TRAVEL TIME ESTIMATION FOR URBAN ROAD NETWORKS USING LOW FREQUENCY PROBE VEHICLE DATA

Erik Jenelius
Corresponding author
KTH Royal Institute of Technology
Department of Transport Science
Email: erik.jenelius@abe.kth.se
Teknikringen 72, SE-100 44 Stockholm, Sweden
Tel: +46 8 790 8032
Fax: +46 8 21 2899

Mahmood Rahmani
KTH Royal Institute of Technology
Department of Transport Science
Email: mahmood.rahmani@abe.kth.se

Haris N. Koutsopoulos
KTH Royal Institute of Technology
Department of Transport Science
Email: haris.koutsopoulos@abe.kth.se

14 July, 2012

ABSTRACT

The paper presents a statistical model for urban road network travel time estimation using low frequency GPS probes as observations, where the vehicles typically cover multiple network links between reports. The network model separates trip travel times into link travel times and intersection delays and allows correlation between travel times on different network links based on a spatial moving average (SMA) structure. The observation model presents a way to estimate the parameters of the network model, including the correlation structure, through low frequency samplings of vehicle traces. We combine link-specific effects with link attributes (speed limit, functional class, etc.) and trip conditions (day of week, season, weather, etc.) as explanatory variables. This makes travel time estimation with sparse probe vehicle data feasible also in areas with very few observations. The approach also reveals the underlying factors behind spatial and temporal variations in speeds, which is useful for traffic management, planning and forecasting. The model is estimated using maximum likelihood and the statistical significance of the results is assessed by reporting standard errors. The model is applied in a case study for the network of Stockholm, Sweden. We find that link attributes and trip conditions (including rain and snow) have significant effects on travel times and that there is significant positive correlation between segments. The case study highlights the potential of using sparse probe vehicle data for monitoring the performance of the urban transport system.

1. INTRODUCTION

Many urban road transport systems today experience increasing congestion that threatens the environment and the transport efficiency. To tackle these problems, knowledge about traffic conditions is critical at many levels of traffic management and transport policy. Through information and personalized advice, individuals and transporters can plan their trips more accurately and increase the efficiency of the system. For traffic management, travel time information at the segment level can reveal problematic locations where new or revised traffic control schemes may be introduced to increase performance. For transport policy, network-wide travel time information provides input for travel demand forecasting and impact assessments of policy instruments such as congestion charges.

There are a number of well established technologies for collecting travel time data, including loop detectors, automatic vehicle identification (AVI) sensors and probe car data. Loop detectors and AVI sensors have the merit that they, once installed, continuously record every vehicle passing the monitored road section. However, the share of segments in the network equipped with these sensors is typically low and not representative of the urban network as a whole, which leaves the traffic conditions in most of the network unknown. Dedicated probe vehicles, meanwhile, have been used in the past to collect travel time and other data at designated segments in the network. However, due to cost considerations the number of traffic studies with probe vehicles is typically small and the number of vehicles involved very low. Hence, they can only cover a limited number of routes for a limited duration of time.

Most recently, GPS devices, already installed for other purposes in vehicle fleets (e.g., taxis, commercial vehicles, service vehicles, etc.) or smartphones, provide a new type of traffic sensor. These opportunistic sensors have a great potential for provision of data for traffic management applications. Unlike stationary sensors, they can collect travel time data for any part of the network where equipped vehicles move. Unlike designated probe cars, they can continuously collect data for any time and day that equipped vehicles are active. However, despite their advantages, an important disadvantage for the widespread use of these data is that more advanced and sophisticated methods are needed to process the data and generate useful information, compared to traditional sensors (Leduc, 2008).

A number of limitations make the use of opportunistic probe vehicle data challenging. First, the penetration rate is still typically low, which means that the collected information represents only a small sample of the full traffic state of the system. Furthermore, there may be systematic differences between the equipped vehicles and the overall population (van Aerde et al., 2003). Second, the accuracy and the frequency of the position reports, while adequate for the original purpose of the GPS device, may be of low quality when used for travel time estimation.

The literature on travel time estimation and forecasting using GPS sensors has grown in recent years as the technology has become more available. Most papers, however, have dealt with high frequency data, (e.g., Zou et al. 2005; Work et al., 2008), which eliminates many of the challenges of interest here. Low sampling frequency creates difficulties in inferring the true path of the vehicle between two position reports, which may involve a considerable number of network segments (Rahmani and Koutsopoulos, 2012; Miwa et al., 2008). It also becomes difficult to identify the fraction of the travel time spent on each individual segment, and different local methods have been developed for this task (Miller et al., 2010; Hellinga et al., 2008, Zheng and van Zuylen, 2012). For short segments in particular, there may be few observations available to estimate the travel time distribution under the conditions of interest.

