prod_{\forall n\in N}\frac{q_n!}{\prod_{\forall r\in R_n} y_r!} \tag{19}$$

Since the value of η is a constant and is independent of $\mathbf{y}$, we remove η from the equation in order to simplify the expression.

On combining Eqs. (3), (15), (17) and (18), we can express the conditional distribution of $\mathbf{y}$ with the prior knowledge $\mathbf{b}$ as

$$P(\mathbf{y}|sue,\mathbf{x}^*,\mathbf{b}) \propto P(\mathbf{b}|\mathbf{y})P(sue,\mathbf{x}^* |\mathbf{y})P(\mathbf{y}) \tag{20}$$
$$\propto \frac{\Gamma(\sum_{\forall n\in N} q_n)}{\prod_{\forall n\in N}\Gamma(q_n)}\prod_{\forall n\in N} {b_n}^{q_n-1} \cdot \prod_{\forall l\in L^*}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{[x_l(\mathbf{y})-x_l^*]^2}{2\sigma^2}} \cdot \prod_{\forall r\in R} p(r|\mathbf{y})^{y_r} \cdot \prod_{\forall n\in N}\frac{q_n!}{\prod_{\forall r\in R_n} y_r!}$$

3. Solution algorithms

The characteristics of $\mathbf{y}$, such as its mean and variance, may be calculated by integrating over $\mathbf{y}$ in $P(\mathbf{y} \mid sue,\mathbf{x}^*)$. However, this method is usually not feasible owing to the high computational overhead involved. Practically, we may use a sampling algorithm to sample the conditional distribution of $\mathbf{y}$, and then ascertain the characteristics of the route/link traffic flows by aggregating the samples of $\mathbf{y}$ (e.g. the sample mean of $\mathbf{y}$ can be used as the estimated mean of $\mathbf{y}$).

In this section, we describe a sampling algorithm that is used to draw samples of $\mathbf{y}$ from the conditional distribution, $P(\mathbf{y} \mid sue,\mathbf{x}^*)$. The sampling algorithm that is designed to sample $P(\mathbf{y} \mid sue,\mathbf{x}^*,\mathbf{b})$ is presented in Appendix A.

Our sampling method uses a blocked Metropolis–Hastings (M-H) algorithm. We use $\mathbf{y}_n$ to denote $y_r\ r\in R_n$ i.e. the route flow vector corresponding to O–D pair n, so that the posterior probability distribution of $\mathbf{y}$ can also be represented as

$$P(\mathbf{y} \mid sue,\mathbf{x}^*) = P\big(\mathbf{y}_1,\dots,\mathbf{y}_{|N|} \mid sue,\mathbf{x}^*\big) \tag{21}$$

Let $\mathbf{y}_n^t$ be the tth samples of $\mathbf{y}_n$. The blocked M–H algorithm draws samples using the following process:

Sampling Scheme A

(0) Specify initial samples $\mathbf{y}_1^0,\dots,\mathbf{y}_{|N|}^0$ for $\mathbf{y}_1,\dots,\mathbf{y}_{|N|}$, set $t\leftarrow 1$ and $n\leftarrow 1$.

(1) For the O–D pair n:
    Execute sampling scheme B to draw $\mathbf{y}_n^t$ using the M–H algorithm.

(2) If $n<|N|$ then $n\leftarrow n+1$, and go to step (1); otherwise, go to step (3).

(3) If $t<T$ then $t\leftarrow t+1$, $n\leftarrow 1$, and go to step (1); otherwise, stop the iteration.

Sampling scheme A implies that the algorithm does not draw samples for all the variables simultaneously. Instead, scheme A samples $\mathbf{y}_n^t$ in turn (for $n=1,\dots,|N|$) at the tth iteration. This strategy makes it possible to avoid directly drawing samples from a high-dimensional distribution, and is more useful for real-world problems (see [20]).

Step (1) of scheme A employs the M–H algorithm to sample $\mathbf{y}_n^t$ from $P(\mathbf{y} \mid sue,\mathbf{x}^*)$ for O–D pair n. The M–H algorithm ([26], [27]) is a Markov chain Monte Carlo (MCMC) method that first treats the probability distribution that we wish to sample as a stationary distribution, and then draws samples based on the construction of a Markov chain that has a stationary distribution. The two major steps involved in MCMC methods are (1) draw a candidate sample from the proposal distribution that is an arbitrary distribution and that can be selected freely, though the support of the proposal distribution needs to include the support of the given distribution, (2) determine whether the candidate sample can be accepted as a sample from the distribution that is to be sampled. The sampling scheme B presented below is designed on the basis of the M–H algorithm in order to draw $\mathbf{y}_n^t$.

Sampling Scheme B

(0) Draw candidate samples $\boldsymbol{y}_n'$:

For route r, draw the candidate sample y_r' from the proposal distribution $\theta_r(g_r|y_r^{t-1})$, in which g_r denotes a random variable, y_r' is a random sample of g_r, and $\theta_r(g_r|y_r^{t-1})$ denotes the conditional probability distribution of g_r given y_r^{t-1} (the PDF of the proposal distribution is given by Eq. 22.).

(1) Let $q_n' = \sum_{\forall r \in R_n} y_r'$.

