DAY-TO-DAY ORIGIN DESTINATION TUPLE ESTIMATION AND PREDICTION WITH HIERARCHICAL BAYESIAN NETWORKS USING MULTIPLE DATA SOURCES

Yinyi Ma
(Consulting author)
Erasmus University Rotterdam
Department of Decision and Information Science
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel. +31 10 4089783
Fax +31 10 4089010
yma@rsm.nl

Roelof Kuik
Erasmus University Rotterdam
Department of Decision and Information Science
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel. +31 10 4082019
Fax +31 10 4089010
rkuik@rsm.nl

Henk J. van Zuylen
Delft University of Technology
Section Transportation and Planning
2600 GA Delft
The Netherlands
Tel. +31 13 5056331
Fax +31 13 5056332
h.j.vanzuylen@tudelft.nl

Date of Submission: 15 Nov, 2012
Words Count: 5497 + 250* (6 figures + 2 tables) = 7497

Submitted for presentation and publication to the 92nd annual meeting of the Transportation Research Board, 13-17 January 2013, Washington D.C.
ABSTRACT
Predicting traffic demand becomes essential, either to understand the traffic state in the future or to take necessary measures for alleviating the congestion in the next time period. Usually, an origin destination matrix (OD) is used to represent traffic demand between two zones in transportation planning. Vehicles are assumed to be homogenous and the trips of each vehicle are examined separately. In fact, this traditional OD-matrix lacks of a behavioral basis and trip based model structure. There is additionally another research stream of travel activity-based research which digs into the individual travel behaviors. This stream really takes care of the trip chain for travelers. But their research scope is on the attributes of the trips, ignoring the road network. In order to link these two fields and to better predict traffic demand, we propose the concept of Origin Destination Tuple (ODT), a sequence of dependent OD pairs. With the help of advanced monitoring systems to identify and track vehicles in the road network, the additional uncertainties from ODTs can be mitigated, reducing the under-specification more specifically. We propose the Hierarchical Bayesian Networks mechanism in Gaussian Space with multi-process to get the posterior of uncertain parameters. The model includes level and trend components to make prediction of future traffic volumes. A case study demonstrates that the proposed method is feasible to predict the demand and the path flow from cameras can reduce the uncertainty in the estimation and prediction process, especially for the OD-tuples.

KEY WORDS: Origin Destination Tuple, Hierarchical Bayesian Networks, Multi-process, Demand Prediction, Multiple Data Sources.
1. INTRODUCTION

Predicting traffic demand becomes essential for policy makers to understand what may happen in the future in the road network. For instance, road congestion happens in a special time and day, such as the peak hours, bad weather and festivals. Consequently, authorities want to predict the travel volume during these special events, so as to alleviate congestion.

The concept of traffic demand is derived by trip generation in transportation planning. It addresses the issue of production and attraction between two zones, represented by origin destination (OD) pairs. The vehicles travelling on among OD pairs to fulfill demand are usually assumed to be homogenous and the trips of each vehicle are taken to be separated. However, trips of vehicles in reality are inter-related. One vehicle may appear in the time dependent OD-matrices several times a day, according to their schedules or travel plans. Commuters travel from home to work in the morning and back home in the afternoon. Trucks with multi-tasks drive every day from distribution center to a store and later to a port area, for instance. The drivers have to find a resting area after two-hour driving, according to the rule. Actually, the traditional definition of OD matrix lacks a behavioral basis and trip-based model structure (1). This setup ignores the behavioral fact that people plan ahead and choose attributes of each trip (including mode, destination, and departure time) while considering the entire trip chain, not each individual trip separately (1). For traffic engineering, the dynamic OD-matrices are mainly taken as inputs to the dynamic traffic assignment, generating traffic flow on each link. Loop detectors as a unique observation instrument on the highways are used to capture the road network information, to calibrate traffic simulation and to further estimate the OD-matrices. Thus, one reason of the ignorance of the trip chain in the OD-matrices could be argued from the anonymous loop detector data, with which vehicles cannot be identified.

Meanwhile, another research stream of travel activity-based research digs into the individual travel behaviors, such as activity schedule and travel choice. Jones et.al. (2) provide a comprehensive definition of activity analysis as: it is a framework in which travel is analyzed as daily or as multi-day patterns of behavior, related to and derived from differences in life styles and activity participation among the population. They take care of the fact that travelers have travel plans as a trip chain, such as from home to work and back (3). Survey data (4, 5) is the main information sources supporting this research. Although surveys may demonstrate some trip chains of travelers, the sparseness of survey data is an issue restricting the presentation of travel behavior. It is consequently hard to replace the OD-matrices by the patterns of behavior as the input of dynamic traffic assignment. Thus, the scope of the patterns of behavior research is normally limited to the demand side, ignoring the road network.

