



Article

Bayesian Inference on Dynamic Linear Models of Day-to-Day Origin-Destination Flows in Transportation Networks

Anselmo Ramalho Pitombeira-Neto ^{1,*},[†] , Carlos Felipe Grangeiro Loureiro ^{2,†}  and Luis Eduardo Carvalho ^{3,†}

¹ Department of Industrial Engineering, Federal University of Ceará, 60455-760 Fortaleza, Brazil

² Department of Transportation Engineering, Federal University of Ceará, 60455-760 Fortaleza, Brazil; felipe@det.ufc.br

³ Department of Mathematics and Statistics, Boston University, Boston, MA 02215, USA; lecarval@math.bu.edu

* Correspondence: anselmo.pitombeira@ufc.br; Tel.: +55-85-9990-46960

† These authors contributed equally to this work.

Received: 3 August 2018; Accepted: 7 December 2018; Published: 10 December 2018



Abstract: Estimation of origin–destination (OD) demand plays a key role in successful transportation studies. In this paper, we consider the estimation of time-varying day-to-day OD flows given data on traffic volumes in a transportation network for a sequence of days. We propose a dynamic linear model (DLM) in order to represent the stochastic evolution of OD flows over time. DLMs are Bayesian state-space models which can capture non-stationarity. We take into account the hierarchical relationships between the distribution of OD flows among routes and the assignment of traffic volumes on links. Route choice probabilities are obtained through a utility model based on past route costs. We propose a Markov chain Monte Carlo algorithm, which integrates Gibbs sampling and a forward filtering backward sampling technique, in order to approximate the joint posterior distribution of mean OD flows and parameters of the route choice model. Our approach can be applied to congested networks and in the case when data are available on only a subset of links. We illustrate the application of our approach through simulated experiments on a test network from the literature.

Keywords: origin–destination flows; transportation networks; dynamic linear models; Bayesian inference

1. Introduction

Given a geographic region, one of the main problems in planning and operating transportation systems is the estimation of origin–destination (OD) flows, i.e., the amount of trips made by people or freight between points in the region over a defined time interval. This is also referred to in the literature as the OD matrix estimation problem, since many initial models represented the vector of OD flows as a matrix.

OD flows are traditionally estimated through surveys, in which households or drivers are inquired about their daily journeys [1]. However, direct surveys are expensive and, due to this reason, they are in general carried out every decade [2]. This low frequency implies that planners may remain many years with no data on the evolution of OD flows over time.

In recent years, governments in urban regions have built traffic control systems in order to manage traffic congestion in transportation networks. Those systems automatically gather large amounts of data on traffic volumes at low cost on a daily or even hourly basis. This allowed in theory the indirect estimation of OD flows by means of mathematical models. The idea is that we

can extract information on OD flows from data on traffic volumes if we have a suitable mathematical model which describes their relationships. Thus, OD matrix estimation from traffic volumes may be regarded as the inverse traffic assignment problem [2]. Within the traditional 4-step model, the traffic assignment problem corresponds to the final step, i.e., to the assignment of trips to a network after determination of generated and attracted trips per geographic region, their distribution among regions and their modal split [1]. The output of the traffic assignment step are predicted traffic volumes on links of a transportation network. OD matrix estimation from traffic volumes aims at estimating or reconstructing the trips that generated observed traffic volumes.

The early models for estimation of OD flows assumed that the transportation network from which one observes traffic volumes is not congested. We can trace back the first attempts to the work of Robillard [3], who applied linear regression to estimate OD flows from traffic volumes. Nguyen [4] proposed a model based on Beckman's optimization model [5] for traffic assignment, whose solution is an estimate of the OD flows.

Van Zuylen and Willumsen [6] proposed a nonlinear programming model based on entropy maximization in which the constraints are the linear conservation flow equations, whose solution corresponds to a maximum entropy OD flow configuration. Cascetta [7] proposed a generalized least squares (GLS) model which can cope with errors in observed traffic volumes by minimizing the Mahalanobis distance between predicted volumes and observed volumes. Cascetta and Nguyen [8] and Brenninger-Göthe and Jörnsten [9] proposed general frameworks which generalize previous models.

Fisk [10,11] proposed models which take into account traffic congestion, by seeking solutions which correspond to an equilibrium state as defined for example by Wardrop's first principle [12,13]. Yang et al. [14] proposed a framework based on bilevel optimization, in which the first level corresponds to an objective function (e.g., maxent or GLS) subject to constraints which are obtained as a solution of a traffic assignment model in the second level. Due to the non-convex nature of the model, the authors proposed a heuristic procedure to solve it [15].

Cascetta and Postorino [16] formulated the OD matrix estimation problem as a compound fixed point problem, in which the solution of an inner fixed point problem corresponds to an equilibrium traffic assignment of an OD matrix obtained as a solution of an outer fixed point problem. In order to solve the problem, they propose a fixed point iteration based on the method of successive averages [17].

A common caveat in the abovementioned models is that they do not take variability of OD flows into account. Assuming that OD flows follow independent Poisson probability distributions, Vardi [18] considered the estimation of mean OD flows given a sample of observed traffic volumes' vectors. Since the mean and variance of Poisson random variables are equal, variance of traffic volumes carry information on mean OD flows. He also assumed the existence of a single route for each OD pair, and proposed maximum likelihood and moment-based estimators for the mean OD flows.

Following Vardi's work, Tebaldi and West [19] proposed Bayesian estimators for the case when a sample of size one is available for the traffic volumes vector. In some applications, it may be the case that a prior OD matrix may be available from older studies. A Bayesian statistical model allows the incorporation of knowledge on OD flows as a prior probability distribution, thereby improving modeling and inference. Hazelton [20–22] extended both Vardi's and Tebaldi and West's works. For example, both latter works assumed a single route in each OD pair, which is not realistic, while Hazelton considered multiple routes. In addition, the Poisson assumption poses computational hurdles for the application of Vardi's or Tebaldi and West's models to large-scale networks. The Hazelton solution was to approximate the distribution of OD flows with a multivariate normal density, which has more tractable computational properties, further assuming that the variances of OD flows are functions of their means as in the Poisson probability function.

More recent approaches consider the development of dynamic models for the estimation of OD flows. As traffic volumes are observed over time, we cannot ignore their temporal nature. Here, we must make a distinction between *within-day* and *day-to-day* dynamics. When considering within-day dynamics, one is concerned with the estimation of OD flows over the course of a single day [23–25].

In contrast, in day-to-day dynamics, we consider the estimation of time-varying OD flows for a reference time period, e.g., the morning peak, given a time series of observed traffic volumes over many consecutive days. It is assumed that this reference time period is long enough so that most trips in the study region start and finish within the observation period.

Hazelton [26] developed one of the first models for the estimation of day-to-day dynamic OD flows. The author assumes that mean OD flows are functions of parameters that do not change over time. Particular cases include constant demand, linear trend and weekday-weekend models. Parry and Hazelton [27] proposed a general Bayesian framework for the estimation of parameters of day-to-day dynamic traffic models. They assumed that OD flows vary according to a Markovian transition kernel, and proposed a Markov chain Monte Carlo (MCMC) algorithm for the estimation of parameters. Hazelton and Parry [28] proposed statistical methods to compare alternative day-to-day dynamic models.

A central aspect in the development of realistic day-to-day dynamic models is the modeling of users' behavior through route choice models. In practice, users choose routes based on the travel times they have experienced over past days. Thus, a suitable route choice model should take into account the influence of past route costs on the decision of users in a given time. Watling and Cantarella [29,30] discuss models for the analysis of dynamic users' behavior.

In this paper, we represent stochastic day-to-day OD flows as a dynamic linear model (DLM), which allows us to capture the time dependencies of OD flows as well as non-stationarity. In the formulation of the DLM, we take into account the variability originating from OD flows, user's route choices and measurement of traffic volumes on links. The main contribution of our proposed model is that we represent route choices through a utility model which takes into account the adaptive behavior of users to experienced past route costs. This allows us to jointly estimate dynamic assignment matrices and OD flows in a statistically principled manner. In addition, our model does not rely on equilibrium assumptions, so that it can be equally applied to noncongested and congested networks. We also propose an MCMC algorithm, based on Gibbs sampling and forward filtering backward sampling, in order to approximate the joint posterior distributions of mean OD flows and parameters of the route choice model. We illustrate its application through numerical studies in a network from the literature.

This paper is organized as follows: in Section 2.1, we describe the proposed dynamic linear model; in Section 2.2, we describe a route choice model and an MCMC algorithm for the approximation of the joint posterior of mean OD flows and route choice parameters; in Section 3, we present numerical studies and discuss the results; finally, in Section 4, we draw some conclusions and propose further developments.

2. Modeling Framework

2.1. A Dynamic Linear Model for Day-to-Day OD Flows

In Section 2.1.1, we define the proposed dynamic linear model and in Section 2.1.2 we describe the procedure for estimating mean OD flows. A nomenclature with all symbols used may be found after the Conclusions section.