A probabilistic model of travel times through the arterial network based on low frequency taxi GPS probes is presented in (Hunter et al., 2009). The model takes into account that the path between two consecutive position reports may contain multiple segments, but not that a reported position may be in the interior of a segment. The authors formulate a maximum likelihood problem to estimate the segment travel time distributions based on the set of observed route travel times. As the segment travel times are typically not directly observed, a simulation based EM estimation algorithm is proposed. The authors assume that the travel times on different segments are independent and briefly report estimation results using normal and log-normal distributions.

A development of the approach in Hunter et al. (2009), more clearly aimed towards travel time forecasting, is presented in Hofleitner et al. (2012). The model assumes that each segment can be in one of two possible states (congested or uncongested), each with its own conditional, independent normal travel time distribution. The transitions between states among neighboring segments are modeled as a dynamic Bayesian network model. The unobserved state and transition probabilities and the travel time distribution parameters are estimated in a simulation based EM approach.

Another approach, using low frequency GPS data from ambulances, is presented in (Westgate et al., 2011). In the paper, path inference and travel time estimation are performed simultaneously using a Bayesian approach. The framework makes use of instantaneous speed information reported by the vehicles. The travel times on the road links are assumed to be independent and log-normally distributed, and the parameters are estimated using Markov chain Monte Carlo methods.

The previous approaches to travel time estimation using low frequency GPS data all assume that the travel times experienced by a vehicle on consecutive road links are independent (Hunter et al., 2009; Westgate et al., 2011; Hofleitner et al., 2012). In practice, however, congestion, weather conditions etc. mean that travel times are often positively correlated across links (Bernard et al., 2006; de Fabritiis et al., 2008). That is, if on a given day the travel time on one segment is unusually high (or low), then the travel time on nearby segments will likely also be unusually high (low). Spatial and temporal dependencies between segment travel times have been incorporated for forecasting purposes using STARMA (Spatio-Temporal Auto-Regressive Moving Average) and similar models (Min et al., 2009; Herring et al., 2010). These models, however, were applied to stationary sensor data and not probe vehicle data.

This paper presents a statistical model for urban road network travel time estimation using observations from low frequency GPS probes. The purpose of the model is to estimate the travel time for any route between any two points in the network under specified trip conditions; we are interested in both the mean travel time and the variability, considering that link travel times along the route may be correlated. The basic idea is to increase the reliability of the estimation by utilizing all observed vehicle trajectories that provide information about some part of the route in the estimation and not only such observations that cover the entire route, of which there may be few or none.

The presented approach extends previous work on travel time estimation using probe vehicle data by considering the effect of explanatory variables on travel times. For the network components themselves this may include attributes such as speed limit, number of lanes, functional class, bus stops, traffic signals, stop signs, left turns, etc. We also consider the effects of the conditions for the trip, such as weather, time of day, weekday, time of year, and so on. This approach is attractive for at least three reasons. First, it allows us to identify the underlying causes for the variability in speeds between links and points in time. This is important for system management and planning, where one needs to know the relationships between possible instruments and network travel times in order to improve the mobility and accessibility in the transport system.

Second, the statistical approach reduces the number parameters to estimate and allows us to estimate travel times with low frequency probe data, even in areas with very few observations. This aspect has not been discussed much in previous work (it is handled implicitly in Bayesian approaches) but proved essential in practical applications of the probe vehicle data source used in this paper. Third, the integration of trip conditions in the model makes it possible to extend the historical estimation to prediction of future travel times based on forecasted conditions.

The methodology further extends previous work by incorporating the correlation experienced by a driver traversing the segments sequentially on a trip. The statistical model consists of two parts: a model for the travel times on network segments, and a model for the probe vehicle observations. The network model assumes that segment travel times are distributed multivariate normal according to a spatial moving average (SMA) structure. The observation model takes into account that the correlation between segments is incorporated not only in each probe vehicle travel time observation, but in the entire sequence of observations from the same vehicle trajectory. We show that the observations are distributed multivariate normal and we derive the analytical likelihood function of the observations expressed in the parameters of the network model, which may thus be estimated. We are able to assess the confidence of the estimates by reporting standard errors.

The paper is organized as follows. The network model is presented in Section 2 and the observation model is presented in Section 3. In Section 4 we discuss some practical considerations regarding the specification and estimation of the model. This is followed by a description of a real-world application in Section 5 and a concluding discussion in Section 6.

2. NETWORK MODEL

2.1 Network travel time components

The travel time of a trip is assumed to consist of two components:

1. Running travel time along links,
2. Delay at intersections and traffic signals (turns).

To begin with, we define a *link* to be the road section between two adjacent intersections or traffic signals (traffic signals are not always located at intersections). A link may be divided into one or more *segments*, where each segment is a part of one specific link. While the links are largely determined by the inherent network structure, we are free to split each link into as many segments as suitable for the analysis. We let N_S and N_L denote the total number of segments and links in the network, respectively. The relationship between segments and links can be described by an $N_S \times N_L$ matrix $\mathbf{S}$, so that element S_{sl} is 1 if segment s belongs to link l and 0 otherwise. Since a segment can only belong to a single link, we must have $\sum_l S_{sl} = 1$ for all s, and we can identify the link $l(s)$ of segment s as the link such that $S_{s,l(s)} = 1$.