Due to the conceptual differences and the data capture limitation, there is no link between the OD-matrices in transportation planning and the trip chains in the behavior activity-based research, although many similarities and potential benefits have been shown (1, 2, 3). In order to fill the gap, we introduce the concept of Origin Destination Tuple (ODT) to represent traffic demand. A tuple as used in set theory is a sequence of elements. An Origin Destination Tuple is a sequence of OD pairs within a certain time period, representing the trip chain in the road network. The traditional OD pair is obviously the simplest case of the OD tuple. However, introducing ODT actually brings the extra challenge to estimate and predict the demand if measurement is based on the link flows only. It is usually assumed that demand for traffic from origin to destination act as an antecedent for the travel volume on links in the network. Since the number of OD pairs is much larger than the number of links, the estimation problem becomes underspecified (6). In this respect, the use of ODT is definitely going to deteriorate the issue.

Fortunately, nowadays has seen massive advanced monitoring systems, such as Automated Number Plate Recognition (ANPR) cameras and Bluetooth scanners, being installed along the road network, which can identify individual vehicles. GPS navigation systems installed in the vehicles record the exact routes of vehicles. These devices can deliver rich traffic information, which is potentially useful to understand and even predict the trip chain of travelers.
in the road network. The good usage of these advanced monitoring systems should decrease the uncertainty of estimating and predicting OD-tuples.

The research questions arise how to connect the macro Origin Destination Tuple and the micro activity-based level; how to predict the ODT; and how to fuse the multiple data sources to reduce the uncertainty from ODT.

In the following section, the literatures of the dynamic OD estimation and prediction, and the travel activity based model are reviewed. The methodology is presented in section 3. The case study in section 4 illustrates the methodology. And Section 5 finalizes the paper with discussions.

2. LITERATURE REVIEW

In this section, the literature of dynamic OD estimation and prediction is reviewed, considering the different methodologies and the diverse data combinations. Also, the papers related to the travel activity based model are analyzed from the concept and data aspects.

2.1 Dynamic Origin-Destination Estimation and Prediction

There are three main methods to estimate and predict dynamic OD-matrices: least squares, Kalman filtering and Bayesian method. Other methods that have drawn attention are entropy maximization (7), variation inequality (8), gradient approximation (9) and Thompson estimator (10). Most of the researchers use loop detector data, while some of the more recent researches combine such with AVI data. The way that they apply the AVI data ignores the vehicle identification, so trip chain in the road network is unaccounted for. TABLE 1 presents an overview of the literatures.

<table>
<thead>
<tr>
<th>TABLE 1 Summary of OD Estimation Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least Squares</td>
</tr>
<tr>
<td>Loop, Survey</td>
</tr>
</tbody>
</table>

Researchers apply least squares to minimize the deviation between historical OD-matrices and the estimated OD-matrices that fits the traffic flow best (11, 12, 13). Estimators based on least squares have the advantages of being mathematically rather easy to solve, especially for large problems.

Kalman filtering method is widely used to adapt model parameters to the measured characteristics of the modeled reality. Considering the state space, this method usually assumes the errors in a Gaussian space where the normality makes the computation easy and efficient. Chang and Wu (14), Ashok and Ben-Akiva (15), Dixon (16), Zhou and Mahmassani (17), and Barcelo (18) use the Kalman Filtering method to estimate and predict the dynamic OD matrices.
The Bayesian method (19, 20) is a classical way to update information by incorporating prior information in a natural manner. It can deal with all kinds of distributions, which means more realistic with less restrictions, but with a quite time-consuming computation due to the requirement of codifying prior knowledge into a prior distribution.

These three methods show great similarity in the basic mechanism if normal distributions of errors are assumed. Kalman filtering where normality dominates all the distributions is one specific case of the recursive Bayesian estimation.

2.2 Travel Activity-Based Model
Travel activity-based model integrates household activities, regional demographics and transportation networks in an explicitly time-dependent fashion (21). The activity based research enriches trip generation in the conventional four-step transportation planning process. In order to understand the individual choice behavior, researchers represent it in a discrete choice model (3, 22) and micro-simulation (21), based on survey data (3, 23), GIS (24, 25) and GPS data (23, 25).

Bowman and Ben-Akiva (3) presented a disaggregated discrete choice activity schedule. The model is designed to capture interactions among individual’s decisions throughout a 24h day by explicitly representing tours and their interrelationships in an activity pattern. They generate the time and mode specific trip matrices for prediction from an available daily survey data in the transportation system level. Wang and Cheng (24) develop a spatial-temporal data model to support activity based transport demand modeling in a GIS environment, identifying spatial and temporal opportunities for activity participation.

2.3 Summing up
Traditional OD-matrix estimation is the input of the road network with anonymous vehicles. In contrast, the focus of the travel activity-based model is on the individual behavior. Both parts of research have been left unconnected yet, due to either the research scope or the lack of individual data in the road network. But actually there is some potential benefit to be reaped for integrating two sides. Our research fills the gap between the two research streams through the concept of OD-tuple and further predicts traffic demand.

3. METHODOLOGY
In this section, the concept of Origin Destination Tuple is elaborated, and then the Hierarchical Bayesian Networks mechanism is applied to obtain the posteriori predicted Origin Destination Tuple. Considering the stochastic feature of the model and the computation efficiency, Kalman filtering in Gaussian space is used to get the mean and variance of ODT.