2.1.1. Model Definition

Dynamic linear models are defined by a state transition equation, which describes how a system evolves in time, and an observational equation, which describes how observed quantities relate to the system's states. They have a Markovian structure and assume multivariate normal probability distributions for the random variables. System's states are regarded as unobserved parameters whose estimation is done through Bayesian updating. We refer to [31] for an excellent introduction to the corresponding theory.

Let $(\mathcal{N}, \mathcal{L})$ be a transportation network in which \mathcal{N} is a set of nodes, \mathcal{L} a set of directed links and $\mathcal{J} \subseteq \mathcal{N} \times \mathcal{N}$ a set of OD pairs. For a sequence of subsequent time periods $t = 1, 2, \dots, T$, we define

$\theta_t = (\theta_{1t}, \theta_{2t}, \dots, \theta_{nt})^T$ as the mean OD flow vector, in which θ_{jt} is the mean OD flow for OD pair $j \in \mathcal{J}$ at time t and $n = |\mathcal{J}|$. We define $z_t = (z_{1t}, z_{2t}, \dots, z_{mt})^T$ as the vector of observed traffic volumes in a subset of links in a network at time t , in which z_{it} is the observed volume on link $i \in \mathcal{I} \subseteq \mathcal{L}$ and $m = |\mathcal{I}|$.

We consider the vector of mean OD flows θ_t as the *unobserved* system state vector, which depends only on its previous state θ_{t-1} at time $t - 1$ plus a stochastic error. The vector of observed link volumes z_t at time t depends only on the current unobserved vector of states, such that

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \omega_t, \tag{1}$$

$$z_t = \mathbf{F}_t \theta_t + \nu_t. \tag{2}$$

We refer to Equation (1) as the dynamic model, while Equation (2) is the observational model. \mathbf{G}_t is the system matrix and $\omega_t \sim N(\mathbf{0}, \mathbf{W}_t)$ is a stochastic error, where the letter “N” stands for the multivariate normal density with the appropriate dimension and \mathbf{W}_t is a covariance matrix also referred to as an *evolution matrix*. It governs the stochastic evolution of the system through time. We notice that $\theta_t | \theta_{t-1} \sim N(\mathbf{G}_t \theta_{t-1}, \mathbf{W}_t)$. \mathbf{F}_t is an *assignment matrix* which relates mean OD flows to observed traffic volumes, while $\nu_t \sim N(\mathbf{0}, \mathbf{V}_t)$ is a zero-mean error term which corresponds to the deviation of observed volumes relative to the expected value $E[z_t | \theta_t] = \mathbf{F}_t \theta_t$. We have that $z_t | \theta_t \sim N(\mathbf{F}_t \theta_t, \mathbf{V}_t)$.

In order to fully determine the model given by Equations (1) and (2), we must specify the corresponding matrices. Regarding the system matrix, we assume that, in the short term, mean OD flows are *locally constant*. This implies that at time t mean OD flows should be approximately equal to mean OD flows at time $t - 1$, so that $\theta_t = \theta_{t-1} + \omega_t$ and $\mathbf{G}_t = \mathbf{I}$, the identity matrix with the appropriate dimension. The evolution matrix \mathbf{W}_t is in general supplied by the modeler.

Regarding matrices \mathbf{F}_t and \mathbf{V}_t of the observational model, they must represent correctly the relationship between mean OD flows and observed traffic volumes. We define x_t as *realized* OD flows around mean OD flows at time t , which distributes among flows on routes y_t , which aggregate into observed traffic volumes z_t . Figure 1 summarizes the hierarchical relationship between variables in our proposed DLM.

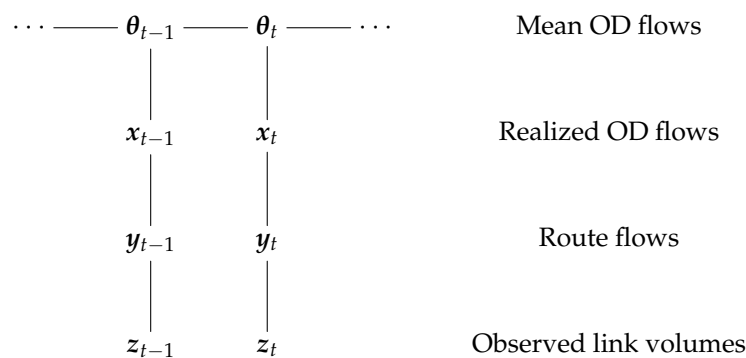


Figure 1. Hierarchical relationship between variables in the DLM. Mean OD flows, realized OD flows and route flows are unobserved variables.

Let $x_t = (x_{1t}, x_{2t}, \dots, x_{nt})^T$ be the vector of realized OD flows at time t . We assume $x_t | \theta_t \sim N(\theta_t, \Sigma_t^x)$ where Σ_t^x is a covariance matrix, which can account for correlations between OD flows or simply be a diagonal matrix in case OD flows are independent. Given a realized vector x_t , for each OD pair j , there is a vector of route flows $y_{jt} = (y_{1jt}, y_{2jt}, \dots, y_{n(j)jt})$, in which $n(j) = |\mathcal{K}_j|$ is the size of the route set \mathcal{K}_j of OD pair j , and let p_{kjt} be the probability of route $k \in \mathcal{K}_j$ being selected at time t . We assume y_{jt} follows a multivariate normal distribution with a multinomial-like covariance structure, i.e.,

$$E[y_{kjt}] = x_{jt}p_{kjt},$$

$$\text{Cov}(y_{kjt}, y_{ljt}) = \begin{cases} x_{jt}p_{kjt}(1 - p_{kjt}) & \text{if } l = k, \\ -x_{jt}p_{kjt}p_{ljt} & \text{if } l \neq k, \end{cases}$$

where $E[\cdot]$ and $\text{Cov}(\cdot)$ denote the expected value and the covariance, respectively. Thus, $\mathbf{y}_{jt}|x_{jt}, \mathbf{p}_{jt} \sim N(x_{jt}\mathbf{p}_{jt}, \Sigma_{jt}^y)$, in which $\mathbf{p}_{jt} = (p_{1jt}, p_{2jt}, \dots, p_{n(j)jt})^T$ is the vector of route choice probabilities of OD pair j and the covariance matrix is given by

$$\Sigma_{jt}^y = x_{jt}(\text{diag}(\mathbf{p}_{jt}) - \mathbf{p}_{jt}\mathbf{p}_{jt}^T), \tag{3}$$

where $\text{diag}(\mathbf{p}_{jt})$ denotes a matrix whose diagonal elements correspond to the vector \mathbf{p}_{jt} and all other elements are zero. Notice also that as $\sum_{k \in \mathcal{K}_j} p_{kjt} = 1$, the covariance matrix Σ_{jt}^y will be singular and the corresponding multivariate normal distribution will be degenerate. In order to avoid this, we assume there is a positive probability π of none of the routes in a route set being chosen, so that $\sum_{k \in \mathcal{K}_j} p_{kjt} < 1$ assuring that Σ_{jt}^y is non-singular. This assumption will be realistic if the route sets include many routes, so that flows on the excluded routes are negligible.

Let $\mathbf{y}_t = (\mathbf{y}_{1t}, \mathbf{y}_{2t}, \dots, \mathbf{y}_{nt})^T$ be the joint vector of route flows across all OD pairs. We also define $\mathbf{P}_t = \text{blockdiag}_{j \in \mathcal{J}} \{\mathbf{p}_{jt}\}$ as a block-diagonal matrix composed of route choice probability vectors and $\Sigma_t^y = \text{blockdiag}_{j \in \mathcal{J}} \{\Sigma_{jt}^y\}$ as a block-diagonal matrix composed of covariance matrices of route flows, i.e.,

$$\mathbf{P}_t = \begin{bmatrix} \mathbf{p}_{1t} & & & \\ & \mathbf{p}_{2t} & & \\ & & \ddots & \\ & & & \mathbf{p}_{nt} \end{bmatrix}, \tag{4}$$

$$\Sigma_t^y = \begin{bmatrix} \Sigma_{1t}^y & & & \\ & \Sigma_{2t}^y & & \\ & & \ddots & \\ & & & \Sigma_{nt}^y \end{bmatrix}. \tag{5}$$

Since we defined multivariate normal distributions for all subvectors \mathbf{y}_{jt} , and assuming that route flow subvectors \mathbf{y}_{jt} are conditionally independent given realized OD flow vector \mathbf{x}_t , we have that $\mathbf{y}_t|\mathbf{x}_t, \mathbf{P}_t \sim N(\mathbf{P}_t\mathbf{x}_t, \Sigma_t^y)$. (For convenience, we omit the explicit dependence on \mathbf{P}_t in the following developments).