In this model the speed of a vehicle can vary between segments but is assumed to be constant along each segment. The travel time on a segment s can always be decomposed as the length of the segment, ℓ_s, multiplied with the inverse speed or *travel time rate* X_s. As described in the following subsections, the travel time rate may depend on observed and unobserved properties of the segment and conditions for the trip.

The second travel time component of a trip is intersection and traffic signal delay. We define the *turn* t as the movement from a link $l_{t,1}$ to the downstream link $l_{t,2}$ and let N_T denote the total number of turns in the network. A turn can thus be defined by the pair $(l_{t,1}, l_{t,2})$. The turn t is assumed to give a travel time penalty h_t. Factors that would influence the magnitude of h_t may include the type of traffic control in the intersection, whether it involves a left or a right turn, time of day, etc.

Conceptually, turns can be seen as links having zero length, as illustrated in Figure 1. This means that the probability of a vehicle reporting its position on a turn link is zero. While one can readily incorporate both types of components in a single set of variables, in this paper we will use separate sets of variables for link running times and turn penalties for ease of exposition.

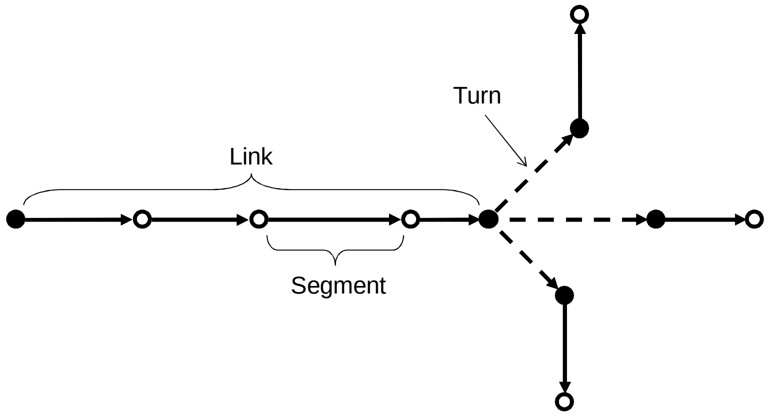

Figure 1: Illustration of the three types of network components: Links, segments and turns.

The segment travel time rates X_s, $s = 1, \dots, N_S$, are modeled as stochastic variables, in general not independent. We are interested in estimating both mean travel times and the variability around the mean values. Assuming that the mean value is finite, we can write $X_s = g_s + \nu_s$, where g_s is the mean travel time rate and ν_s is a stochastic component with $E[\nu_s] = 0$ capturing the variability around the mean.

For compact notation, it is convenient to introduce the $N_S \times 1$ vector $\boldsymbol{X}$ with elements X_s. We then have

$$\boldsymbol{X} = \mathbf{g} + \boldsymbol{\nu}, \tag{1}$$

where $\mathbf{g}$ is the vector of mean segment travel time rates, and $\boldsymbol{\nu}$ is the vector of zero-mean stochastic terms. The model assumes that the stochastic components of the segment travel time rates $\boldsymbol{\nu}$ follow a multivariate normal distribution according to a covariance structure defined in Section 2.3.

The turn penalties can be treated as deterministic or stochastic variables. In reality the delay at an intersection certainly varies according to unobserved changes in traffic flows, signal cycles etc. For estimation purposes, however, our practical applications show that treating turn delays as stochastic makes it difficult to separate the total travel time variability into variability in running travel times and in turn delays. Hence, we treat turn delays as deterministic travel time penalties in the model. The turn penalties h_t are collected in the $N_T \times 1$ vector $\mathbf{h}$.

2.2 Mean structure

The vectors $\mathbf{g}$ and $\mathbf{h}$ may be further expressed as functions of a number of factors with associated parameters to be estimated from data. The parametric structure should reflect the way that different factors affect travel times, while also allowing for convenient and efficient estimation. The structure can also be used to partition the segments into larger groups to ensure that all parameters can be identified through the available observations. We divide the explanatory variables for the mean segment travel time rates $\mathbf{g}$ into two different categories: *segment characteristics* that vary across the network and *trip conditions* that vary in time.

The explanatory factors capturing the segment characteristics could include regulatory factors such as speed limit and classification, link length, nearby land use and fixed effects for specific segments. In the model the N_B different attributes are collected in the $N_S \times N_B$ design matrix $\mathbf{B}$. The baseline segment travel time rates are then $\mathbf{B}\boldsymbol{\beta}_B$, where $\boldsymbol{\beta}_B$ is an $N_B \times 1$ parameter vector to be estimated. Note that the segments can be modeled as having fully distinct means, without any other explanatory variables, by setting $\mathbf{B}$ equal to the $N_S \times N_S$ identity matrix.

