Transportation Research Part B 68 (2014) 52–75

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Estimation of mean and covariance of peak hour origin–destination demands from day-to-day traffic counts

Hu Shao [a,b], William H.K. Lam [a,*], Agachai Sumalee [a,e], Anthony Chen [c], Martin L. Hazelton [d]

[a] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[b] Department of Mathematics, School of Sciences, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China
[c] Department of Civil and Environmental Engineering, Utah State University, Logan, UT 84322-4110, USA
[d] Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand
[e] Department of Civil Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

ARTICLE INFO

Article history:
Received 2 April 2012
Received in revised form 5 June 2014
Accepted 6 June 2014

Keywords:
OD demand variation
OD estimation
Traffic counts
Second-order statistical property
Bi-level optimization problem
Least squares method

ABSTRACT

This paper proposes a generalized model to estimate the peak hour origin–destination (OD) traffic demand variation from day-to-day hourly traffic counts throughout the whole year. Different from the conventional OD estimation methods, the proposed modeling approach aims to estimate not only the mean but also the variation (in terms of covariance matrix) of the OD demands during the same peak hour periods due to day-to-day fluctuation over the whole year. For this purpose, this paper fully considers the first- and second-order statistical properties of the day-to-day hourly traffic count data so as to capture the stochastic characteristics of the OD demands. The proposed model is formulated as a bi-level optimization problem. In the upper-level problem, a weighted least squares method is used to estimate the mean and covariance matrix of the OD demands. In the lower-level problem, a reliability-based traffic assignment model is adopted to take account of travelers' risk-taking path choice behaviors under OD demand variation. A heuristic iterative estimation-assignment algorithm is proposed for solving the bi-level optimization problem. Numerical examples are presented to illustrate the applications of the proposed model for assessment of network performance over the whole year.

1. Introduction

The origin–destination (OD) traffic demand matrix is an important input data for transportation planning and management. Among the conventional OD demand estimation methods, it is generally assumed that the traffic demands of all OD pairs are deterministic. However, due to daily variations in activity patterns, the OD traffic demands of the same hour stochastically fluctuate from day to day over the whole year (Clark and Watling, 2005). Moreover, there exists correlation between traffic demands of different OD pairs particularly when there are major changes in the land use and/or day-to-day activity variations over the year. In view of this, this paper proposes a new model for estimating the mean and variation (in terms of covariance matrix) of peak hour (say from 8:00 am to 9:00 am) OD demands from day-to-day hourly traffic counts.

Recent studies on transportation network reliability and uncertainty issues demonstrate the need for estimating the OD demand variation. In these studies, various models have been formulated under the assumption that the probability

* Corresponding author. Tel.: +852 27666045; fax: +852 23659291.
E-mail address: william.lam@polyu.edu.hk (W.H.K. Lam).

http://dx.doi.org/10.1016/j.trb.2014.06.002
0191-2615/

distribution of OD demands is given as an input data. For example, Chen et al. (2002) used a simulation-based method to examine travelers' risk-taking behaviors under demand uncertainties by assuming that the OD demands follow mutually independent normal distributions. Nakayama and Takayama (2003) proposed a stochastic network equilibrium model to estimate variances of link travel times using Binomial distributions of the OD demands. Clark and Watling (2005) proposed a stochastic network model for estimating network travel time reliability with mutually independent Poisson distributions of the OD demands. Shao et al. (2006a) proposed a reliability-based user equilibrium (RUE) model to capture the travel time reliability-based path choice behaviors under OD demand uncertainty. In their model, the OD demands and path flows are assumed to follow mutually independent normal distributions. Shao et al. (2006b) extended the RUE model to consider the perception errors and multiple risk-taking behaviors. Lam et al. (2008) modified the RUE model to consider both demand and supply uncertainties by a reliability-based stochastic user equilibrium (RSUE) model. Zhou and Chen (2008) analytically derived the mean and variance of link travel times under assumption that the OD demands follow mutually independent log-normal distributions. Chen and Zhou (2010) presented an α-reliable mean-excess traffic equilibrium model to explicitly consider both reliability and unreliability aspects of travel time variability in the path choice decision process. Chen et al. (2011) and Xu et al. (2013) considered the issue of perception error in the stochastic version of the METE model and demonstrated the impacts of stochastic perception error under travel time variability.