Now, we are able to obtain the conditional $\mathbb{P}(\mathbf{y}_t|\boldsymbol{\theta}_t)$ by marginalizing \mathbf{x}_t . Since both the densities $\mathbb{P}(\mathbf{x}_t|\boldsymbol{\theta}_t)$ and $\mathbb{P}(\mathbf{y}_t|\mathbf{x}_t)$ are multivariate normal, and ignoring the dependence of Σ_t^y on \mathbf{x}_t (we will take the dependence of Σ_t^y on \mathbf{x}_t into account in the inference of $\boldsymbol{\theta}_t$ in next section), from the properties of the multivariate normal [32], we have

$$\mathbb{P}(\mathbf{y}_t|\boldsymbol{\theta}_t) = N(\mathbf{P}_t\boldsymbol{\theta}_t, \mathbf{P}_t\Sigma_t^x\mathbf{P}_t^T + \Sigma_t^y).$$

In order to complete the formulation of our model, we must obtain the conditional density $\mathbb{P}(\mathbf{z}_t|\boldsymbol{\theta}_t)$ of observed traffic volumes given mean OD flows. First, we assume that $\mathbf{z}_t|\mathbf{y}_t \sim N(\Delta\mathbf{y}_t, \Sigma_t^z)$, where Σ_t^z is the covariance matrix of measurement errors when observing volumes on links. Δ is the link-path incidence matrix for observed links, whose entries $a_{ik} = 1$ if link i make part of route k and $a_{ik} = 0$, otherwise. It is a deterministic parameter which is a function of the topology of the network and of the route choice sets.

Since both $\mathbb{P}(z_t|y_t)$ and $\mathbb{P}(y_t|\theta_t)$ are multivariate normal, by marginalizing y_t , we have

$$\mathbb{P}(z_t|\theta_t) = N(\Delta P_t \theta_t, \Delta(P \Sigma_t^x P_t^T + \Sigma_t^y) \Delta^T + \Sigma_t^z). \tag{6}$$

From Equation (6), we can finally specify F_t and V_t by defining

$$F_t = \Delta P_t, \tag{7}$$

as an assignment matrix and the covariance matrix by

$$V_t = F_t \Sigma_t^x F_t^T + \Delta \Sigma_t^y \Delta^T + \Sigma_t^z. \tag{8}$$

It is noteworthy in Equation (8) that the variability in OD flows, in route choices and in volume measurement are all represented in the covariance matrix. In the next section, we describe the estimation equations for mean OD flows.

2.1.2. Estimation of Mean OD Flows

Estimation of mean OD flows θ_t from observed link volumes can be made by Bayesian updating. We assume that the covariance matrices Σ_t^x , Σ_t^z and W_t , the link-path incidence matrix Δ , the route choice matrices P_t and the probability π of none of the routes in a route set being chosen are known parameters and are part of the prior knowledge.

At time t , let $\mathbb{P}(\theta_{t-1}|z_{0:t-1}) = N(m_{t-1}, C_{t-1})$ be the posterior distribution in the previous time step $t - 1$, where $z_{0:t-1} = \{z_0, z_1, \dots, z_{t-1}\}$ is the set of observed volume vectors in past time periods. We have a prior distribution $\mathbb{P}(\theta_t|z_{0:t-1}) = N(\bar{m}_t, \bar{C}_t)$ before the observation of z_t . Since, in our local constant model given by Equation (1), the system matrix $G_t = I$, then

$$\bar{m}_t = m_{t-1}, \tag{9}$$

$$\bar{C}_t = C_{t-1} + W_t. \tag{10}$$

Notice that uncertainty on mean OD flows increases by addition of the evolution matrix W_t to the covariance matrix C_{t-1} . Next, we define $\mathbb{P}(z_t|z_{0:t-1}) = N(f_t, Q_t)$ as the one-step forecast distribution of the vector of observed link volumes, where

$$f_t = F_t \bar{m}_t, \tag{11}$$

$$Q_t = F_t \bar{C}_t F_t^T + V_t. \tag{12}$$

It is worth noting that, in computing V_t by Equation (8), the covariance matrix of route flows Σ_t^y is computed based on the predicted mean OD flows $\bar{m}_t = (\bar{m}_{1t}, \bar{m}_{2t}, \dots, \bar{m}_{nt})^T$, i.e., $\hat{\Sigma}_t^y = \text{blockdiag}_{j \in \mathcal{J}} \{\hat{\Sigma}_{jt}^y\}$ in which from Equation (3)

$$\hat{\Sigma}_{jt}^y = \bar{m}_{jt} (\text{diag}(p_{jt}) - p_{jt} p_{jt}^T). \tag{13}$$

Finally, the parameters of the posterior distribution $p(\theta_t|z_{0:t}) = N(m_t, C_t)$ are given by the Bayesian updating of the parameters, which in this case coincide with Kalman filter updating equations (see [31])

$$m_t = \bar{m}_t + A_t (z_t - f_t), \tag{14}$$

$$C_t = \bar{C}_t - A_t Q_t A_t^T, \tag{15}$$

$$A_t = \bar{C}_t F_t^T Q_t^{-1}, \tag{16}$$

where A_t is an adjustment matrix, which controls how the parameters from the posterior distribution are modified according to the new observation z_t . In particular, the adjustment matrix is a function of the prior covariance matrix \bar{C}_t and of the inverse Q_t^{-1} of the covariance matrix of the one-step forecast

distribution of link volumes, so that the adjustment matrix gives more or less weight to observed link volumes according to their uncertainty relative to the uncertainty in OD flows.

At a time t , $\hat{\theta}_t = m_t$ is an estimator of the mean OD flows and C_t is a measure of uncertainty of the estimate. At $t = 0$, we define $\mathbb{P}(\theta_0|z_0) = N(m_0, C_0)$, where z_0 are not actually observed link volumes but symbolically represents the modeler’s prior knowledge on OD flows. The estimation procedure can be summarized as follows:

1. Starting from $t = 0$, set $\mathbb{P}(\theta_0|z_0) = N(m_0, C_0)$;
2. For $t = 1$ to T , do:
 - (a) Determine prior distribution parameters \bar{m}_t and \bar{C}_t by means of Equations (9) and (10), respectively.
 - (b) Compute the assignment matrix F_t by means of Equation (7).
 - (c) Compute the covariance matrix V_t by means of Equations (8) and (13).
 - (d) Determine the parameters of the one-step forecast distribution f_t and Q_t by using Equations (11) and (12).
 - (e) Determine posterior distribution parameters m_t and C_t by means of Equations (14)–(16).

In the next section, we model route choice probabilities based on a utility model in order to estimate jointly mean OD flows and the route choice matrix.

2.2. Bayesian Inference on Route Choice Parameters

In Section 2.2.1, we treat route choice probabilities as unobserved quantities and represent them through a utility model. In Section 2.2.2, we propose an MCMC algorithm to approximate the joint posterior distribution of route choice parameters and mean OD flows.

2.2.1. Utility Model and Route Choice Probabilities

In congested networks, route choice probabilities greatly depend on past route costs. Users take into account their previous experiences in order to evaluate the utility associated with each route. Utilities will vary over time depending on user’s memory, their sensitivity to route costs, fluctuations in mean OD flows, among other possibly influencing factors. As route choice probabilities are functions of utilities, these will also dynamically change.

We assume that the utility u_{kjt} of a route k in an OD pair j at a time t as perceived by users is a linear function of its past route that costs

$$u_{kjt} = - \sum_{s=1}^r \phi_s c_{kj(t-s)}. \tag{17}$$

ϕ_s are parameters, $c_{kj(t-s)}$ are observed past route costs and r is the length of users’ memory. The minus sign in Equation (17) is used since costs are disutilities, in order that routes with lower costs are preferable. We may interpret the parameters $\phi = (\phi_1, \phi_2, \dots, \phi_r)^T$ as users’ sensitivities to past route costs. It is reasonable to assume that they have non-negative values; otherwise, higher costs will contribute to higher utilities. We also expect users to be more sensitive to recent route costs than to older costs, which implies that $\phi_1 \geq \phi_2 \geq \dots \geq \phi_r$.

We adopt a multinomial logit model for route choice probabilities [33] in order that the probability p_{kjt} of a route $k \in \mathcal{K}_j$ in OD pair j at time t being chosen by a user is given by

$$p_{kjt} = (1 - \pi) \frac{e^{u_{kjt}}}{\sum_{l \in \mathcal{K}_j} e^{u_{ljt}}}, \tag{18}$$

where π is the probability of a user not following a route in a route choice set and e denotes Euler’s number.

Let $c_{jt} = (c_{1jt}, c_{2jt}, \dots, c_{n(j)jt})^T$ be the vector of route costs for routes in OD pair j at time t and $c_t = (c_{1t}, c_{2t}, \dots, c_{nt})^T$. We notice that, given route costs c_t for $t = -(r - 1), \dots, 0, 1, \dots, T$, and parameters ϕ and π , we uniquely determine the vectors $p_{jt} = (p_{1jt}, p_{2jt}, \dots, p_{n(j)jt})^T$ and the route choice matrix P_t for $t = 1$ to T by means of Equations (4), (17) and (18).

2.2.2. An MCMC Algorithm

In this section, we develop an MCMC algorithm in order to sample from the joint posterior distribution of mean OD flows and the parameters of the route choice utility model given in (17).