The observed trip conditions are assumed to act as a multiplier to the baseline segment travel time rates. Thus, a certain trip condition, say, a trip during vacation months, multiplies the baseline travel time rates $\mathbf{B}\boldsymbol{\beta}_B$ on all segments according to a certain percentage. Other relevant trip attributes could include temporal variations within the day, week and year, weather conditions, etc. For a given trip, the attributes are collected in the $1 \times N_o$ design vector $\mathbf{o}$. The multiplier for the trip conditions is then $(1 + \mathbf{o}\boldsymbol{\beta}_o)$, where $\boldsymbol{\beta}_o$ is an $N_o \times 1$ parameter vector to be estimated. In total, the mean segment travel time rates can be written as

$$\mathbf{g} = (1 + \mathbf{o}\boldsymbol{\beta}_o) \cdot \mathbf{B}\boldsymbol{\beta}_B. \tag{2}$$

The turn penalties $\mathbf{h}$ are assumed to be described by a corresponding structure among the explanatory variables. The explanatory factors for the turn penalties, which may include indicators for traffic signals, left/right turns, congestion etc., are collected in the $N_T \times N_E$ design matrix $\mathbf{E}$. We make the simplifying assumption that the trip conditions influence the turn penalties in the same way as the travel time rates. Thus, with the $N_E \times 1$ parameter vector $\boldsymbol{\beta}_E$, we have

$$\mathbf{h} = (1 + \mathbf{o}\boldsymbol{\beta}_o) \cdot \mathbf{E}\boldsymbol{\beta}_E. \tag{3}$$

The model can be extended to also allow the impact of trip conditions to differ between segments or turns. This would be appropriate, e.g., if network-coded information is used about traffic incidents or construction works that do not cover the whole period of observations, or if temporal variations are known to be more dominant in some parts of the network than others.

2.3 Variance structure

2.3.1 Variance components

The stochastic components $\boldsymbol{\nu}$ represent the variability in segment travel time rates due to unobserved heterogeneity in traveler characteristics, traffic conditions and local network characteristics. For our purposes, we consider *links* between intersections and traffic signals to be a more intuitive and convenient unit for representing travel time variability than segments, especially when considering the correlation between units. The travel time variability is thus modeled at the link level; travel time rates on different segments in a link are allowed to differ in means but are assumed to have the same variability around the mean, which implies that segment travel time rates are perfectly correlated within the link. The stochastic component of link l is denoted u_l, and we have $\nu_s = u_{l(s)}$, or in vector notation,

$$\boldsymbol{\nu} = \boldsymbol{S}\boldsymbol{u}. \tag{4}$$

To begin with, each link has an independent stochastic travel time rate component ϵ_l with zero mean and variance σ_l^2, which captures the variability in travel time rates that originates on the particular link. Independence implies that $\mathrm{E}[\epsilon_l \epsilon_{l'}] = 0$ for $l \neq l'$. Using vector notation, we have the $N_L \times 1$ vector of independent stochastic components $\boldsymbol{\epsilon}$ with zero means and variances $\boldsymbol{\sigma}^2$. Similarly to the means, the variances $\boldsymbol{\sigma}^2$ may be decomposed into a number of explanatory factors with associated parameters to be estimated from data. We assume again that the explanatory variables can be divided into two different categories: static link characteristics and dynamic trip conditions.

The link characteristics may include geometric properties and fixed effects for specific links or groups of links. The N_U variance components are represented by the $N_L \times N_U$ design matrix $\mathbf{U}$. The baseline link variances are then $\mathbf{U}\boldsymbol{\sigma}_U^2$**,** where $\boldsymbol{\sigma}_U^2$ is an $N_U \times 1$ parameter vector to be estimated. Note that the most simple model formulation would be that all links share a single variance parameter σ^2, in which case $\mathbf{U}$ is an $N_L \times 1$ vector of ones. In the other extreme, each link may have an individual variance parameter σ_l^2, in which case $\mathbf{U}$ is the $N_L \times N_L$ identity matrix.