(2) Calculate τ as

$$\tau = \frac{P(\dots, \mathbf{y}_n', \dots | sue, \mathbf{x}^*)\theta_r(y_r^{t-1}|y_r')}{P(\dots, \mathbf{y}_n^{t-1}, \dots | sue, \mathbf{x}^*)\theta_r(y_r'|y_r^{t-1})}$$

where

$$P(\dots, \mathbf{y}_n', \dots | sue, \mathbf{x}^*) = P\left(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n', \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1} | \, sue, \mathbf{x}^*\right)$$
$$\propto P(sue, \mathbf{x}^* \mid \mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n', \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1}) P(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n', \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1})$$

and

$$P(\dots, \mathbf{y}_n^{t-1}, \dots | sue, \mathbf{x}^*) = P\left(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n^{t-1}, \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1} | \, sue, \mathbf{x}^*\right)$$
$$\propto P(sue, \mathbf{x}^* \mid \mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n^{t-1}, \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1}) P(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n^{t-1}, \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1})$$

To calculate the value of $\theta_r(y_r^{t-1}|y_r')$, we only need to replace g_r and y_r^{t-1} with y_r^{t-1} and y_r', respectively in Eq. (22), and then calculate the value of the equation; the value of $\theta_r(y_r'|y_r^{t-1})$ is calculated in the same manner.

(3) Accept $\mathbf{y}_n'$ as $\mathbf{y}_n^t = \mathbf{y}_n'$ with a probability of $\min(1, \tau)$; otherwise, accept them as $\mathbf{y}_n^t = \mathbf{y}_n^{t-1}$.

Note that the normalizing constant of $P(\dots, \mathbf{y}_n', \dots | sue, \mathbf{x}^*)$ (or $P(\dots, \mathbf{y}_n^{t-1}, \dots | sue, \mathbf{x}^*)$) does not need to be known for calculating the value of τ.

We choose $\theta_r(g_r|y_r^{t-1})$ as a binomial distribution whose PDF is given as

$$\theta_r(g_r|y_r^{t-1}) = \frac{m_r!}{g_r!(m_r-g_r)!} \left(\frac{y_r^{t-1}}{m_r}\right)^{g_r} \left(\frac{m_r-y_r^{t-1}}{m_r}\right)^{m_r-g_r} \tag{22}$$

where m_r is a parameter of the binomial distribution, which should be specified as a value that can cover the range of the traffic flow on route r. Since $\theta_r(g_r|y_r^{t-1})$ is a binomial distribution, it is easy to sample $\theta_r(g_r|y_r^{t-1})$ using programming toolkits such as C++ TR1. One benefit of the M–H algorithm is that it can work based only on the joint distribution of $\mathbf{y}_1, \dots, \mathbf{y}_{|N|}$ and does not require further derivation of the conditional distribution of $\mathbf{y}_n$ given all the other elements.

The sampling scheme A calls the sampling scheme B at step (1) in order to draw $\mathbf{y}_n^t$ from the conditional distributions of $\mathbf{y}_n$. The combination of schemes A and B is usually referred to as a blocked M–H sampler.

Along the same lines as the sampling algorithm described above, scheme C is designed to sample $P(\mathbf{y}|sue, \mathbf{x}^*, \mathbf{b})$, and is presented in Appendix A.

4. Numerical example

4.1. Testing of the proposed model on the Sioux–Falls network

In this section, we test the performance of the proposed method on the Sioux–Falls network (Fig. 2). The network consists of 24 nodes and 76 links. A set of 60 non-zero O–D pairs is considered in this test case. The feasible routes between each O–D pair are predefined using the method described by [28], which follows the principle of Dial's algorithm [29] (We coded the program using MATLAB, which takes a few seconds to generate routes for the 60 O–D pairs in the Sioux–Falls network.). To focus on the basic idea, the route sets are considered to remain constant over iterations. We calculate the travel time d_l on link l, using the Bureau of Public Road (BPR) function:

$$d_l = d_l^0 \cdot (1 + 0.15 \cdot (x_l/k_l)^4) \tag{23}$$

where d_l^0 is the free-flow travel time on link l, and k_l is the capacity of link l whose average is 1135 vehicles per time period.

We use a SUE traffic assignment model to generate a set of "true" link flows. The SUE model calculates the choice probability of route r that connects O–D pair n using a Logit route choice model as follows:

$$p(r|\mathbf{y}) = \frac{\exp(-\beta z_r(\mathbf{y}))}{\sum_{\forall h \in R_n} \exp(-\beta z_h(\mathbf{y}))} \tag{24}$$

where $z_r(\mathbf{y})$ is the cost of route r that is determined by $\mathbf{y}$, and β is the coefficient of the Logit model. The value of β is set to 2. The same Logit model is also used in the estimation model, but we would note that the proposed model does not impose specific requirements on the application of user behaviour models.

As shown in Fig. 2, a set of 23 links (i.e. about 30% of the links) is randomly selected to be the observed links. The "true" flows on these links are given in Table 1. Note that the "observed" flow on link l, x_l^* may be different from the "true" flow, $x_l^\#$ due to observational errors, so that inconsistencies can arise in the "observed" link flows (see [30]). For illustrative purposes, we created the "observed" flow, x_l^* by drawing a sample from the Poisson distribution as $x_l^* \sim \text{Poisson}(\lceil x_l^\# \rceil)$ (see [31]), where $\lceil x_l^\# \rceil$ denotes the nearest integer greater than or equal to $x_l^\#$.