3.1 Origin Destination Tuple
The proposed concept of Origin Destination Tuple is an extension of Origin Destination Pair. An ODT is an ordered set of OD pairs, a number of vehicles with the same entries and exits of the road network. Clustering travelers in the same travel pattern from the geographic point of view during a certain time period actually takes the individual travel behavior into account. ODT brings the travel demand from the aggregated level to the individual behavior level. It not only addresses the issue of the number of anonymous vehicles from an origin to a destination, but specifically focuses on the trip chain of the vehicles with the same travel pattern as well.

For example, in the network of FIGURE 3, the traditional OD data for a whole day could be 6000 vehicles for 3→4 and 5000 for 6→7. Actually among these demand data, there are 500 vehicles travelling with the trips of first 3→4 and then 6→7 as an ODT. Consequently, the demand data should be with 3 OD tuples instead of 2 OD pairs: 5500 vehicles for 3→4, 4500 for 1→4, and 500 for 3→4→6→7. Additionally, the way to assign ODT to the network follow the same rules as All-or-Nothing or Stochastic-User-Equilibrium, except that the mapping from loop
detectors should be multi-counted if vehicles pass the same link more than once during a certain time interval.

In the short term, predicting ODT can help to better understand the interaction between activity driven travel and real travel behavior in the network. For the long term, transport policies such as road pricing or tolling system may refer ODT.

3.2 Hierarchical Bayesian Networks to Predict Origin Destination Tuple
Hierarchical Bayesian Networks represent the probabilistic dependencies between variables as a directed acyclic graph, where each node of the graph corresponds to a random variable and is linked by the conditional probability of that variable given the value of its parents in the graph (27). FIGURE 1 illustrates the diagram of the Hierarchical Bayesian Networks for predicting the Origin Destination Tuple with three layers: hyper-parameter, parameter and data layers. In the hyper-parameter layer, the survey data and an a priori ODT distribution are located. The variables corresponding to volumes of ODT are in the parameter layer. The observations in the data layers are link flows and path flows.

3.2.1 Hyper-Parameter Level
The setup in the hyper-parameter level is first to have the individual trip chain from survey data. After aggregating the travelers who have the same travel pattern, the aggregated trip chain can be obtained. Since the aggregated trip chain is a sample data, up-scaling is carried out to generate the a priori demand.

The individual activity trip chain $C_{r,i,t,d}$ with individual traveler $r$ having an ODT pattern $i$ at departure time interval $t$ in day $d$, is related to both scheduled time and the activity location. With the scheduled time, travelers determine the departure time, and with the activity locations, travelers decide the travel patterns. Here we assume that travel modes are either cars or trucks.
Aggregating the number of vehicles \( r \) with the same ODT pattern, \( S_{i,t,d} = \sum_r C_{r,i,t,d} \), is to get the sample demand of ODT \( i \) and at time interval \( t \) and day \( d \). Then a random growth factor \( \beta_{i,t,d} \) is used to scale up the sample demand, obtaining an a priori ODT \( U_{i,t,d} \). This random factor may consequently bring uncertainty to the further estimation, especially for the extreme case that only one or zero vehicle in a specific ODT.

\[
U_{i,t,d} = \beta_{i,t,d} \cdot S_{i,t,d}
\]  

Furthermore, the distribution of an a priori ODT \( U_{i,t,d} \) may have the feature of multiple peaks. Kernel density estimation, as one of the non-parametric approaches, is applied to smooth the density function covering the whole observation time period and to estimate density function directly from the available a priori data along time. These known data are treated as stochastic following a certain type such as Gaussian, triangular, rectangular, biweight and Epanechnikov (26). The individual kernels are mixed to generate one main kernel density to represent the density function of a random variable. Here, we assume that each kernel density of a priori ODT \( i \) at departure time interval \( t \), \( k(\hat{U}_{i,t,d}) \), follows the normal distribution. It illustrated in Equation 2, with average a priori ODT over the day \( \overline{U}_{i,t,d} \), the total number of time intervals \( H \) and smoothing parameter \( w \).

\[
k(\hat{U}_{i,t,d}) = \frac{1}{Hw} k \left( \frac{\overline{U}_{i,t,d} - U_{i,t,d}}{w} \right)
\]

**3.2.2 Parameter Level**

The ODT denoted as \( D_{i,t,d} \) is a parameter in this layer. Predicting ODT in the next day \( D_{i,t,d+1} \) uses exponential smoothing with the weights \( \alpha \). To examine the shares of the past demands for predicting the future demand, we treat the weights \( \alpha \) as unknown. Thus, the applied trend model of the dynamics is a Multi-process Model (28). Here we treat the weights constant in the demand dynamics model, but the uncertainties on their values are updated through the flow observations in the measurement model. It is discussed in the subsection 3.2.3.

\[
D_{i,t,d+1} = \sum_{z=0}^{x-1} \alpha_z D_{i,t,d-z} + \epsilon_{i,t,d+1}
\]