From Bayes theorem, we have:

$$\mathbb{P}(\boldsymbol{\theta}_{1:T}, \boldsymbol{\phi} | \mathbf{z}_{1:T}, D_0) \propto \mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\theta}_{1:T}, \boldsymbol{\phi}, D_0) \mathbb{P}(\boldsymbol{\theta}_{1:T}, \boldsymbol{\phi} | D_0), \tag{19}$$

where $\boldsymbol{\theta}_{1:T}$ is a shorthand referring to the vectors $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_T$ and $\mathbf{z}_{1:T}$ a shorthand for z_1, z_1, \dots, z_T . $\mathbb{P}(\boldsymbol{\theta}_{1:T}, \boldsymbol{\phi} | \mathbf{z}_{1:T}, D_0)$ is the joint posterior, $\mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\theta}_{1:T}, \boldsymbol{\phi}, D_0)$ is the likelihood function of the observed link volumes and $\mathbb{P}(\boldsymbol{\theta}_{1:T}, \boldsymbol{\phi} | D_0)$ is the prior distribution. We include in prior knowledge set D_0 all other parameters, such as $\pi, \Delta, \mathbf{W}_t, \boldsymbol{\Sigma}_t^x, \boldsymbol{\Sigma}_t^z$ for $t = 1, \dots, T$, and route costs c_t for $t = -(r-1), \dots, 0, 1, \dots, T$.

We notice that Gibbs sampling [34] allows us to sample from the joint distribution $\mathbb{P}(\boldsymbol{\theta}_{1:T}, \boldsymbol{\phi} | \mathbf{z}_{1:T}, D_0)$ by alternately sampling from conditional distributions $\mathbb{P}(\boldsymbol{\theta}_{1:T} | \boldsymbol{\phi}, \mathbf{z}_{1:T}, D_0)$ and $\mathbb{P}(\boldsymbol{\phi} | \boldsymbol{\theta}_{1:T}, \mathbf{z}_{1:T}, D_0)$. We must then devise how to sample from each distribution.

Sampling from $\mathbb{P}(\boldsymbol{\theta}_{1:T} | \boldsymbol{\phi}, \mathbf{z}_{1:T}, D_0)$ can be implemented through the *forward filtering backward sampling* (FFBS) method [31,35,36]:

1. (forward filtering) For $t = 1, \dots, T$, apply Kalman filtering recurrences given by Equations (14) and (15). Save $\mathbf{m}_t, \bar{\mathbf{m}}_t, \mathbf{C}_t$ and $\bar{\mathbf{C}}_t$ for $t = 1, \dots, T$;
2. (backward sampling) Starting from $\boldsymbol{\theta}_T \sim \mathcal{N}(\mathbf{m}_T, \mathbf{C}_T)$, for $t = T-1, \dots, 1$, sample each $\boldsymbol{\theta}_t$ backwards from the conditionals $\mathbb{P}(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t+1}, \boldsymbol{\phi}, \mathbf{z}_{1:T}) = \mathcal{N}(\mathbf{h}_t, \mathbf{H}_t)$, in which:

$$\mathbf{h}_t = \mathbf{m}_t + \mathbf{B}_t(\boldsymbol{\theta}_{t+1} - \bar{\mathbf{m}}_{t+1}), \tag{20}$$

$$\mathbf{H}_t = \mathbf{C}_t - \mathbf{B}_t \bar{\mathbf{C}}_t \mathbf{B}_t^T, \tag{21}$$

and $\mathbf{B}_t = \mathbf{C}_t \bar{\mathbf{C}}_{t+1}^{-1}$. We notice that step 1 (forward filtering) of the FFBS algorithm is possible since, given $\boldsymbol{\phi}$, observed link volumes \mathbf{z}_t and prior knowledge D_0 , we have the route choice matrix \mathbf{P}_t and all other parameters determined for $t = 1, \dots, T$, so that we can apply recurrence Equations (14) and (15). Then, step 2 generates a sample of mean OD flows $\boldsymbol{\theta}_{1:T}$.

Regarding the conditional $\mathbb{P}(\boldsymbol{\phi} | \boldsymbol{\theta}_{1:T}, \mathbf{z}_{1:T}, D_0)$, sampling can be performed through a Metropolis–Hastings (MH) step [37]. First, we notice that:

$$\begin{aligned} \mathbb{P}(\boldsymbol{\phi} | \boldsymbol{\theta}_{1:T}, \mathbf{z}_{1:T}, D_0) &\propto \mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\phi}, \boldsymbol{\theta}_{1:T}, D_0) \mathbb{P}(\boldsymbol{\theta}_{1:T} | \boldsymbol{\phi}, D_0) \mathbb{P}(\boldsymbol{\phi} | D_0) \\ &\propto \mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\phi}, \boldsymbol{\theta}_{1:T}, D_0) \mathbb{P}(\boldsymbol{\phi} | D_0), \end{aligned} \tag{22}$$

as $\boldsymbol{\theta}_{1:T}$ does not depend on $\boldsymbol{\phi}$, and $\mathbb{P}(\boldsymbol{\phi} | D_0)$ is the prior distribution of $\boldsymbol{\phi}$. From (6), the likelihood term is given by:

$$\mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\phi}, \boldsymbol{\theta}_{1:T}, D_0) \propto \prod_{t=1}^T \|\mathbf{V}_t\|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{z}_t - \mathbf{F}_t \boldsymbol{\theta}_t)^T \mathbf{V}_t^{-1} (\mathbf{z}_t - \mathbf{F}_t \boldsymbol{\theta}_t) \right\}, \tag{23}$$

in which matrices \mathbf{F}_t and \mathbf{V}_t are dependent on $\boldsymbol{\phi}$ since they are functions of the route choice matrices \mathbf{P}_t , which are dependent on $\boldsymbol{\phi}$ through Equations (17) and (18).

The MH step is then performed as follows: given a candidate vector $\boldsymbol{\phi}' \sim \mathcal{Q}(\boldsymbol{\phi}' | \boldsymbol{\phi}^{(i)})$, where $\boldsymbol{\phi}^{(i)}$ is the current vector at iteration i and $\mathcal{Q}(\cdot)$ is a proposal distribution, sample $u \sim \text{Uniform}(0, 1)$ and accept $\boldsymbol{\phi}'$ as next sample $\boldsymbol{\phi}^{(i+1)}$ if:

$$u < \min \left(1, \frac{\mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\phi}', \boldsymbol{\theta}_{1:T}, D_0) \mathbb{P}(\boldsymbol{\phi}' | D_0) \mathcal{Q}(\boldsymbol{\phi}^{(i)} | \boldsymbol{\phi}')}{\mathbb{P}(\mathbf{z}_{1:T} | \boldsymbol{\phi}^{(i)}, \boldsymbol{\theta}_{1:T}, D_0) \mathbb{P}(\boldsymbol{\phi}^{(i)} | D_0) \mathcal{Q}(\boldsymbol{\phi}' | \boldsymbol{\phi}^{(i)})} \right); \tag{24}$$

otherwise, make $\boldsymbol{\phi}^{(i+1)} = \boldsymbol{\phi}^{(i)}$.

Finally, the proposed MCMC algorithm is summarized below:

1. Initialize vectors $\phi^{(0)}$ and $\theta_{1:T}^{(0)}$.
2. From iteration $i = 0$ onwards, repeat until convergence:
 - (a) (FFBS step) Sample $\theta_{1:T}^{(i+1)}$ by applying updating Equations (14) and (15) and sampling backwards through Equations (20) and (21);
 - (b) (MH step) Sample a candidate ϕ' according to a proposal distribution $Q(\phi'|\phi^{(i)})$ and make $\phi^{(i+1)} = \phi'$ according to acceptance test given in Equation (24), otherwise make $\phi^{(i+1)} = \phi^{(i)}$.

In the next section, we illustrate the application of the proposed MCMC algorithm through some numerical studies.

3. Results and Discussion

In the following subsections, we describe the generation of simulated data and present the results from the application of our approach to a test network from the literature.

3.1. Generation of Simulated Data

In order to illustrate the application of our proposed DLM and MCMC algorithm, we generated data for a transportation network from [28], given in Figure 2, which has eight nodes and 10 links. Nodes 1 and 2 are the origins and nodes 7 and 8 are the destinations, so that we have four OD pairs: (1,7), (1,8), (2,7) and (2,8).

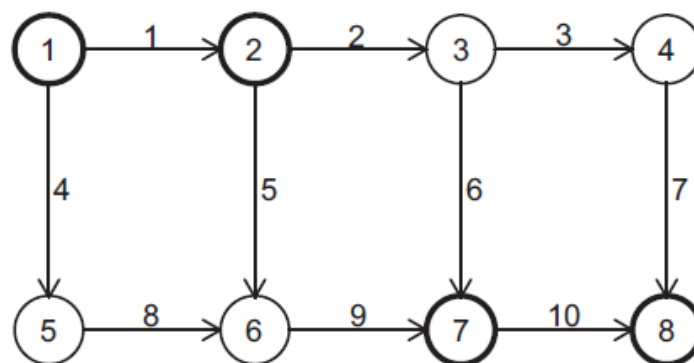


Figure 2. Network used in the numerical studies (adapted from [28]).