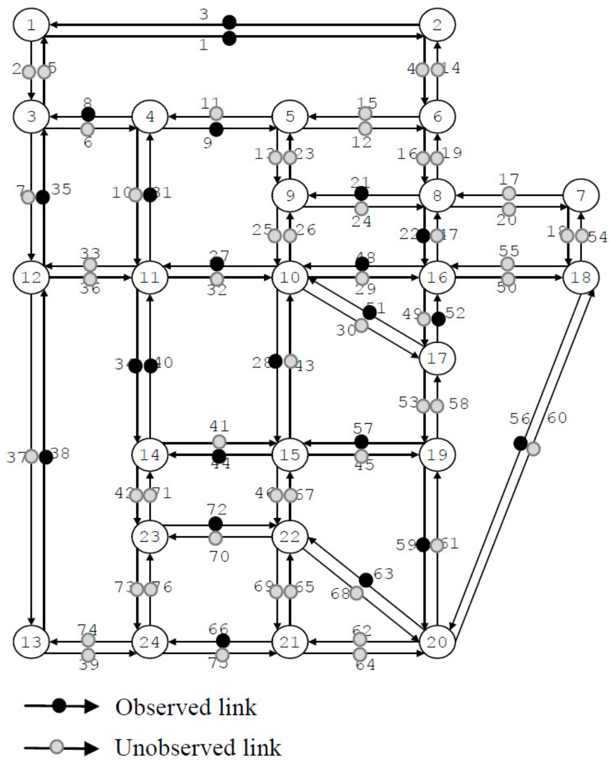

Fig. 2. Observed and unobserved links in the Sioux–Falls network

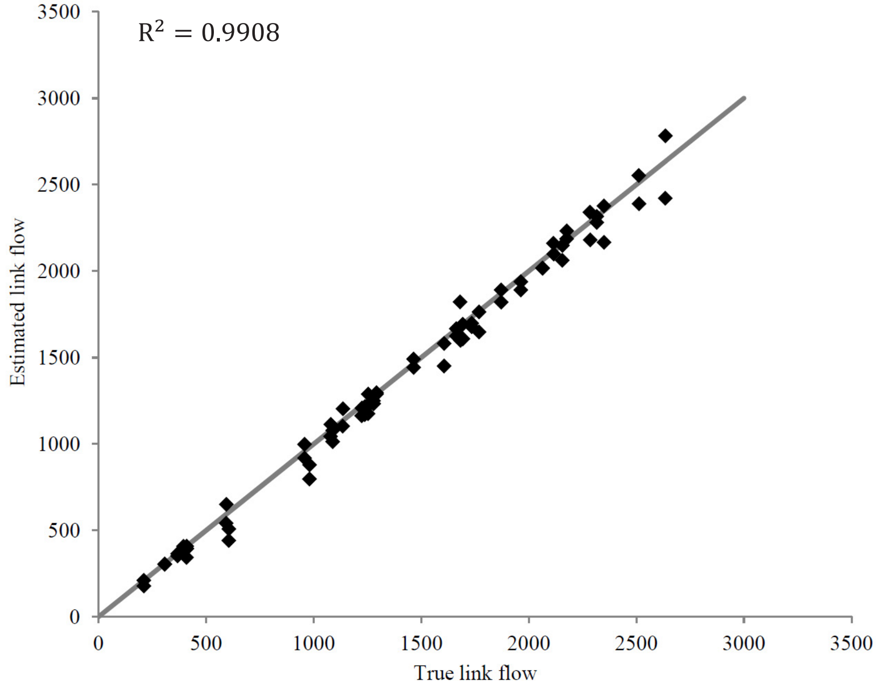

Fig. 3. True link flows versus link flow estimates without prior knowledge

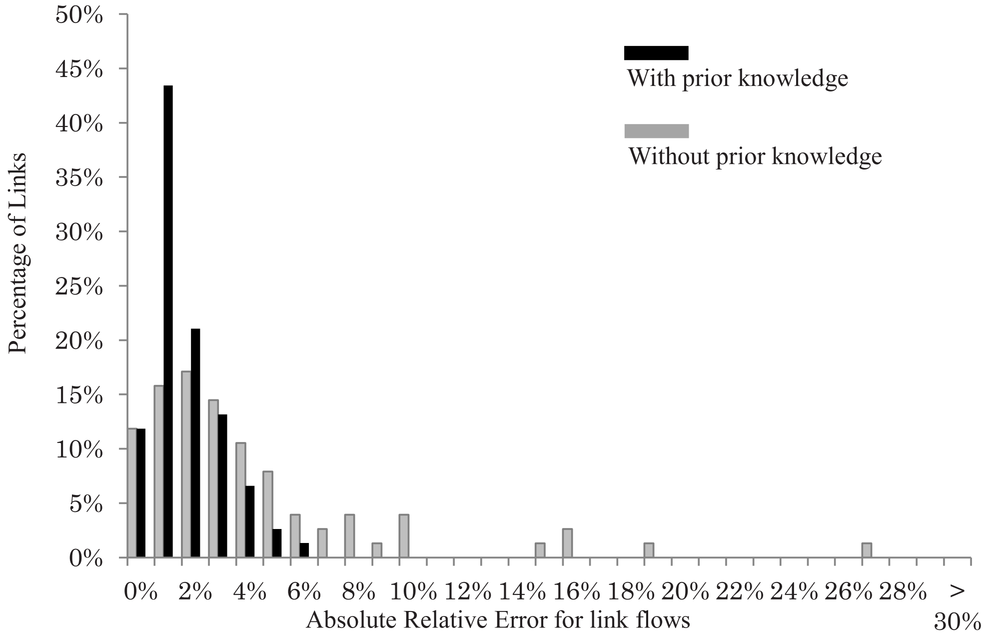

Fig. 4. Distribution of AREs for estimates with and without prior knowledge

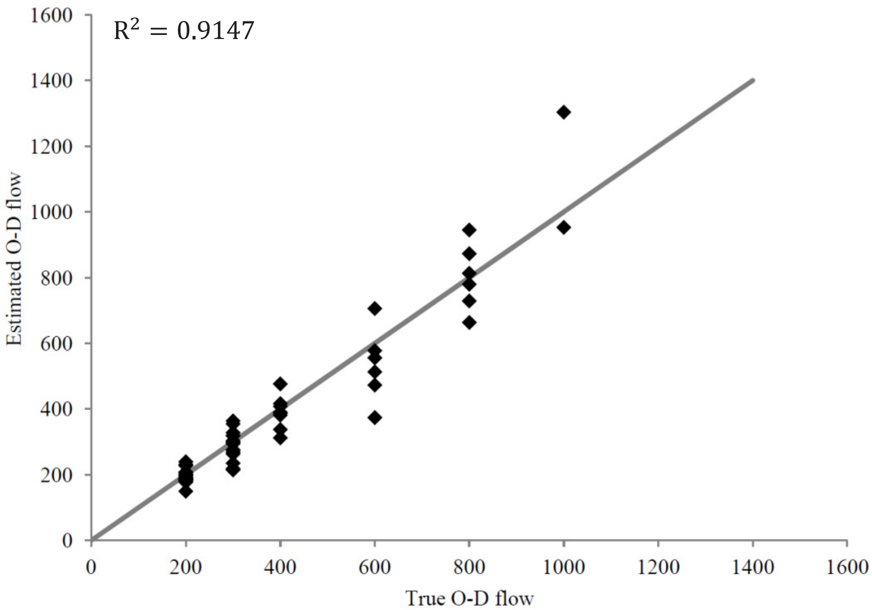

Fig. 5. True O–D flows versus O–D flow estimates without prior knowledge

We use the absolute relative error (ARE) to evaluate the accuracy of the estimates produced by the proposed model. This measure is defined below:

$$\mathrm{ARE}_l = \frac{|\bar{x}_l - x_l^{\#}|}{x_l^{\#}} \cdot 100\% \tag{25}$$

where $\bar{x}_l$ denotes the sample mean of x_l, which is calculated using the samples simulated by the proposed sampling algorithm.

4.2. Traffic estimation without prior knowledge

We first conducted traffic estimation without prior knowledge of traffic patterns. The estimation is based on the model summarized by Eq. (19).

To run the sampling algorithm, we set the initial traffic flow on each route to 40 (The true traffic flows in the network range from 79 to 118. We also investigated the performance of the proposed model when the initial traffic flow on each route was set to 20 and we find the performance is independent on the change of the initial traffic flow.). We drew 50,000 samples from the probability distribution, $P(\mathbf{y} \mid sue, \mathbf{x}^*)$ using sampling scheme A and then produced the sample means of $\mathbf{y}$ from the samples. We monitored the acceptance rate for candidate sample and found the rate is around 58%.

Fig. 3 gives an outline of the flow estimates of all the links. The horizontal axis of the figure denotes the "true" link flows, while the vertical axis denotes the sample means of the link flows, which are calculated using the samples. The straight line in the figure represents the 45° line.

The histogram shown in Fig. 4 presents the estimated results in more detail. The horizontal axis of the figure represents the values of the ARE. Fig. 4 shows that about 12% of the links have a corresponding value of ARE that is in a range of 0% to 1%, and about 17% of the links have a corresponding value in the range of 2% to 3%. It can be seen that 70% of the links have an ARE that is less than 5%.

Table 1 lists the estimated results obtained without prior knowledge of the observed links. We that find link 1 has the largest ARE (15.04%). On the other hand, we also note that the "observed" flow on link 1, x_1^*, is very different from the "true" flow, $x_1^{\#}$ due to observational errors. Although the sample mean of the flow on link 1 is extremely close to the "observed" flow, x_1^*, this estimate is far from the "true" flow, $x_1^{\#}$. In Section 4.3 we show how prior