In order to represent this evolution process in a convenient way, we collect the demands in different days into vectors as \( \hat{D}_{i,t,d}^T = (D_{i,t,d}, D_{i,t,d-1}, \ldots, D_{i,t,d-(x-1)}) \), including the demands from the day \( d \) to the day \( d-(x-1) \), where \( x \) is the length of recursive process. The first component of the error vector \( \tilde{\epsilon}_{i,t,d+1} \) is assumed to follow a normal distribution with null mean and variance \( \sigma_{i,t,d+1} \).
Additionally, the trend model can produce the demand at day $d+k$ with the recursive Equation 4. Analytically, the expression is derived.

\[
\bar{D}_{t,d,d+k} = B_d \cdot \bar{D}_{t,d,d} + \sum_{z=1}^{k} B_d^{k-z} \begin{pmatrix} e_{t,d+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}
\]  

(5)

where,

\[
B = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{x-1} & \alpha_x \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}
\]

Further, if we want to have the future demand in the next $k$ days, the approach based on Equation 3 is as follows. Given the demand in the past days, taking today and yesterday for instance, we would like to have the expression of the demand forecasting $k$ days further in the future, $D_{d+k}$. For reasons of the expositional ease we work on the situation of a two-day case in some more detail. First, considering the fact that the eigenvalues of the evolution matrix $B = \begin{pmatrix} \alpha & 1-\alpha \\ 1 & 0 \end{pmatrix}$ are 1 and $-(1-\alpha)$, we propose two new variables $X_d$ and $\Delta D_d$ which are defined as below.

\[
X_d = D_d + (1-\alpha)D_{d-1}
\]

\[
\Delta D_d = D_d - D_{d-1}
\]

These variables manifest the two eigenvalues and so have simple dynamics as follows. It shows that one variable $X_d$ keeps constant, and the other one decreases over time with damped trend guaranteeing convergence of demand.

\[
X_{d+1} = X_d
\]

\[
\Delta D_{d+1} = -(1-\alpha)\Delta D_d.
\]

Then, the variables can be generalized to the $k$-day future where the deviations between demands flip flop over time depending on the parity of $k$.

\[
X_{d+k} = X_d
\]

\[
\Delta D_{d+k} = [-(1-\alpha)]^k \Delta D_d
\]

Based on this, we can derive the future demand in Equation 6.
\[ D_{d+k} = \frac{1}{2-\alpha} [X_{d+k} + (1-\alpha)\Delta D_{d+k}] \]

\[ = \frac{1}{2-\alpha} \left\{ D_d + (1-\alpha)D_{d-1} + (1-\alpha)[-(1-\alpha)^k (D_d + D_{d-1})] \right\} \]

\[ = \frac{1}{2-\alpha} [1 + (-1)^k (1-\alpha)^{k+1}]D_d - \frac{1}{2-\alpha} [-(1-\alpha) + (-1)^k (1-\alpha)^{k+1}]D_{d-1} \]

Further, setting \( r = \alpha - 1 \) to have

\[ D_{d+k} = \frac{1}{1-r} (1-r^{k+1})D_d + \frac{r(1-r^k)}{1-r} D_{d+1} \]

Due to \( \sum_{l=0}^{k-1} r^l = \frac{1-r^k}{1-r} \), the relation between demand in the future with \( k \) days ahead and the measured demands in the two consecutive days is expressed in Equation 7.

\[ D_{d+k} = \sum_{l=0}^{k} (\alpha-1)^l \cdot D_d - (\alpha-1) \sum_{l=0}^{k-1} (\alpha-1)^l \cdot D_{d-1} \]

\[ = D_d + \sum_{l=0}^{k-1} (\alpha-1)^l \cdot (D_d - D_{d-1}) \]

\[ = D_d + \sum_{l=0}^{k-1} (\alpha-1)^l \cdot \Delta D_d \]

In the extreme situation of predicting the demand in the far future \( D_{d+\infty} \), the predicted ODT converges to a certain number as Equation 8, if the deviation of demand \( \Delta D_d \) is small. And the weight \( \alpha \) can be in the range of \((0, 2)\).

\[ D_{d+\infty} = D_d + \frac{\alpha-1}{1-(\alpha-1)} \Delta D_d = D_d - \frac{1-\alpha}{2-\alpha} \Delta D_d \]

3.2.3 Data Level with Multiple Data Sources

For the measurement model on the data level, there are two types of flow data generated from different devices. First is the link flow, representing the traffic counts on links during a certain time period. Second is path flow, the traffic counts which pass a particular path with multiple links. Actually, the path flow may indicate the origin-destination information, which may reduce the uncertainty from the under-specification issue.