In order to adapt the network to our experiment, we did not use the original data on demand and link parameters. We assumed that travel times on links are given by BPR functions according to Equation (25):

$$\tau(z) = \tau_0 \left[1 + \alpha \left(\frac{z}{z^{\max}} \right)^\beta \right], \tag{25}$$

where $\tau(z)$ is the travel time on a link with traffic volume z , τ_0 denotes the travel time in “free flow” and z^{\max} is the capacity of the link. α and β are parameters of the function, for which we adopt the typical values from the literature 0.15 and 4, respectively [38]. Note that the capacity of the link is not treated as a hard constraint, i.e., simulated volumes z are allowed to be greater than z^{\max} . In the test network, we set $\tau_0 = 1$ and $z^{\max} = 130$ on all links. The procedure used to generate the simulated data was the following:

1. At $t = 0$, set values for Δ , θ_0 , \mathbf{W}_t , Σ_t^x , Σ_t^z , ϕ , π for $t = 1$ to T and past route costs $c_{1-r} = c_{1-r+1} = \dots = c_0 = c_{\text{freeflow}}$, where r is the size of users’ memory and c_{freeflow} denotes route costs computed assuming free flow in the network.
2. For $t = 1$ to T do:
 - (a) Sample mean OD flows $\theta_t \sim N(\theta_{t-1}, \mathbf{W}_t)$.

- (b) Sample OD flows $x_t | \theta_t \sim N(\theta_t, \Sigma_t^x)$.
- (c) Compute u_{kt} and p_{kt} for all routes $k \in \mathcal{K}_j$ for all OD pairs $j \in \mathcal{J}$ from Equations (17) and (18), respectively, and compute the route choice matrix $\mathbf{P}_t = \text{blockdiag}_{j \in \mathcal{J}} \{ \mathbf{p}_{jt} \}$.
- (d) Sample route flows $\mathbf{y}_t | x_t \sim N(\mathbf{P}_t x_t, \Sigma_t^y)$, where Σ_t^y is calculated from Equation (3).
- (e) Calculate route costs $c_t = \Delta^T g(\Delta \mathbf{y}_t)$, where $g(\cdot)$ is a vector function which returns a vector of costs on links based on BPR functions given by Equation (25).
- (f) Sample observed traffic volumes $z_t | \mathbf{y}_t \sim N(\Delta \mathbf{y}_t, \Sigma_t^z)$.

OD flows vary within the interval [10.0, 100.0] and follow the locally constant model given in Equation (1), starting at time $t = 0$ from the OD flow vector $\theta_0 = (50.0, 50.0, 50.0, 50.0)^T$. The OD pairs are ordered lexicographically in vector θ_t , so that mean OD flows in OD pair (1,7) corresponds to θ_{1t} , (1,8) corresponds to θ_{2t} and so on. The evolution covariance matrix is constant and given by $\mathbf{W}_t = 10\mathbf{I}$, where \mathbf{I} is the appropriate identity matrix. We assume that variability around mean OD flows is negligible, so as to $\Sigma_t^x = \mathbf{I}$ and $\theta_t \approx x_t$ for all t , and we also assume that $\Sigma_t^z = \mathbf{I}$. These assumptions imply that variability in observed link volumes will be mostly due to variability in route choices. We simulate OD flows for $T = 100$ time periods.

In addition, we set a users' memory length of $r = 2$ for the utility model, so that utility of a route k of OD pair j at time t is given by $u_{kjt} = -\phi_1 c_{kj(t-1)} - \phi_2 c_{kj(t-2)}$ (Equation (17)) and $\phi = (\phi_1, \phi_2)^T = (0.5, 0.3)^T$ is the vector of route choice parameters. We enumerated exhaustively all 12 routes between origins and destinations, resulting in a link-path incidence matrix Δ with 10 rows and 12 columns. Since we enumerated all routes, we assumed a small probability $\pi = 0.01$ of a trip not following one of the routes. Table 1 shows the mean congestion level (CL) for a link i , given by the ratio between mean simulated traffic volumes and the link capacities:

$$CL_i = \sum_{t=1}^T z_{it} / Tz^{\max}. \tag{26}$$

As can be seen in Table 1, links 2, 5, 6 and 9, which are in the central area of the network, have high congestion levels.

Table 1. Mean congestion level (CL) on links from simulated OD flows.

Link	1	2	3	4	5	6	7	8	9	10
CL	0.5500	0.8940	0.2482	0.2535	0.6378	0.6473	0.2481	0.2542	0.8914	0.5685

The simulations, the proposed DLM and MCMC algorithm were implemented in the Python programming language version 2.7 by using Numerical Python module version 1.10.

3.2. Application of the MCMC Algorithm

When applying the MCMC algorithm, we assume that we observe link volumes $z_{1:T}$ on all links and route costs $c_{1:T}$. In practice, these data on traffic volumes are routinely collected through modern ITS systems and route costs can be estimated from data on travel times on links or can be collected from a sample of cars following selected routes. Except for the mean OD flows θ_t and the parameters ϕ , we assume all other parameters are known. We use an uninformative multivariate normal prior for the initial mean OD flows vector θ_0 , with $m_0 = 100\mathbf{1}$, where $\mathbf{1}$ is the unit vector with the appropriate dimension and $\mathbf{C}_0 = 1000\mathbf{I}$, and an uninformative improper prior for ϕ .

The MCMC algorithm was ran through 10,000 iterations with a multivariate normal proposal distribution for candidate ϕ vectors with covariance matrix equals to $0.04\mathbf{I}$, which resulted in an acceptance percentage of about 22%, and we discarded the initial 2000 samples as burn-in. We used as starting values for $\phi^{(0)} = (1.0, 1.0)^T$ and the starting values for $\theta_{1:T}^{(0)}$ are sampled from a multivariate normal with mean $100\mathbf{1}$ and covariance matrix $100\mathbf{I}$. It took about 10 min to run the MCMC algorithm in an Intel Core i7 machine with 3.1 GHz and 8GB RAM.

The resulting Markov chain of sampled values is given in Figure 3. Figure 4 shows kernel density estimation of the marginal posteriors of ϕ_1 and ϕ_2 obtained from the sampled values. It can be seen that the sampled marginal posteriors have high densities around the true simulated values for ϕ_1 and ϕ_2 . The sample mean values for ϕ_1 and ϕ_2 were 0.5250 and 0.3651, respectively, and highest posterior density regions with 95% probability were (0.3546, 0.7067) and (0.1979, 0.5437), respectively. Figure 5 shows the simulated and estimated mean OD flows, with mean squared error (MSE) between simulated and estimated mean OD flows of 15.83. From these results, we can conclude that the proposed MCMC algorithm was able to estimate with good accuracy both mean OD flows and route choice parameters.

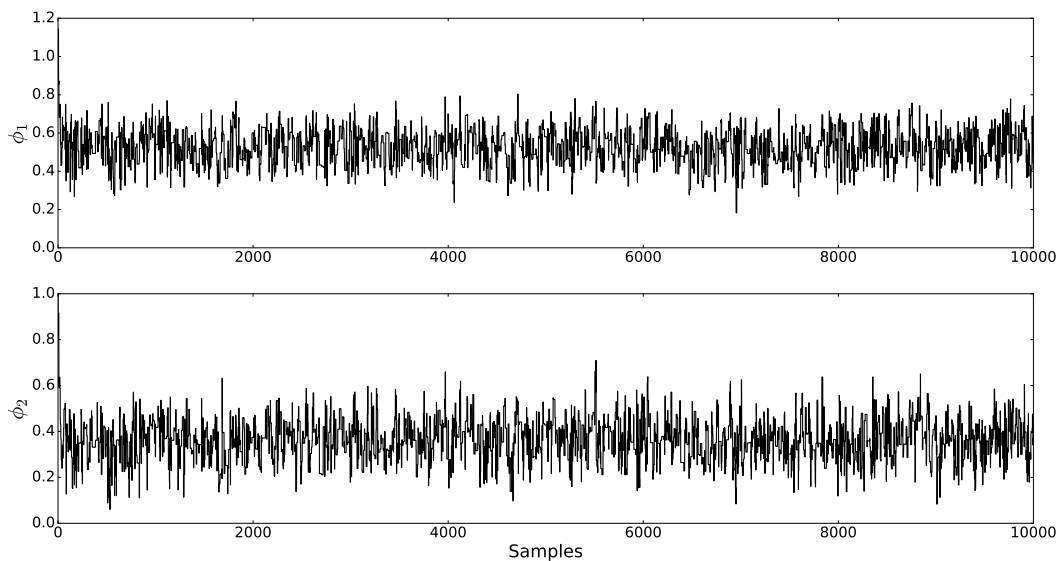


Figure 3. Markov chain for ϕ_1 and ϕ_2 .