In summary, the above studies show that an increasing attention has been directed to the development of reliability-based network equilibrium models under the assumption that the probability distribution of OD demands is known. However, to the best of our knowledge, less attention has been paid to estimation of the probability distribution of the OD demand which is a necessary input for the application of the above models. It is known that the mean and covariance matrix are two key parameters to characterize the OD demand probability distribution. Therefore, the estimation of mean and covariance of the OD demands is an important extension of the current research work on reliability-based network equilibrium models.

In practice, the OD demand can be obtained by survey-based methods, such as home-based survey, road-side interview and license plate recognition technique. These methods can yield accurate results when conducted carefully. However, it is often costly and requires a large amount of resources for implementation. In view of this, the network-wide survey-based method is difficult or impossible to undertake frequently, notwithstanding the advancement in traffic survey technologies (Badoe and Steuart, 2002). OD demand estimation using traffic counts is a cheaper and effective method. Many researchers have developed a variety of OD matrix estimation methods using traffic counts, with great economic advantages in comparison to the costly traditional direct estimation methods based on traffic surveys (Shafahi and Faturechi, 2009).

Although the OD estimation problems using traffic counts have been considerably investigated in the past decades, the corresponding methods cannot fully capture the statistical properties of the OD demand variation. It is assumed in these methods that the OD demands are either deterministic variables or mutually independent random variables.

- It was generally assumed in the early developed and commonly used methods that the OD demand is deterministic. They are the entropy maximizing model (Van Zuylen and Willumsen, 1980; Chen et al., 2009), maximum likelihood model (Spiess, 1987; Watling, 1994), generalized least squares (GLS) model (Cascetta, 1984; Bell, 1991), and Bayesian inference estimation model (Maher, 1983). Recently, works have still been carried out on the assumption of deterministic OD demands, such as Bayesian Networks to predict traffic flows (Sun et al., 2006; Castillo et al., 2008a,b) and Markov chain model (Li, 2009).
- A few studies have recognized the stochastic nature of the OD demands. However, these studies assumed that the OD demands are independently distributed (Hazelton, 2000; Hazelton, 2001b; Parry and Hazelton, 2012). For example, Hazelton assumed that the OD demand follows mutually independent Poisson distributions and proposed a maximum likelihood estimation method (Hazelton, 2000) and a Bayesian inference method (Hazelton, 2001b) for OD estimation. He also provided some valuable comments and inferences regarding OD demand estimation problems under the assumption that the OD demand is mutually independent Poisson distributed (Hazelton, 2001a,b). Li (2005) reinvestigated Bayesian inference for estimating the population means of traffic flows, reconstructing traffic flows, and predicting future traffic flows through the EM (expectation–maximization) algorithm.

Above all, a common feature of the existing OD estimation models is that the correlation between random traffic demands of different OD pairs (covariance) is ignored.

In road transportation networks, the assumption of independent OD demands may result in inaccurate results and the ignoring the correlation between random variables may lead to very different outputs (Haas, 1999). For example, Waller et al. (2001) found that correlation level of OD demands plays a major role in determining the degree of error relating to expected total travel time of the road network. Zhao and Kockelman (2002) discussed the propagation of errors through the four-step traffic demand forecast model. They stated that neglecting the correlation of data (e.g. OD demands) would ultimately reduce the reliability of traffic forecasts, and in turn affect the policy-making and infrastructure decisions. Duthie et al. (2006) relaxed the independence assumption of OD demands in solving the conventional user equilibrium (UE) assignment problem by some simulation-based methods. Their findings indicated that the independence assumption of OD demands can lead to significant over- or under-estimation of network performance. From all the above studies, it is evident that the correlation between OD demands should not be ignored. This is particularly significant for the OD demand