Loop detectors measuring link flows, only include anonymous counts. In principle, they cannot distinguish trip chains of vehicles. Denoting the flow observation on link \( l \) at observation time \( h \) in day \( d \) as \( V_{l,h,d} \), the relation between observed flows and ODT in the previous days is linked with the route proportion \( A_{i(l,h,d)} \). Error \( \sigma_{l,h,d} \) is assumed to be white noise.

\[ V_{l,h,d} = \sum_{i,l} A_{i(l,h,d)} \cdot D_{i,l,d} + \sigma_{l,h,d} \]
Path flows are generated by the devices which can identify vehicles. For instance, cameras track the trajectories of each vehicle, which is a rich information source to obtain the traveling routes and even ODT information. Through the identification of individual vehicles, the cameras offer the pair-wise flow, which is a part of vehicle trajectory. Denoting path flow as $V_{cc,h,d}$ with multiple passed cameras $cc$, and the route proportion as $A_{(cc,h,d|i,t)}$, the linear relation between path flow and ODT is expressed in Equation 10. Actually, here the route proportion is changing over time if the camera recognizes the vehicle is back to the road system on a secondary trip. It may bring complexity to the real-time simulation. But for the day-to-day situation, it is not an issue even if a multi-day trip chain is broken up per day. Error $\xi_{cc,h,d}$ is a white noisy, independent on the error $\zeta_{i,h,d}$ in Equation 9.

$$V_{cc,h,d} = \sum_{i,j} A_{(cc,h,d|i,t)} D_{i,t,d} + \xi_{cc,h,d} \quad (10)$$

3.3 Posterior Estimation Method in the Gaussian Space

After having the hierarchical Bayesian Model with relations among layers, the estimation and prediction of the posterior ODT is carried out with the stochastic features. We concern Gaussian space for the errors, in which an analytical approach is effective. Specifically, Kalman filtering is a special method for recursive Bayesian inference in such Gaussian space with the assumptions that all the error terms have multivariate normal distributions. It operates recursively on the streams of noisy input data to produce a statistically optimal estimate of the underlying system state.

Kalman filtering has two main updating steps: predicting and updating. The advantage of the two-step updating is to decrease computation complexity. The updating procedure is as following, illustrated in FIGURE 2. First, the observation flow at day $d$ is used to estimate demand $D_{(d|d)}$, which is based on the a-priori information. And then through dynamic prediction with a trend model, the demand at day $d+1$ can be predicted. The predicted demand predicts the flow data at the day $d+1$ as $f_{(d+1|d)}$ based on Equation 9 and 10. Further, once the observed flow at the predicted day $d+1$ is obtained, there is likely a deviation between the simulated flow $f_{(d+1|d)}$ and the observed flow $W_{(d+1)}$. Then the deviation and the prior predicted state at day $d+1$ is used to get the posterior demand. Actually, this updating procedure has the same mechanism as predicting flow directly. But the computation time of updating ODT in the Kalman filtering framework is much less than the time to update flows.

![FIGURE 2 Kalman Filter for ODT Estimation and Prediction in the Day Level.](image)
The Kalman filtering updating in our case follows four steps (28) to get the posterior ODT at day \( d+1 \).

**Step 1:** Initializing with the posterior at days \( d \) with normal distribution, having mean \( m_d \) and covariance \( \sigma_d \).

\[
(D_d | C) \sim N(m_d, \sigma_d)
\]

**Step 2:** Computing the predicted ODT with mean \( a_{d+1} \) and covariance \( R_{d+1} \) with covariance matrix of dynamics error \( \varepsilon_{d+1} \).

\[
(D_{d+1} | W_d) \sim N(a_{d+1}, R_{d+1})
\]

Where, \( a_{d+1} = B_{d+1} m_d \), \( R_{d+1} = B_{d+1} \sigma_d B_{d+1}^T + \varepsilon_{d+1} \)

**Step 3:** One-step forecast of flow data with mean \( f_{d+1} \) and covariance \( Q_{d+1} \) with covariance matrix of measurement errors \( v_{d+1} \).

\[
(W_{d+1} | W_d) \sim N(f_{d+1}, Q_{d+1})
\]

Where, \( f_{d+1} = A_{d+1} a_{d+1} \), \( Q_{d+1} = A_{d+1} R_{d+1} A_{d+1}^T + v_{d+1} \)

**Step 4:** Posterior at day \( d+1 \) updating mean \( m_{d+1} \) and covariance \( \sigma_{d+1} \).

\[
(D_{d+1} | W_{d+1}) \sim N(m_{d+1}, \sigma_{d+1})
\]

\[
m_{d+1} = a_{d+1} + X_{d+1} e_{d+1} \quad \text{and} \quad \sigma_{d+1} = R_{d+1} - X_{d+1} X_{d+1}^T Q_{d+1}
\]

Where, \( X_{d+1} = R_{d+1} A_{d+1} Q_{d+1}^{-1} \) and \( e_{d+1} = W_{d+1} - f_{d+1} \)

### 3.4 Approach to Handle the Evolution Parameters \( \alpha \) in the Multi-Process Model

The evolution parameters \( \tilde{\alpha} \) in the Multi-Process Model are constant over time but unknown. We assume to have an a priori probability distribution on a finite set of possible values for \( \tilde{\alpha} \). The set of the evolution parameters, denoted as \( \Lambda \), is discrete with \( \alpha \in \Lambda \). Given any weights \( \alpha \), the trend model in Equation 3 can be analyzed in the Gaussian space to produce sequences of prior, posterior and forecast distributions of ODT that are sequentially updated over time as flow observations are processed (28). The means and variances of the distributions all depend on the specific weight value \( \alpha \) under consideration. The inference about the ODT at day \( d+1 \) is based on the density of demand given all the available flow observation over historical time, \( p(D_{d+1} | \alpha, V_d, V_{d-1}, \ldots) \).