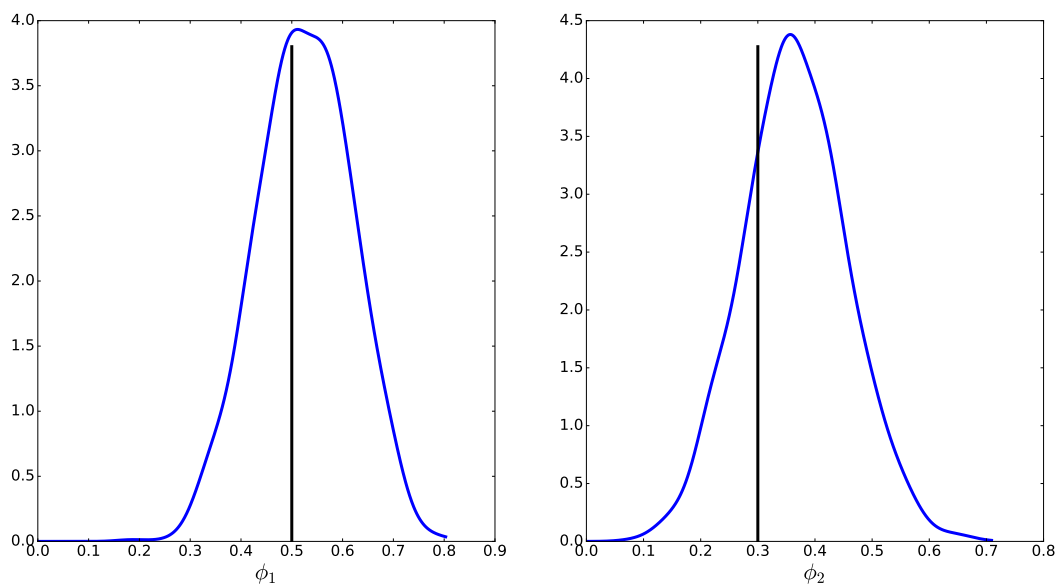


Figure 4. Kernel density estimation for the marginal posteriors of ϕ_1 and ϕ_2 . Vertical axes show probability density. Vertical bars show true simulated values.

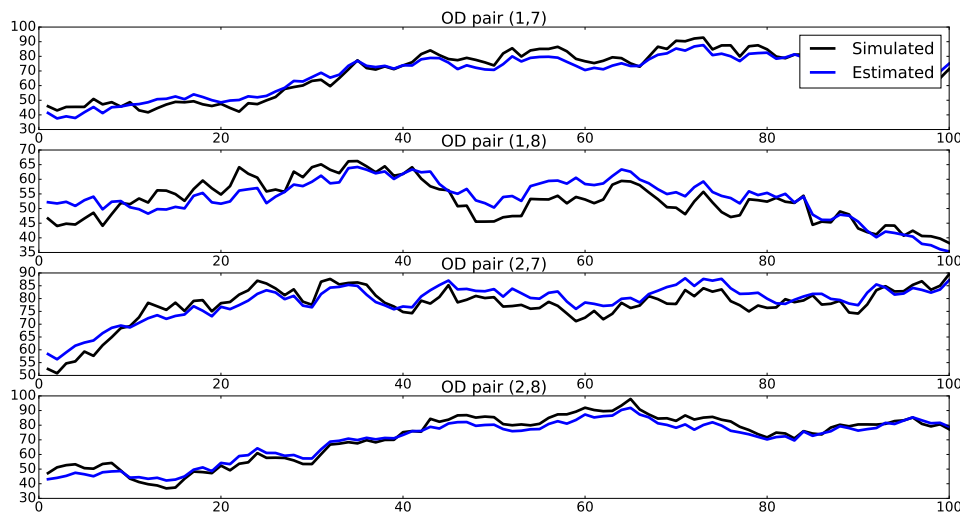


Figure 5. Simulated and estimated mean OD flows. Horizontal axes show time periods, while vertical axes show magnitude of OD flows.

3.3. Unknown Evolution Matrix

Although in theory the evolution matrix W_t in DLMs may be jointly estimated along with other parameters, it is often treated as a known parameter provided by the analyst. A convenient way of specifying the evolution matrix is through discount factors [31]. According to Equation (9), the prior covariance matrix of the OD flows at time t is given by $\bar{C}_t = C_{t-1} + W_t$, i.e., it is the posterior covariance matrix at the previous time period amplified by the evolution covariance matrix, which corresponds to the increase in uncertainty due to time. Then, we can write the prior covariance matrix as $\bar{C}_t = 1/\delta C_{t-1}$, where $0 < \delta \leq 1$, so as to $W_t = (1 - \delta)/\delta C_{t-1}$. The term δ is a discount factor. Notice that, when $\delta = 1.0$, we do not have an increase in uncertainty from time $t - 1$ to t , thus corresponding to a static model. Lower δ values correspond to higher increase in uncertainty from time $t - 1$ to time t .

We ran our proposed MCMC algorithm with varying discount factors δ , starting from covariance matrix $C_0 = 1000 I$, in order to assess its impact on the quality of the estimation of the route choice parameters and mean OD flows. In Table 2, we show the corresponding sample means $\hat{\phi}_1$ and $\hat{\phi}_2$, 95% highest posterior density (HPD) regions for ϕ_1 and ϕ_2 , and mean squared error (MSE) between simulated and estimated mean OD flows.

Table 2. Estimation results given different discount factors (HPD—95% highest posterior density region).

δ	$\hat{\phi}_1$	HPD	$\hat{\phi}_2$	HPD	MSE (OD Flows)
0.7	0.4941	[0.3265, 0.6659]	0.3340	[0.1475, 0.5007]	137.27
0.8	0.4960	[0.3178, 0.6600]	0.3380	[0.1674, 0.5288]	82.84
0.9	0.4898	[0.3121, 0.6606]	0.3313	[0.1480, 0.5057]	33.07

Figure 6 illustrates the effect of using $\delta = 0.9$ on the estimation of mean OD flows. As can be seen from the results, there was little difference in estimation quality between the cases with known and unknown evolution matrix regarding the parameters ϕ . The difference is more pronounced with relation to estimation quality of mean OD flows, which exhibited higher MSE with the evolution matrix specified by means of a discount factor.

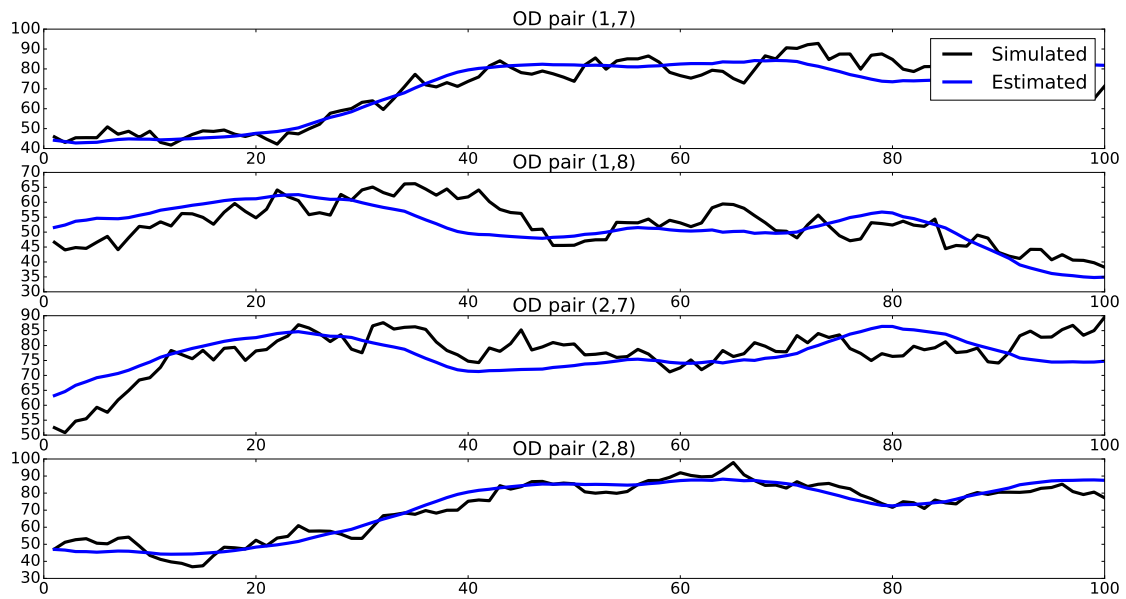


Figure 6. Simulated and estimated mean OD flows for $\delta = 0.9$. Horizontal axes show time periods, while vertical axes show magnitude of OD flows.

3.4. Observation of Traffic Volumes on Partial Links

In some real transportation networks, there are data on traffic volumes only on a few links. We consider the application of our MCMC algorithm in these settings. In this set of experiments, all parameters are equal to the simulated ones, except for the mean OD $\theta_{1:T}$ flows and the parameters ϕ of the utility model, which are estimated by the MCMC algorithm. We vary only the number of links and in which of them we have traffic volume data.