Further, the observed travel conditions for the trip may impact the travel time variances as well as the means. Relevant trip attributes may be similar as for the mean travel time rates, but the impact of each attribute may be different. For a given trip, the attributes are collected in the $1 \times N_p$ design vector $\mathbf{p}$. The multiplier for the trip conditions is then $\left(1 + \mathbf{p}\boldsymbol{\beta}_p\right)^2$, where $\boldsymbol{\beta}_p$ is an $N_p \times 1$ parameter vector to be estimated. Note that the parameters actually capture the effect on the square root of the variance, i.e., the standard deviation, which is convenient since it has the same dimension as the mean. The variances of the independent stochastic components $\boldsymbol{\epsilon}$ are thus in total

$$\boldsymbol{\sigma}^2 = (1 + \mathbf{p}\boldsymbol{\beta}_P)^2 \cdot \mathbf{U}\boldsymbol{\sigma}_U^2. \tag{5}$$

2.3.2 Covariance structure

In general, the segment-level covariance matrix can be obtained from the link-level stochastic components as

$$\boldsymbol{\Omega} = \mathrm{E}[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}] = \mathrm{E}[\boldsymbol{S}\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}]. \tag{6}$$

In the special case where links are assumed to be independent, i.e., when $\boldsymbol{u} = \boldsymbol{\epsilon}$, the segment-level covariance matrix is obtained from (5) and (6) as

$$\boldsymbol{\Omega} = (1 + \mathbf{p}\boldsymbol{\beta}_P)^2 \cdot \sum_{n=1}^{N_U} \sigma_{U,n}^2 \, \mathbf{S}\mathbf{U}_n\mathbf{S}^{\mathrm{T}}, \tag{7}$$

where $\sigma_{U,n}^2$ is the nth link variance component and $\mathbf{U}_n$ is the diagonal matrix with the nth column of $\mathbf{U}$ along its diagonal.

In reality, correlations between link travel time rates may arise due to common characteristics in unobserved traffic movements and congestion, geometric properties etc. Our model allows travel time rates to be correlated between links. Note that we are considering the correlation faced by drivers traversing the links sequentially during a trip, as opposed to the correlation at a given instant in time.

To model the covariance between links we adapt an approach from spatial econometrics to an urban road network setting (LeSage and Pace, 2009; Cheng et al., 2011). The general approach is to specify the structure for how the independent stochastic components $\boldsymbol{\epsilon}_l$ interact to determine the total stochastic travel time rate components $\boldsymbol{v}$. A common model in spatial econometrics is the spatial error model (SEM) (Anselin, 1988). In the SEM, the stochastic component u_l of each link is expressed as the independent term ϵ_l plus a linear combination of the stochastic components of the other links $u_{l'}$, $l' \neq l$. The relative dependence of link l on link l', denoted $w_{ll'}$, is specified by the analyst, while the overall magnitude of the covariance is captured by a parameter ρ to be estimated. We require that $w_{ll} = 0$ for all l. The total stochastic component of link l is thus

$$u_l = \epsilon_l + \rho \sum_{l' \neq l} w_{ll'} u_{l'}. \tag{8}$$

Introducing the $N_L \times N_L$ weight matrix $\mathbf{W}$ with elements $w_{ll'}$ and moving the dependent stochastic components to the left-hand side, the structure can be written in vector notation as $\boldsymbol{u} = (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\epsilon}$, which can be expanded as $\boldsymbol{u} = (\mathbf{I} + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \cdots)\boldsymbol{\epsilon}$. Thus, even if there is no direct influence from one link on another according to $\mathbf{W}$, dependence will exist through common neighbors, common neighbors of neighbors, etc. In the end, correlation exists between every pair of links if the network is connected. Because of this property, the SEM is typically interpreted as suitable for systems that have converged to some kind of equilibrium or steady state over time (LeSage and Pace, 2009). This may not be a good characterization of urban link travel times, where we expect correlation to be more directly and locally determined by the dynamic traffic flows connecting pairs of link, as opposed to higher-order network topological effects. In this paper, therefore, we instead employ a *spatial moving average* (SMA) specification (Hepple, 2004). In the SMA model the stochastic component u_l is determined directly by the independent components of itself and other links,

$$u_l = \epsilon_l + \rho \sum_{l' \neq l} w_{ll'} \epsilon_{l'}, \tag{9}$$

or in matrix notation, $\boldsymbol{u} = (\mathbf{I} + \rho\mathbf{W})\boldsymbol{\epsilon}$. Of course, we do not need to use the same weights $\mathbf{W}$ as we would have in the SEM above. By specifying $\mathbf{W}$, we have explicit control of the influences between links. This is important since our practical applications show that the structure of $\mathbf{W}$ must be defined with much care in order to represent the dependencies between link travel time rates properly and allow a meaningful estimation of the correlations between links. As shown in Section 4, there are also computational advantages of the SMA model when using probe vehicle observations for the estimation.