In order to have the density function of the weight \( \alpha \) given all the observed flow data \( p(\alpha | V_d, V_{d-1}, \ldots) \), we start with an initial a priori density \( p(\alpha | V_0) \). Information is sequentially processed to provide inference about \( \alpha \) via posterior \( p(\alpha | V_d, V_{d-1}, \ldots) \). This is sequentially updated using Bayes’ theorem as Equation 11.

\[
p(\alpha | V_d, V_{d-1}, \ldots) \propto p(\alpha | V_{d-1}, V_{d-2}, \ldots) p(V_d | \alpha, V_{d-1}, V_{d-2}, \ldots)
\]

And then to make inferences about \( V_d \) without reference to any particular value of \( \alpha \), the required unconditional density is in Equation 12.

\[
p(D_d | V_d, V_{d-1}, \ldots) = \int_{\Lambda} p(D_d | \alpha, V_d, V_{d-1}, \ldots) p(\alpha | V_d, V_{d-1}, \ldots) d\alpha
\]
4. CASE STUDY ON A15 MOTORWAY IN THE NETHERLANDS

The proposed method is tested in a real network of a part of the A15 motorway (between entry 17 and exit 15 from east to west) in the Netherlands. There are seven highway sections, four on-ramps as origins and four off-ramps as destinations. Loop detectors are installed on each highway section. Cameras are on highway section 3, 4, and 6.

The travel survey of individual trip chain in this area is obtained from Statistics Netherlands, including the departure time, travel pattern and so on. The travel patterns information is used to understand the types of trip chain of these travelers within one day. Besides the OD pairs in this network, two types of the trip chains are presented in the survey: one is in3-out4 and then in6-out7; the other is in3-out5 and then in6-out7. These two types of trip chains are the introduced Origin Destination Tuples. The 14 OD pairs in this case are the simplest case of ODT. In addition, through clustering the travelers with the same travel pattern within one day, the initial sampled demand data is derived. The expectation of the a priori demand \((m_d)\) is available, after up-scaling the sampled demand. And the variance of the a priori demand \((\sigma_d)\) is ODT-wise, assumed as one thousandth of each ODT demand to be relatively realistic. Based on this one day demand, given a random factor between 0.8 and 1.2, the mean of the initial four-day demands are available.

Three scenarios of the stationary weights \((\vec{\alpha})\) of predicting demand are designed as \((0, 0.1, 0.9), (0.25, 0.25, 0.25, 0.25)\) and \((0.2, 0.4, 0.1, 0.3)\), with the probability of 10%, 40% and 50% for each scenario. These probabilities represent the trust level of these scenarios. For instance, people believe that scenario 1 is almost impossible to happen, thus associated with the probability of 10%.

To demonstrate the behavior of the model, two tests are carried out. First, we generate demand data with OD tuples based on the stationary weights within the designed scenarios, and the second is the stationary weights beyond the designed scenarios. The first test is to show that the multi-process model can help to find out the right scenario. The second test is more realistic. In the real demand forecasting, people can design different scenarios of weights, but the real weight is unknown and may not be included in the designed ones. Thus, this test intents to examine whether the model can help to find out one scenario which leads to best predicted demand. And for both tests, the contribution of cameras is presented as well.
4.1 Test One: Stationary Weights within the Designed Scenarios

We generate the day-to-day demand of 500 days based on a certain dynamics in Equation 3 with the evolution variance of 1 ($\epsilon_{d+1}$). The stationary weights are from scenario 2 (0.25 0.25 0.25 0.25), which means that the shares of the previous four days to predict the future demand are all 0.25. With these 500-day demand data, the traffic flow of loop detectors and cameras are derived by the linear relation as in Equation 9 and 10. The measurement errors ($\nu_{d+1}$) of flow data are multivariate normal random numbers with the means of zeros and the variances of 50 and 0.1 for loop and cameras, respectively. Thus, the ratio of the evolution variance of demand and the measurement variance of loop flow is 0.02 (1/50), which means the most recent real-time estimate receives relatively small weighting eventually (17). And the ratio for camera is 10, implying that the camera flow data actually play a role with a less randomness and a high accuracy.

After running the model in the situation with cameras plus loops and the one with only loops, the probabilities of scenario 2 jump from 40% to 100% immediately and stay in 100%, as in FIGURE 4, while other scenarios will not happen although people believe them with 10% and 50% at the beginning. It indicates that the proposed method of the Hierarchical Bayesian Networks with the multi-process trend model is feasible to find out the exact scenario as designed.