We test three cases: observations on one link, on two links, and on three links in the network in Figure 2. Links were selected so that we have a representative subset of the network. In Table 3, we show the corresponding sample means $\hat{\phi}_1$ and $\hat{\phi}_2$, 95% highest posterior density regions for ϕ_1 and ϕ_2 , and MSE between simulated and estimated mean OD flows.

Table 3. Estimation results with observation on partial links (HPD—95% highest posterior density region).

Observed Links	$\hat{\phi}_1$	HPD	$\hat{\phi}_2$	HPD	MSE (OD Flows)
1	0.6088	[0.0065, 1.1342]	0.3904	[0.0014, 0.8312]	471.67
2	0.6283	[0.3897, 0.8247]	0.4707	[0.2509, 0.6827]	256.53
9	0.5877	[0.3697, 0.7972]	0.3644	[0.1608, 0.5752]	233.34
2 and 5	0.5115	[0.3194, 0.6817]	0.3681	[0.1692, 0.5422]	246.70
1 and 9	0.5728	[0.3608, 0.7456]	0.3640	[0.1760, 0.5546]	175.43
2, 5 and 9	0.5023	[0.3241, 0.6838]	0.3323	[0.1383, 0.5039]	86.51
1, 7 and 9	0.5933	[0.3947, 0.7936]	0.3676	[0.1870, 0.5870]	59.82

In Table 3, we see that MSE of mean OD flows decreases as we observe more links. This result is in agreement with similar results of experiments with static models in the literature. The best results were obtained with data on links 1, 7 and 9, for which $MSE = 59.82$. It is noteworthy that this result is not very far from the case studied in Section 3.2, where we have traffic volume data on all links and $MSE = 15.83$. Figure 7 shows simulated and estimated mean OD flows for this case with data on links 1, 7 and 9.

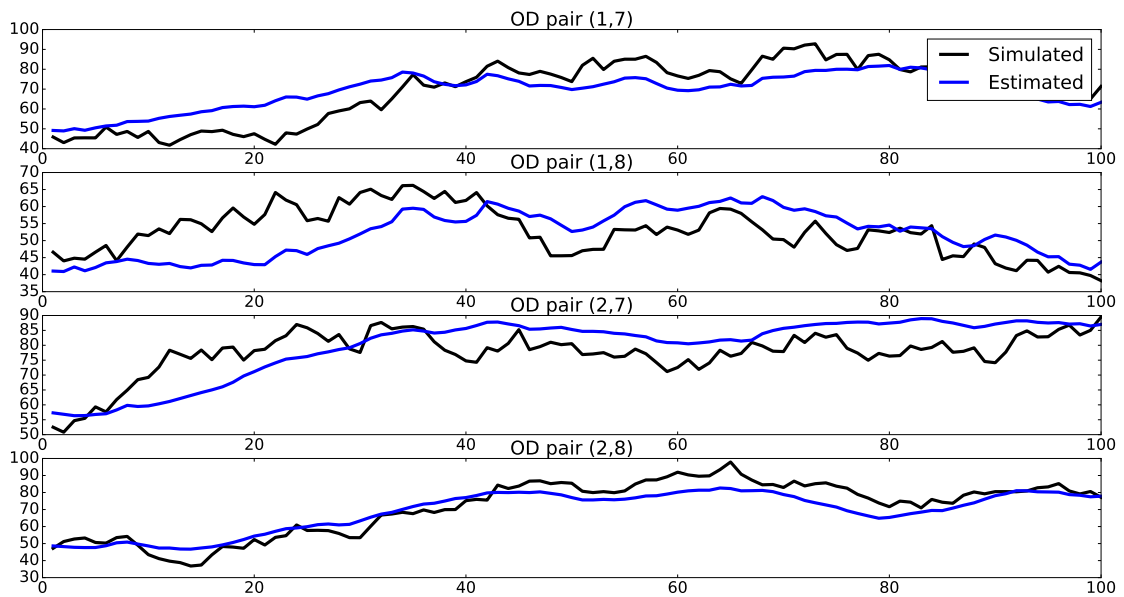


Figure 7. Simulated and estimated mean OD flows with observed traffic volumes on links 1, 7 and 9. Horizontal axes show time periods, while vertical axes show magnitude of OD flows.

Estimation quality of route choice parameters ϕ also increases with the observation of more links, as the HPD regions get tighter around the simulated values $\phi = (0.5, 0.3)^T$. Nevertheless, this effect is not so pronounced for OD flows. The best result among the tested cases was obtained when observing links 2, 5 and 9, which are links located in the central and more congested region of the network. It is also worth noting that we could estimate route choice parameters to a good accuracy by observing only link 9.

In addition, we notice that the estimation results for the case when we have traffic volumes data only on link 1 are the worst. This may be possibly due to the fact that there is no route in OD pairs (2,7) and (2,8) which includes link 1. Thus, traffic volumes on link 1 give no information on OD flows for these OD pairs. In Figure 8, we see that estimated mean OD flows remain at 100 in this case, which is our prior mean at $t = 0$.

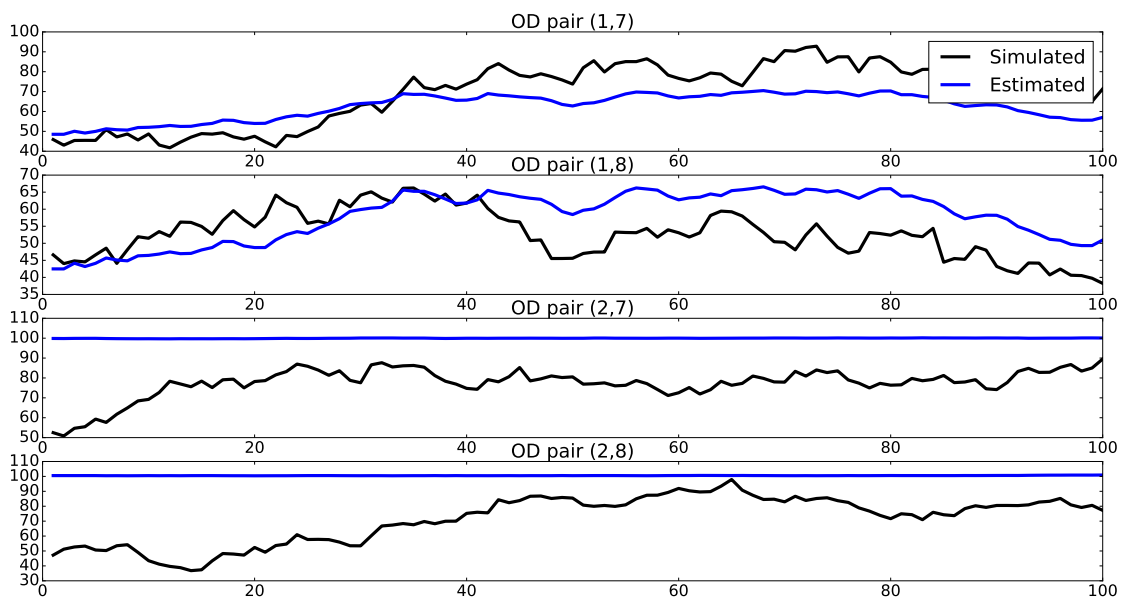


Figure 8. Simulated and estimated mean OD flows with observed traffic volumes on link 1. Horizontal axes show time periods, while vertical axes show magnitude of OD flows.

Finally, we should emphasize that we obtained good estimates when observing only a subset of links even though we used uninformative priors for both mean OD flows and route choice parameters. Static models often resort to prior knowledge, in the form of prior OD matrices, in order to “regularize” estimates when data are available only on a few links.

4. Conclusions

In this paper, we proposed a dynamic linear model for day-to-day OD flows in transportation networks. In our model, we took into account variability in OD flows, route choices and traffic volume measurements. We also modeled route choices through a utility model based on past route costs. We proposed a Markov chain Monte Carlo algorithm in order to sample from the joint posterior distribution of mean OD flows and route choice parameters by using data on traffic volumes on links and past route costs. Our model can be applied to congested networks and when data are available on only a subset of links.

We illustrated the application of the DLM and MCMC algorithm on a test network from the literature using simulated data. In our experiments, we were able to estimate with good accuracy both mean OD flows and route choice parameters using uninformative prior distributions and data on a subset of links in the test network. These are very promising results and indicate that dynamic linear modeling of day-to-day OD flows may provide a valuable tool in analyzing and estimating OD demand in transportation networks. Although we carried out computational experiments on a small network, due to the convenience of testing the model and presenting the results, we made no limiting assumptions on the size of the network when developing the model and algorithm, so that it may be applied to networks of any size.

As further extensions to our work we suggest the consideration of trend and seasonality in OD flows. This can be implemented through the definition of additional parameters such as trend and seasonal factors, whose relationship with mean OD flows can be easily established by means of the system matrix \mathbf{G}_t . Another promising research direction is the use of our proposed DLM to forecast future traffic volumes, after setting up the model with estimated route choice parameters.

Author Contributions: Conceptualization and Methodology by all the authors; Writing—Original Draft Preparation, A.R.P.-N.; Writing—Review and Editing by all the authors.