Table 1. Estimation results for the observed links

Link	True flow	Observed flow	without prior knowledge		with prior knowledge	
			Estimated	ARE	Estimated flow	ARE
1	210.685	179	179.002	15.04%	208.042	1.25%
3	210.403	207	211.09	0.33%	212.42	0.96%
8	1253.39	1292	1288.82	2.83%	1271.2	1.42%
9	2316.02	2283	2282.09	1.47%	2300.67	0.66%
21	593.879	661	650.674	9.56%	615.811	3.69%
22	1871.37	1890	1891.23	1.06%	1903.18	1.70%
27	1962.69	1887	1890.96	3.65%	1952.76	0.51%
28	2064.01	2011	2016.7	2.29%	2047.31	0.81%
31	1663.24	1622	1624.98	2.30%	1627.48	2.15%
34	2114.33	2167	2161.45	2.23%	2136.75	1.06%
35	1258.61	1235	1228.51	2.39%	1237.37	1.69%
38	1677.88	1625	1625.71	3.11%	1629.27	2.90%
40	2115.6	2103	2099.24	0.77%	2101.78	0.65%
44	2175.68	2187	2187.2	0.53%	2172.97	0.12%
48	2511.19	2558	2553.2	1.67%	2524.63	0.54%
51	1079.61	1111	1113.19	3.11%	1078.85	0.07%
52	2284.24	2344	2340.98	2.48%	2311.07	1.17%
56	2156.3	2154	2149.26	0.33%	2146.3	0.46%
57	1692.66	1695	1692.67	0.00%	1687.28	0.32%
59	307.818	306	304.697	1.01%	301.015	2.21%
63	1224.35	1206	1207.92	1.34%	1215.01	0.76%
66	1088.33	1080	1078.23	0.93%	1079.93	0.77%
72	1134.92	1103	1103.61	2.76%	1118	1.49%
			Average	2.66%	Average	1.19%

knowledge can improve the accuracy of the estimate of link 1.

Table 2 provides a summary for the link flow estimates of all the links. For the case without prior knowledge, the average ARE is 4.24%, and the max ARE is 26.99%.

The samples of $\mathbf{y}$ from $P(\mathbf{y} \mid sue, \mathbf{x}^*)$ can be aggregated to yield the sample mean of $\mathbf{q}$. Fig. 5 presents the estimates of $\mathbf{q}$. The horizontal axis and vertical axis represent the “true” and the sample means of the O–D demands, respectively. Table 2 shows that the average and the max AREs for the O-D estimates are 10.46% and 37.65%, respectively. The results imply that the accuracy of the O–D demand estimates is not as good as that of the link flow estimates. However, we should bear in mind that this estimation was carried out without prior knowledge, while only 30% of links in the network could be observed.

4.3. Traffic estimation using prior knowledge

As described in Section 2.3, a hierarchical form can be employed to integrate the prior knowledge $\mathbf{b}$ into the estimation model. As mentioned in that section, the route flow vector $\mathbf{y}$ may not absolutely satisfy the relative magnitudes given by $\mathbf{b}$. This leads us to formulate the stochastic relationship between $\mathbf{b}$ and $\mathbf{y}$ as a Dirichlet distribution. In this test scenario, we created $\mathbf{b}$ by introducing Poisson-perturbed errors to the true O–D matrix (see Section 4.1).

We drew 50,000 samples from the probability distribution, $P(\mathbf{y} \mid sue, \mathbf{x}^*, \mathbf{b})$ using sampling scheme C (see Appendix A). The acceptance rate for candidate sample is around 52%.