![FIGURE 4 Probabilities of Three Scenarios in Testing One.](image)

Meanwhile, comparing the converge of the posterior demands, taking ODT3-4–6-7 for instance, the deviation of the posterior and true demands in scenario 2 with the solid lines in FIGURE 5 converge the fastest among the scenarios in both situations with and without cameras. FIGURE 5 also illustrates that the randomness of the deviations without cameras is larger than the one with cameras. And with only loops installed, the posterior demands of ODT3-4–6-7 with Scenario 1 and 3 even cannot converge to the true demand value because of the randomness.
FIGURE 5 Posterior Demand of ODT3-4-6-7 in 3 Scenarios With and Without Cameras.

In addition, the absolute deviations in TABLE 2 (a) and (b) between true (generated) demand and estimated demand in scenario 2, as the right scenarios to generate demand, are expectedly much lower than the other two scenarios. The summation over the difference ratios of real demand and posterior demand in scenario 2, as 0.04% and 0.05% in the situations with and without cameras, are also much lower than the rest: 1.21% and 6.38% from scenario 1; 0.35% and 1.93% from scenario 3.

Furthermore, cameras do play a significant role to predict demands. They help to increase the prediction accuracy. In general, the absolute deviations with cameras and loops are lower than the ones with only loops for all the scenarios. Especially for the ODTs identified by cameras as marked in TABLE 2 (a) and (b), the predicted demands are the same as the real demands in the situation with cameras. It is also indicated in FIGURE 5.

In a nutshell, the Hierarchical Bayesian Networks with the multi-process trend model is able to reach the weight scenario with which the 500-day demand is generated. The posterior demand with the right weight converges fastest. Cameras can help to get more accurate results, especially for the ODTs.
<table>
<thead>
<tr>
<th></th>
<th>Dev</th>
<th>See 1</th>
<th>See 2</th>
<th>See 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODT 1-4</td>
<td>1.34</td>
<td>0.10</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>ODT 1-5</td>
<td>1.23</td>
<td>0.02</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>ODT 1-7</td>
<td>1.12</td>
<td>0.05</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>ODT 1-8</td>
<td>3.83</td>
<td>0.01</td>
<td>1.12</td>
<td></td>
</tr>
<tr>
<td>ODT 2-4</td>
<td>1.92</td>
<td>0.11</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>ODT 2-5</td>
<td>2.04</td>
<td>0.05</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>ODT 2-7</td>
<td>1.70</td>
<td>0.08</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>ODT 2-8</td>
<td>5.21</td>
<td>0.00</td>
<td>1.42</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4</td>
<td>0.00</td>
<td>0.01</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5</td>
<td>0.01</td>
<td>0.00</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>ODT 3-7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>ODT 3-8</td>
<td>3.06</td>
<td>0.27</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>ODT 6-7</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>ODT 6-8</td>
<td>6.14</td>
<td>0.13</td>
<td>1.56</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4-6-7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5-6-7</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>sum(%)</td>
<td>1.21</td>
<td>0.04</td>
<td>0.35</td>
<td></td>
</tr>
</tbody>
</table>

(c) Extra Weights

<table>
<thead>
<tr>
<th></th>
<th>Dev</th>
<th>See 1</th>
<th>See 2</th>
<th>See 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODT 1-4</td>
<td>3.37</td>
<td>1.68</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>ODT 1-5</td>
<td>2.85</td>
<td>1.70</td>
<td>1.22</td>
<td></td>
</tr>
<tr>
<td>ODT 1-7</td>
<td>2.64</td>
<td>1.48</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>ODT 1-8</td>
<td>8.45</td>
<td>4.20</td>
<td>3.11</td>
<td></td>
</tr>
<tr>
<td>ODT 2-4</td>
<td>4.65</td>
<td>2.20</td>
<td>1.85</td>
<td></td>
</tr>
<tr>
<td>ODT 2-5</td>
<td>4.54</td>
<td>2.57</td>
<td>1.86</td>
<td></td>
</tr>
<tr>
<td>ODT 2-7</td>
<td>3.86</td>
<td>2.06</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>ODT 2-8</td>
<td>11.68</td>
<td>5.92</td>
<td>4.37</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4</td>
<td>0.08</td>
<td>0.03</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5</td>
<td>0.14</td>
<td>0.02</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>ODT 3-7</td>
<td>0.14</td>
<td>0.04</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>ODT 3-8</td>
<td>6.96</td>
<td>3.73</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>ODT 6-7</td>
<td>0.04</td>
<td>0.03</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>ODT 6-8</td>
<td>14.65</td>
<td>7.43</td>
<td>5.59</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4-6-7</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5-6-7</td>
<td>0.00</td>
<td>0.01</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>sum(%)</td>
<td>2.91</td>
<td>1.51</td>
<td>1.15</td>
<td></td>
</tr>
</tbody>
</table>