We can extend the SMA model to include multiple weight matrices $\mathbf{W}_i$, $i = 1, \dots, N_\rho$, with each matrix representing a separate dimension of spatial dependence (Hepple, 2004). This more general model is useful when we believe that there are multiple factors that contribute to correlations and are distributed differently in the network. The structure for the stochastic travel time components is then

$$\boldsymbol{u} = \left(\mathbf{I} + \sum_{i=1}^{N_\rho} \rho_i \mathbf{W}_i\right)\boldsymbol{\epsilon}. \tag{10}$$

With the (extended) SMA model (10), the covariance between two links l_1 and l_2 is

$$\mathrm{E}\left[u_{l_1} u_{l_2}\right] = \sum_{i=1}^{N_\rho} \rho_i \left(w_{i,l_2,l_1}\sigma_{l_1}^2 + w_{i,l_1.l_2}\sigma_{l_2}^2\right) + \sum_{i=1}^{N_\rho}\sum_{j=1}^{N_\rho} \rho_i\rho_j \sum_{l' \neq \{l_1,l_2\}} w_{i,l_1,l'} w_{j,l_2,l'} \sigma_{l'}^2. \tag{11}$$

As can be seen, there is a first-order term that arises from the direct influences between the links, and a second-order term that arises from influences through common neighbors in the different dimensions. We can now formulate the full covariance matrix for the network segments, denoted $\boldsymbol{\Omega}$. Inserting (10) into (6), the covariance matrix is obtained as

$$\mathbf{\Omega} = (1 + \mathbf{p}\boldsymbol{\beta}_P)^2 \cdot \sum_{n=1}^{N_U} \sigma_{\tilde{U},n}^2 \left(\mathbf{S}\mathbf{U}_n\mathbf{S}^\mathrm{T} + \sum_{i=1}^{N_\rho} \rho_i \left(\mathbf{S}\mathbf{W}_i\mathbf{U}_n\mathbf{S}^\mathrm{T} + \mathbf{S}\mathbf{U}_n\mathbf{W}_i^\mathrm{T}\,\mathbf{S}^\mathrm{T} \right) + \sum_{i=1}^{N_\rho}\sum_{j=1}^{N_\rho} \rho_i\rho_j \mathbf{S}\mathbf{W}_i\mathbf{U}_n\mathbf{W}_j^\mathrm{T}\,\mathbf{S}^\mathrm{T} \right). \tag{12}$$

It follows from the covariance matrix that the correlation between two segments is independent of the trip conditions, since the factor $(1 + \mathbf{p}\boldsymbol{\beta}_P)^2$ enters both the covariance and the variances and cancels out. In the next section we show how the covariance between segments is manifested in probe vehicle travel time observations.

The SMA model (10) does not take into account that there may be dependencies in travel times due to unobserved variations between vehicles or drivers, or perhaps more importantly, between different days. It is straightforward to extend the model to capture such differences by including a stochastic component at the day level or vehicle level.

3. OBSERVATION MODEL

The travel time measurements considered in this paper consist of sparsely sampled vehicle trajectories through the network obtained from GPS devices or similar sensor technologies. In general, GPS location measurements are associated with errors. Here we assume that the most likely network location corresponding to each GPS measurement, as well as the path (i.e., the sequence of network segments) taken between each pair of consecutive measurements, have been determined by a map-matching and path inference process (Rahmani and Koutsopoulos, 2012). A basic observation then consists of

1. a vehicle identification number,
2. a pair of time stamps τ_1, τ_2,
3. a path representing the trajectory of the vehicle between the two time stamps, involving a sequence of segments $(s_1, \ldots, s_n)$ and two offsets δ_1, δ_2 specifying the vehicle locations at times τ_1, τ_2 in relation to the upstream nodes of the first and last segments of the path, respectively.

The concepts are illustrated in Figure 2. Note that the sequence of segments define two corresponding sequences of links $\big(l(s_1), \ldots, l(s_n)\big)$ and turns $(t_1, \ldots, t_m)$, respectively, according to the network model.

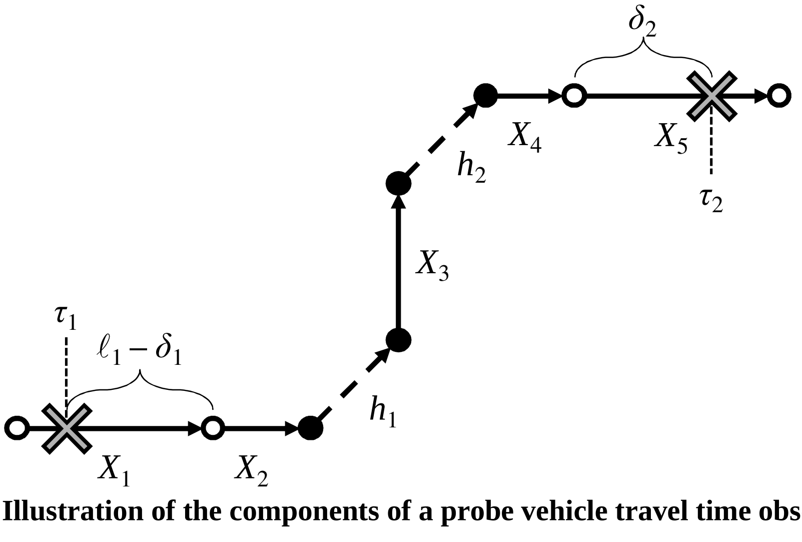

Figure 2: Illustration of the components of a probe vehicle travel time observation.