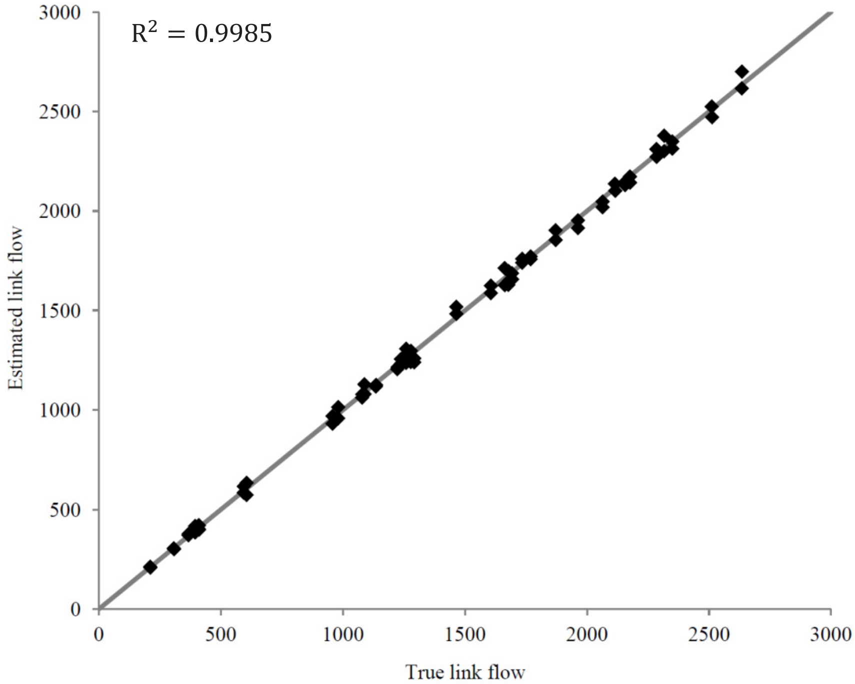

Fig. 6. True link flows versus link flow estimates with prior knowledge

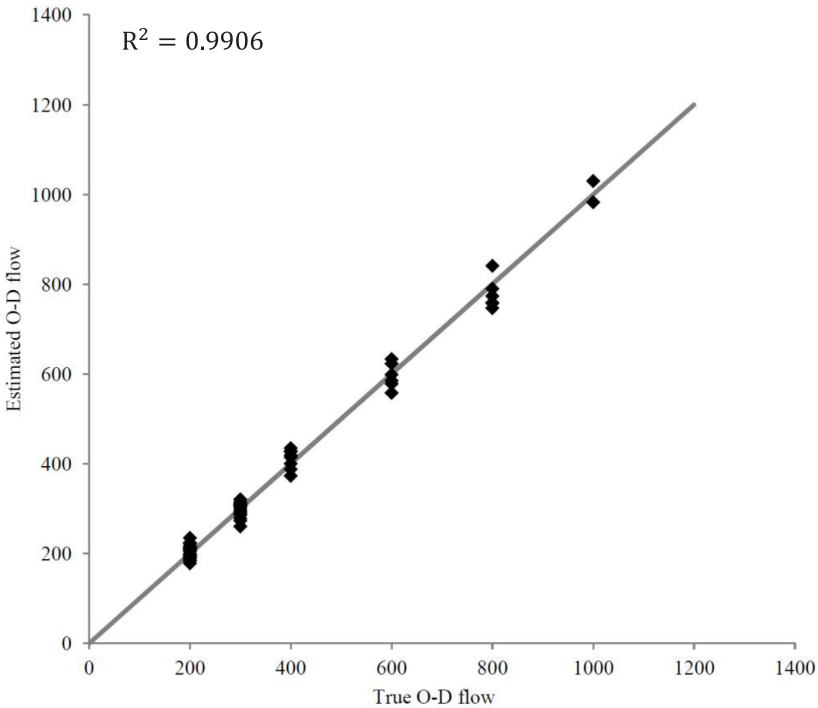

Fig. 7. True O–D flows versus O–D flow estimates with prior knowledge

Table 2. Summary of the estimation results

Prior knowledge	Link flow		O-D flow	
	Average ARE	Max ARE	Average ARE	Max ARE
Without	4.24%	26.99%	10.46%	37.65%
With	1.74%	5.85%	4.97%	17.41%

Similar to Fig. 3, Fig. 6 gives an outline of the link flow estimate results with prior knowledge. One can see that almost all the data points are closely aligned with the 45° line.

Fig. 4 also shows the distribution of the values of ARE for the estimated results with prior knowledge. We can observe that the peak of the histogram shifts to the left significantly, which reflects the fact that more than half the links (about 55%) have an ARE of less than 2%.

The set of the perturbed observed link flows described in Section 4.1 is used in this test scenario as well. Table 1 shows the estimated results of these observed links with prior knowledge. Compared to the case without prior knowledge, one can observe that the estimated error of link 1 is reduced considerably. This result implies that the use of appropriate prior knowledge with the hierarchical form can relax the negative effect of observation errors.

Fig. 7 plots the true O–D demand versus the sample mean O–D demand. As expected, the prior knowledge can significantly improve the accuracy of the O–D demand estimates. As summarized in Table 2, the average and the max AREs for the O–D demand estimates are 4.97% and 17.41%, respectively, while these measurements in the case without prior knowledge are 10.46% and 37.65%, respectively.

Table 2 also compares the estimated link flows without prior knowledge with those obtained with prior knowledge. As can be seen in the table, the use of prior knowledge can reduce the average ARE of the link flow estimates over all the links from 4.24% to 1.74%, while reducing the max ARE from 26.99% to 5.85%.

In addition to the sample means, we were also able to obtain the Bayesian confidence intervals from the samples drawn by the blocked M–H algorithm. Fig. 8 shows 95% Bayesian confidence interval. We used dots to indicate the estimated means of the link flows, and crosses to indicate the “true” link flows. One can see that almost all the “true” link flows fall within the corresponding Bayesian confidence intervals.

Fig. 9 investigates the correlation using the autocorrelation function (ACF) in samples. We calculated the absolute value of the ACF in samples for each route. As can been seen, the average absolute value of the ACFs over the routes is less than 0.1 when the lag between samples larger than 30. The ACF values indicated that the correlations between the samples are low, thus the samples can be considered to be drawn from the conditional distribution, $P(\mathbf{y} \mid sue, \mathbf{x}^{*}, \mathbf{b})$ independently.