(d) Extra Weights

<table>
<thead>
<tr>
<th></th>
<th>Dev</th>
<th>See 1</th>
<th>See 2</th>
<th>See 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODT 1-4</td>
<td>3.73</td>
<td>2.13</td>
<td>1.66</td>
<td></td>
</tr>
<tr>
<td>ODT 1-5</td>
<td>2.80</td>
<td>1.82</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td>ODT 1-7</td>
<td>1.65</td>
<td>1.08</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>ODT 1-8</td>
<td>7.92</td>
<td>4.31</td>
<td>3.37</td>
<td></td>
</tr>
<tr>
<td>ODT 2-4</td>
<td>5.09</td>
<td>2.67</td>
<td>2.30</td>
<td></td>
</tr>
<tr>
<td>ODT 2-5</td>
<td>4.42</td>
<td>2.60</td>
<td>2.04</td>
<td></td>
</tr>
<tr>
<td>ODT 2-7</td>
<td>2.41</td>
<td>1.34</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>ODT 2-8</td>
<td>10.99</td>
<td>6.24</td>
<td>4.63</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4</td>
<td>8.96</td>
<td>5.04</td>
<td>4.04</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5</td>
<td>14.59</td>
<td>8.60</td>
<td>6.63</td>
<td></td>
</tr>
<tr>
<td>ODT 3-7</td>
<td>13.20</td>
<td>7.69</td>
<td>6.00</td>
<td></td>
</tr>
<tr>
<td>ODT 3-8</td>
<td>4.30</td>
<td>2.19</td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td>ODT 6-7</td>
<td>14.21</td>
<td>8.13</td>
<td>6.17</td>
<td></td>
</tr>
<tr>
<td>ODT 6-8</td>
<td>24.09</td>
<td>13.90</td>
<td>10.51</td>
<td></td>
</tr>
<tr>
<td>ODT 3-4-6-7</td>
<td>9.71</td>
<td>5.73</td>
<td>4.53</td>
<td></td>
</tr>
<tr>
<td>ODT 3-5-6-7</td>
<td>14.72</td>
<td>8.65</td>
<td>6.85</td>
<td></td>
</tr>
<tr>
<td>sum(%)</td>
<td>16.00</td>
<td>9.32</td>
<td>7.29</td>
<td></td>
</tr>
</tbody>
</table>
4.2 Test Two: Stationary Weights beyond the Designed Scenarios

In reality, people do not know the real weights for the demand prediction, which are most likely beyond the designed scenarios. Keeping the model settings as in the first testing, we generate 500-day demand data with an extra stationary weight of (0.1 0.1 0.7 0.1).

First, the probability of the weights in scenario 3 converges to unity, and for the rest to zeros in both situations with and without cameras. It means that the weights in scenario 3 are able to lead the posterior demands to the most likely ones.

Second, even the designed weights do not include the true weights, scenario 3 with the unity probability is able to lead the low summation of the difference ratios, in TABLE 2 (c) and (d), between the real and posterior demand: 1.15% in the situation with cameras and 7.29% without cameras. Comparing the absolute deviations of the scenario 3 in TABLE 2 (c) and (d), especially for the ODTs which cameras can identify (marked with shadow), the posterior demands can achieve almost real demand with the absolute deviation, while there are more than 4 absolute deviations in the situation with only loops.

Third, the convergence rates of the posterior demands when cameras are installed are significantly higher than the one without cameras, especially for the ODTs where cameras can identify. An example from ODT 3-4~5-6 is illustrated in FIGURE 6. The dashed line represents the posterior demand along 500 days with only loop data. It takes almost 400 days to a convergence and this converged demand has 4.53 deviations from the real demand as illustrated in TABLE 3 (d).

In a nutshell, the Hierarchical Bayesian Networks with the multi-process trend model are able to find the scenario which achieves the lowest deviations between real (generated) and posterior demand, even if the real scenario is beyond the designed ones. Cameras play a significant role in this test. The deviations in the indicated scenario with cameras are much lower than the ones without cameras installed, although they cannot reach zero deviations as in the testing one. The convergence rate in the situation with cameras is much higher than the one without cameras.

4.3 Summing up

The Hierarchical Bayesian Networks with the multi-process trend model is a feasible method to find the right weight scenario with the lowest deviation between the real and posterior demands. The path flows by cameras are very essential in the real situation, where the right weights are not in the designed ones. Camera data result in the fast convergence and low deviations.
5. CONCLUSION

There are three main contributions in our paper, which also answer the research questions in the section 1. First, we propose the new concept of Origin Destination Tuple as the sequential dependence of OD matrix, which fills the gap between the transportation modeling and the activity-based model research. In order to connect these micro and macro levels, the kernel density estimation is applied to smooth the probability density of the a priori demand.

Second, we take the advantages of the monitoring systems to identify the trip chain of vehicles. The path flow from the identification devices such as cameras, significantly decrease the uncertainty from the OD tuples, which brings more serious under-specified problem to the estimation and prediction. The case study demonstrates that the path flow leads to the more accurate prediction with the almost zero absolute deviation and also results in the fast convergence of the predicted demand during the long term.

Third, the Hierarchical Bayesian Networks with the multi-process trend model is suitable to predict demand. The method is able to find the right weight-scenario with the unity probability and generate the lowest deviation between the true and posterior demands.

ACKNOWLEDGEMENTS

We would like to thank Statistics Netherlands for funding this research.

REFERENCE