For a given travel time observation $y = \tau_2 - \tau_1$, we define d_s as the distance traversed on segment s and a_t as equal to 1 if turn t was undertaken and 0 otherwise. With ℓ_s denoting the length of segment s, we have

$$d_s = \begin{cases} \delta_2 - \delta_1 & \text{if } n = 1,\ s = s_1, \\ \ell_s - \delta_1 & \text{if } n > 1,\ s = s_1, \\ \ell_s & \text{if } n > 1,\ s = s_2, \dots, s_{n-1}, \\ \delta_2 & \text{if } n > 1,\ s = s_n, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

We can then write

$$y = \sum_{s=1}^{N_S} d_s X_s + \sum_{t=1}^{N_T} a_t h_t. \tag{14}$$

A foundation for our model is thus that a probe vehicle travel time observation is a linear combination of the segment travel time rates and the turn delays. Furthermore, under the assumption of multivariate normal segment travel time rates and deterministic turn delays, the observed travel time is normal. We may write $y = \mu + \eta$, where μ is the mean travel time given by

$$\mu = \mathrm{E}[y] = \sum_{s=1}^{N_S} d_s g_s + \sum_{t=1}^{N_T} a_t h_t, \tag{15}$$

that is, the sum of the mean segment travel times plus the turn delays. Let $d_l = \sum_s S_{sl} d_s$ be the distance traveled on link l. The zero-mean stochastic term η is distributed normal and given by

$$\eta = \sum_{s=1}^{N_S} d_s v_s = \sum_{l=1}^{N_L} d_l u_l. \tag{16}$$

The variance of η, and hence of the observation y, is calculated from the variances and covariances of the traversed links as

$$\begin{aligned} \mathrm{Var}[\eta] = \sum_{l=1}^{N_L} d_l^2 \sigma_l^2 + \sum_{i=1}^{N_\rho} \rho_i \sum_{l_1=1}^{N_L} \sum_{l_2=l_1}^{N_L} d_{l_1} d_{l_2} \left(w_{i,l_2,l_1} \sigma_{l_1}^2 + w_{i,l_1.l_2} \sigma_{l_2}^2 \right) \\ + \sum_{i=1}^{N_\rho} \sum_{j=1}^{N_\rho} \rho_i \rho_j \sum_{l_1=1}^{N_L} \sum_{l_2=l_1}^{N_L} d_{l_1} d_{l_2} \sum_{l' \neq \{l_1, l_2\}} w_{i,l_1,l'} w_{j,l_2,l'} \sigma_{l'}^2. \end{aligned} \tag{17}$$

The first term is the sum of the link travel time variances, which would be the only term if the links were independent. The second term contains the direct dependencies between the traversed links, and the third term contains the second-order dependencies through common neighbors, on or off the traversed path.

Furthermore, we take into account the fact that consecutive observations from the same vehicle are correlated whenever the links traversed in the different observations are correlated. Incorporating this in the model helps the consistent estimation of the link dependence parameters ρ_i. We define a *trace* to be a contiguous sequence of observations from the same vehicle as it moves through the network. Observations from the same trace are correlated through the link correlation structures, while we treat observations from different traces as independent. Figure 3 illustrates the relationship between observations and traces.

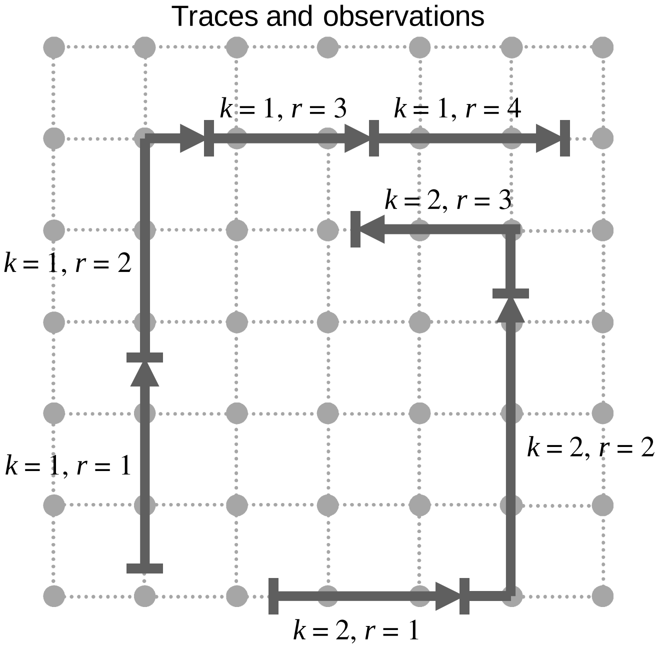

Figure 3: Illustration of two traces ($k = 1, 2$) each containing four and three observations, respectively.