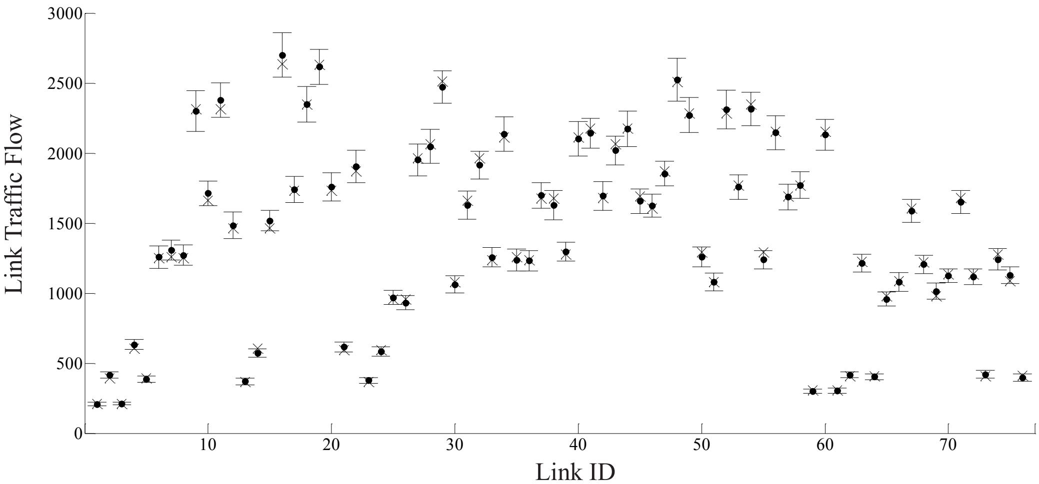

Fig. 8. 95% confidence intervals for link flow estimates with prior knowledge

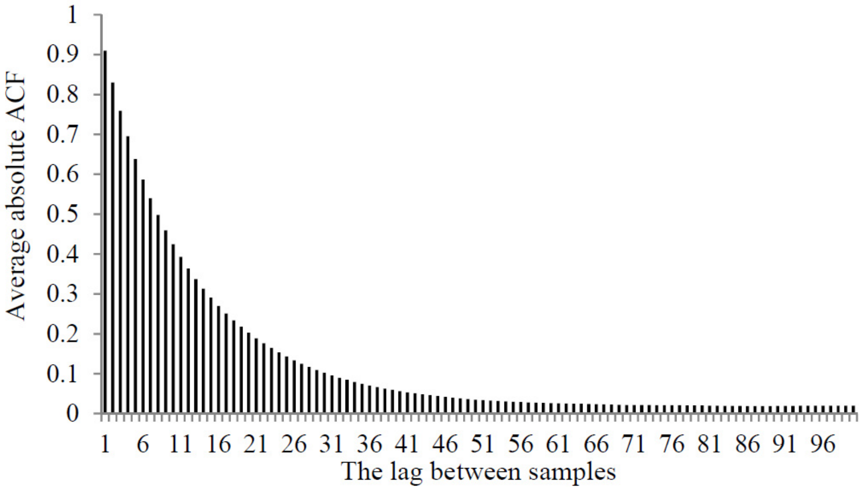

Fig. 9. The absolute value of the ACFs over routes under different lags

5. Conclusions

In this paper, we proposed a likelihood-based statistical model to estimate route/link traffic flows in congested networks. The proposed model represents route flows and O–D demand in a network as a conditional probability distribution that is given for observed link traffic counts in conformity with the SUE principle.

We formulated this conditional distribution from a Bayesian perspective. To facilitate the use of prior knowledge, we designed a hierarchical form of $\mathbf{q}$ that provides a straightforward way to integrate prior knowledge into flow estimates.

Since the conditional distribution was precisely specified in this paper, a blocked M–H algorithm could be developed to estimate the characteristics of the route/link flow variables through sampling from this probability distribution. In addition, another benefit of knowing the probability distribution of traffic is that it can provide insight into the reliability of traffic networks for traffic planning and/or traffic management purposes.

We provided a series of numerical examples to illustrate the performance of the proposed method as applied to the Sioux–Falls network. In these tests, only approximately 30% of links could be observed. As a result, the links were estimated with an ARE of 4.24% when prior knowledge was not available, while this ARE was reduced to 1.74% by using prior knowledge with the hierarchical form.

The proposed model uses observed link counts as input but does not require consistency among the observations. While the proposed model constrains link flow estimates in order to satisfy the SUE principle, it does so without using bi-level formulations. Therefore, the proposed method does not find an equilibrium solution in each iteration. The formulation of the model also implies that there are no specific requirements for the application of user behaviour models. Although this study focuses on static problems, the proposed method does not impose constraints on the use of network loading methods. The dynamics of traffic patterns can also be accounted by using the underlying framework described in study.

Although the Sioux–Falls network has been widely used to test transport models in previous studies, we recognize the need for future works that would test the proposed method on large networks using real world data.

Appendix A. Sampling scheme for $P(\mathbf{y}|sue, \mathbf{x}^*, \mathbf{b})$

Sampling Scheme C

(0) Specify initial samples $\mathbf{y}_1^0, \dots, \mathbf{y}_{|N|}^0$ for $\mathbf{y}_1, \dots, \mathbf{y}_{|N|}$ set $t \leftarrow 1$ and $n \leftarrow 1$.

(1) For the O–D pair n:

(1.1) Draw candidate samples, $\mathbf{y}_n'$, from the proposed distribution $\theta_r(g_r|y_r^{t-1})$ (see Eq. 22).

(1.2) Let $q_n' = \sum_{\forall r \in R_n} y_r'$.

(1.3) Calculate τ as

$$\tau = \frac{P(\dots, \mathbf{y}_n', \dots | sue, \mathbf{x}^*, \mathbf{b})(y_r^{t-1}|y_r')}{P(\dots, \mathbf{y}_n^{t-1}, \dots | sue, \mathbf{x}^*, \mathbf{b})\theta_r(y_r'|y_r^{t-1})}$$

where

$$P(\dots, \mathbf{y}_n', \dots | sue, \mathbf{x}^*, \mathbf{b}) = P\left(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n', \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1} | \, sue, \mathbf{x}^*, \mathbf{b}\right)$$

and

$$P(\dots, \mathbf{y}_n^{t-1}, \dots | sue, \mathbf{x}^*, \mathbf{b}) = P\left(\mathbf{y}_1^t, \dots, \mathbf{y}_{n-1}^t, \mathbf{y}_n^{t-1}, \mathbf{y}_{n+1}^{t-1}, \dots, \mathbf{y}_{|N|}^{t-1} | sue, \mathbf{x}^*, \mathbf{b}\right)$$

(1.4) Accept $\mathbf{y}_n'$ as $\mathbf{y}_n^t = \mathbf{y}_n'$ with a probability of $\min(1, \tau)$; otherwise, $\mathbf{y}_n^t = \mathbf{y}_n^{t-1}$.