Funding: This research was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant No. 422464/2016-3.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Nomenclature

In the text and in the equations, lowercase italic letters represent scalars, while vectors are represented by bold lowercase letters and matrices by uppercase bold letters. Greek letters are in general reserved to parameters, but there are exceptions such as some covariance matrices which are represented by the Greek capital letter Σ . Sets are typed in calligraphic letters, while blackboard letters (such as \mathbb{P}) denote probability density functions. We provide below a list of symbols used in the paper:

θ_t	Vector of mean OD flows at time t ;
x_t	Vector of realized origin–destination (OD) flows at time t ;
y_t	Vector of route flows at time t ;
z_t	Vector of traffic volumes on links at time t ;
p_t	Vector of route choice probabilities at time t ;
c_t	Vector of route costs at time t ;
\mathcal{K}_i	Set of routes for OD pair i ;

ϕ	Vector of parameters in the utility equation of a route given past route costs;
π	Probability of not using a route in a given route choice set;
$\tau(\cdot)$	Performance (cost) function of a link;
v_t	Vector of random errors in the observational model at time t ;
ω_t	Vector of evolution errors in the as the dynamic model at time t ;
G_t	State evolution matrix (system matrix);
F_t	Assignment matrix at time t ;
W_t	Evolution covariance matrix of mean OD flows at time t ;
V_t	Covariance matrix of traffic volumes on links at time t ;
Σ_t	A covariance matrix at time t ;
P_t	Route choice matrix at time t ;
Δ	A link-path incidence matrix;
\bar{m}_t^v	Location vector of the prior distribution of mean OD flows at time t ;
\bar{C}_t	Covariance matrix of the prior distribution of mean OD flows at time t ;
m_t	Location vector of the posterior distribution of mean OD flows at time t ;
C_t	Covariance matrix of the posterior distribution of mean OD flows at time t ;
I	The identity matrix;
$\text{diag}(x)$	A matrix whose main diagonal is the vector x and entries off the main diagonal are all zero;
$\text{blockdiag}(X_i)$	A block-diagonal matrix composed of submatrices X_i ;
$N(\mu, \Sigma)$	Multivariate normal distribution with mean vector μ and covariance matrix Σ ;
$MN(x, p)$	Multinomial distribution with parameters x and p .

References

- De Dios Ortúzar, J.; Willumsen, L. *Modelling Transport*, 4th ed.; Wiley: Chichester, UK, 2011.
- Cascetta, E. *Transportation Systems Analysis: Models and Applications*, 2nd ed.; Springer: New York, NY, USA, 2009.
- Robillard, P. Estimating the OD matrix from observed link volumes. *Transp. Res.* **1975**, *9*, 123–128. [[CrossRef](#)]
- Nguyen, S. *Estimating an OD Matrix from Network Data: A Network Equilibrium Approach*; Technical Report; Centre de Recherche sur les Transports, Université de Montreal: Montreal, QC, Canada, 1977.
- Beckmann, M.; McGuire, C.; Winsten, C.B. *Studies in Economics of Transportation*; Yale University Press: Ann Arbor, MI, USA, 1956.
- Van Zuylem, H.J.; Willumsen, L.G. The most likely trip matrix estimated from traffic counts. *Transp. Res. Part B* **1980**, *14*, 281–293. [[CrossRef](#)]
- Cascetta, E. Estimation of Trip Matrices from Traffic Counts and Survey Data: A Generalized Least Squares Estimator. *Transp. Res. Part B* **1984**, *16*, 289–299. [[CrossRef](#)]
- Cascetta, E.; Nguyen, S. A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transp. Res. Part B* **1988**, *22B*, 437–455. [[CrossRef](#)]
- Brenninger-Göthe, M.; Jörnsten, K.O. Estimation of origin–destination matrices from traffic counts using multiobjective programming formulations. *Transp. Res. Part B* **1989**, *23B*, 257–269. [[CrossRef](#)]
- Fisk, C. On combining maximum entropy trip matrix estimation with user optimal assignment. *Transp. Res. Part B* **1988**, *22B*, 69–79. [[CrossRef](#)]
- Fisk, C. Trip matrix estimation from link counts: the congested network case. *Transp. Res. Part B* **1989**, *23B*, 331–336. [[CrossRef](#)]
- Wardrop, J.G. Some theoretical aspects of traffic research. In *Proceedings of the Institution of Civil Engineers Part II*; Institution of Civil Engineers: London, UK, 1952; Volume 1, pp. 325–378.
- Smith, M. The existence, uniqueness and stability of traffic equilibria. *Transp. Res. Part B* **1979**, *13B*, 295–304. [[CrossRef](#)]
- Yang, H.; Sasaki, T.; Iida, Y.; Asakura, Y. Estimation of origin–destination matrices from link counts on congested networks. *Transp. Res. Part B* **1992**, *26B*, 417–434. [[CrossRef](#)]
- Yang, H. Heuristic algorithms for the bilevel origin–destination matrix estimation problem. *Transp. Res. Part B* **1995**, *29B*, 231–242. [[CrossRef](#)]
- Cascetta, E.; Postorino, N. Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks. *Transp. Sci.* **2001**, *35*, 134–147. [[CrossRef](#)]

17. Sheffi, Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1985.
18. Vardi, Y. Network tomography: Estimating source-destination traffic intensities from link data. *J. Am. Stat. Assoc.* **1996**, *91*, 365–377. [[CrossRef](#)]
19. Tebaldi, C.; West, M. Bayesian inference on network traffic using link count data. *J. Am. Stat. Assoc.* **1998**, *93*, 557–573. [[CrossRef](#)]
20. Hazelton, M.L. Estimation of origin–destination matrices from link flows on uncongested networks. *Transp. Res. Part B* **2000**, *34*, 549–566. [[CrossRef](#)]
21. Hazelton, M.L. Some comments on origin–destination matrix estimation. *Transp. Res. Part A* **2003**, *37*, 811–822. [[CrossRef](#)]
22. Hazelton, M.L. Inference for origin–destination matrices: Estimation, prediction and reconstruction. *Transp. Res. Part B* **2001**, *35*, 667–676. [[CrossRef](#)]
23. Cremer, M.; Keller, H. A new class of dynamic methods for the identification of origin–destination flows. *Transp. Res. Part B* **1987**, *21*, 117–132. [[CrossRef](#)]
24. Cascetta, E.; Inaudi, D.; Marquis, G. Dynamic estimators of origin–destination matrices using traffic counts. *Transp. Sci.* **1993**, *27*, 363–373. [[CrossRef](#)]
25. Ashok, K.; Ben-Akiva, M.E. Estimation and prediction of time-dependent origin–destination flows with a stochastic mapping to path flows and link flows. *Transp. Sci.* **2002**, *36*, 184–198. [[CrossRef](#)]
26. Hazelton, M.L. Statistical inference for time varying origin–destination matrices. *Transp. Res. Part B* **2008**, *42*, 542–552. [[CrossRef](#)]
27. Parry, K.; Hazelton, M. Bayesian inference for day-to-day dynamic traffic models. *Transp. Res. Part B* **2013**, *50*, 104–115. [[CrossRef](#)]
28. Hazelton, M.L.; Parry, K. Statistical methods for comparison of day-to-day traffic models. Special issue: Day-to-Day Dynamics in Transportation Networks. *Transp. Res. Part B: Methodol.* **2016**, *92*, 22–34. [[CrossRef](#)]
29. Watling, D.P.; Cantarella, G.E. Modelling sources of variation in transportation systems: Theoretical foundations of day-to-day dynamic models. *Transp. B Transp. Dyn.* **2013**, *1*, 3–32. [[CrossRef](#)]
30. Cantarella, G.E.; Watling, D.P. Modelling road traffic assignment as a day-to-day dynamic, deterministic process: A unified approach to discrete- and continuous-time models. *EURO J. Transp. Logist.* **2015**, *5*, 69–98. [[CrossRef](#)]
31. West, M.; Harrison, J. *Bayesian Forecasting and Dynamic Models*, 2nd ed.; Springer: New York, NY, USA, 1997.
32. Särkkä, S. *Bayesian Filtering and Smoothing*; Cambridge University Press: New York, NY, USA, 2013.
33. Ben-Akiva, M.; Lerman, S.R. *Discrete Choice Analysis: Theory and Application to Travel Demand*; MIT Press: Cambridge, MA, USA, 1985.
34. Geman, S.; Geman, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *PAMI-6*, 721–741. [[CrossRef](#)]
35. Carter, C.K.; Kohn, R. On Gibbs Sampling for State Space Models. *Biometrika* **1994**, *81*, 541–553. [[CrossRef](#)]
36. Frühwirth-Schnatter, S. Data augmentation and dynamic linear models. *J. Time Ser. Anal.* **1994**, *15*, 183–202. [[CrossRef](#)]
37. Hastings, W.K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109. [[CrossRef](#)]
38. Bureau of Public Roads. *Traffic Assignment Manual for Application with a Large, High Speed Computer*; U.S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division: Washington, DC, USA, 1964.