(2) If $n < |N|$, then $n \leftarrow n + 1$ and go to step (1); otherwise, go to step (3).

(3) If $t < T$, then $t \leftarrow t + 1$, $n \leftarrow 1$ and go to step (1); otherwise, stop the iteration.

References

[1] Bell, M.G.H. (1991). The estimation of origin-destination matrices by constrained generalised least squares. Transportation Research Part B 39 (2), 169-185.

[2] Sasaki, T. (1967). Probabilistic models for trip distribution. Traffic Engineering 2 (6), 12-21 (in Japanese).

[3] Van Zuylen, H.J., & Willumsen, L.G. (1980). The most likely trip matrix estimated from traffic counts. Transportation Research Part B 14 (3), 281-293.

[4] Van Zuylen, H.J. (1981). Some improvements on the estimation of an origin-destination matrix from traffic counts. Proceedings of the 8th International Symposium on Transportation and Traffic Theory, Toronto University, June 1981.

[5] Yang, H., Sasaki, T., Iida, Y., & Asakura, Y. (1992). Estimation of origin-destination matrices from link traffic counts on congested networks. Transportation Research Part B 26 (6), 417-434.

[6] Bell, M.G.H., Sield, C.M., Busch, F., & Kruse, G. (1997). A stochastic user equilibrium path flow estimator. Transportation Research Part C 5 (3-4), 197-210.

[7] Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. Transportation Research Part B 18 (4-5), 289-299.

[8] Xie, C., Kockelman, K.M., & Waller, S.T. (2011). A maximum entropy-least squares estimator for elastic origin-destination trip matrix estimation. Transportation Research Part B 45 (9), 1465-1482.

[9] Spiess, H. (1987). A maximum likelihood model for estimating origin-destination matrices. Transportation Research Part B 21 (5), 395-412.

[10] Watling, D.P. (1994). Maximum likelihood estimation of an origin-destination matrix from a partial registration plate survey. Transportation Research Part B 28 (4), 289-314.

[11] 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 30 (4), 309-324.

[12] Hazelton, M.L. (2000). Estimation of origin–destination matrices from link flows on uncongested networks. Transportation Research Part B 34 (7), 549-566.

[13] Parry, K., & Hazelton, M. (2012). Estimation of origin–destination matrices from link counts and sporadic routing data. Transportation Research Part B 46 (1), 175-188.

[14] Maher, M.J. (1983). Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. Transportation Research Part B 17 (6), 435-447.

[15] Castillo, E., Menéndez, J.M., & Sánchez-Cambronero, S. (2008). Predicting traffic flow using Bayesian networks. Transportation Research Part B 42 (5), 482-509.

[16] Hazelton, M.L. (2008). Statistical inference for time varying origin–destination matrices. Transportation Research Part B 42 (6), 542-552.

[17] Li, B. (2009). Markov models for Bayesian analysis about transit route origin–destination matrices. Transportation Research Part B 43 (3), 301-310.

[18] Yamamoto, T., Miwa, T., Takeshita, T., & Morikawa, T. (2009). Updating dynamic origin-destination matrices using observed link travel speed by probe vehicles, In W.H.K. Lam, S.C. Wong and H.K. Lo (eds.) Transportation and Traffic Theory 2009: Golden Jubilee, pp. 723-738, 2009.

[19] Perrakis, K., Karlis, D., Cools, M., Janssens, D., Vanhoof, K., & Wets, G. (2012). A Bayesian approach for modeling origin–destination matrices. Transportation Research Part A 46 (1), 200-212.

[20] Hazelton, M.L. (2010). Bayesian inference for network-based models with a linear inverse structure. Transportation Research Part B 44 (5), 647-685.

[21] Wei, C., Asakura, Y., & Iryo, T. (2012). The posterior probability distribution of traffic flow: A new scheme for the assignment of stochastic traffic flow. Transportmetrica (in press) DOI:10.1080/18128602.2012.661799.

[22] Hazelton, M.L. (1998). Some remarks on stochastic user equilibrium. Transportation Research Part B 32 (2), 101-108.

[23] Train, K. (2009). Discrete Choice Methods with Simulation (Second edition). Cambridge, UK: Cambridge University Press.

[24] Sheffi, Y. (1985). Urban transportation networks: equilibrium analysis with mathematical programming methods. Englewood Cliffs, NJ: Prentice-Hall.

[25] Caggiani, L., Ottomanelli, M., Sassanelli, D. A fixed point approach to origin‑destination matrices estimation using uncertain data and fuzzy

programming on congested networks.Transportation Research Part C. (in press）

[26] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., & Teller, E. (1953). Equation of state calculations by fast computing machines. Journal of Chemical Physics 21 (6), 1087-1092.
[27] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57 (1), 97-109.
[28] Han, S. (2007). A route-based solution algorithm for dynamic user equilibrium assignments. Transportation Research Part B 40 (1), 16-34.
[29] Dial, R.B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research Part B 5 (2), 83-111.
[30] Van Zuylen, H.K., & Branston, D.M. (1982). Consistent link flow estimation from counts. Transportation Research Part B 16 (6), 473-476.
[31] Chen, A., Chootinan, P., & Recker, W. (2009). Norm approximation method for handling traffic count inconsistencies in path flow estimator. Transportation Research Part B 43 (8-9), 852-872.