In order to derive the full covariance structure of the observations, we first introduce the index $k = 1, \dots, N_K$ to represent a certain vehicle trace, where N_K (upper-case K) is the number of traces in the data. We further let the index $r = 1, \dots, N_k$ represent a certain observation in a certain trace k, where N_k (lower-case k) is the number of observations in the trace. The total number of observations in the data is $N_R = \sum_{k=1}^{N_K} N_k$. We then define $\mathbf{D}_k$ as the $N_k \times N_S$ matrix where element d_{rs} is the distance traversed on segment s for observation r. We also define $\mathbf{A}_k$ as the $N_k \times N_T$ matrix where element a_{rt} is equal to 1 if turn t is made in observation r and 0 otherwise. Further, we introduce the $N_k \times 1$ dependent variable vector $\mathbf{y}_k$ where element y_r is the travel time of observation r. The vector version of (14) is then

$$\mathbf{y}_k = \mathbf{D}_k \boldsymbol{X} + \mathbf{A}_k \mathbf{h}. \tag{18}$$

The travel times $\mathbf{y}_k$ are a linear transformation of multivariate normal stochastic variables and are thus distributed multivariate normal. We may write $\mathbf{y}_k = \boldsymbol{\mu}_k + \boldsymbol{\eta}_k$, where $\boldsymbol{\mu}_k$ is a vector of mean travel times and $\boldsymbol{\eta}_k$ is a vector of correlated zero-mean stochastic terms.

Trip conditions are assumed to be the same for all observations in a trace. Thus, each trace k is associated with two vectors of trip attributes $\mathbf{o}_k$ and $\mathbf{p}_k$ (weather conditions, weekday, season, etc.) affecting the means and variances, respectively. The vector of mean travel times for the trace is then simply the vector form of (15), where the mean segment travel time rates can be expressed in the model parameters using (2) and (3),

$$\boldsymbol{\mu}_k(\boldsymbol{\beta}_B, \boldsymbol{\beta}_E, \boldsymbol{\beta}_o) = \mathrm{E}[\mathbf{y}_k] = (1 + \mathbf{o}_k\boldsymbol{\beta}_o) \cdot (\mathbf{D}_k\mathbf{B}\boldsymbol{\beta}_B + \mathbf{A}_k\mathbf{E}\boldsymbol{\beta}_E). \tag{19}$$

Note how the trip conditions affect all segments and observations uniformly and can be moved outside the other terms.

The vector of stochastic terms for the observations in trace k is given by the vector form of (16),

$$\boldsymbol{\eta}_k = \mathbf{D}_k\boldsymbol{\nu} = \mathbf{D}_k\boldsymbol{S}\boldsymbol{u}. \tag{20}$$

For two different traces k and k', we have $\mathrm{E}\left[\boldsymbol{\eta}_k\boldsymbol{\eta}_{k'}^{\mathrm{T}}\right] = 0$ by assumption. Within the same trace k, meanwhile, (6), (12) and (20) gives the $N_k \times N_k$ covariance matrix

$$\begin{aligned}\boldsymbol{\Sigma}_k\left(\boldsymbol{\beta}_p, \boldsymbol{\sigma}_U^2, \boldsymbol{\rho}\right) &= \mathrm{E}\left[\boldsymbol{\eta}_k\boldsymbol{\eta}_k^{\mathrm{T}}\right] \\ &= \left(1 + \mathbf{p}_k\boldsymbol{\beta}_p\right)^2 \\ &\quad \cdot \sum_{n=1}^{N_U} \sigma_{U,n}^2 \left(\mathbf{D}_k\mathbf{S}\mathbf{U}_n\mathbf{S}^{\mathrm{T}}\mathbf{D}_k^{\mathrm{T}} + \sum_{i=1}^{N_\rho} \rho_i \mathbf{D}_k\mathbf{S}\left(\mathbf{W}_i\mathbf{U}_n + \mathbf{U}_n\mathbf{W}_i^{\mathrm{T}}\right)\mathbf{S}^{\mathrm{T}}\mathbf{D}_k^{\mathrm{T}} \right. \\ &\quad \left. + \sum_{i=1}^{N_\rho}\sum_{j=1}^{N_\rho} \rho_i\rho_j \mathbf{D}_k\mathbf{S}\mathbf{W}_i\mathbf{U}_n\mathbf{W}_j^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\mathbf{D}_k^{\mathrm{T}} \right).\end{aligned} \tag{21}$$

Again, the trip conditions affect all segments and observations in the same way and acts as a scalar multiplier to the entire covariance matrix. This means, for example, that the correlation between two observations is independent of the trip conditions.

4. ESTIMATION

4.1 Maximum likelihood estimation

Together, equations (19) and (21) provide the way to estimate the network model parameters $\boldsymbol{\beta}_B$, $\boldsymbol{\beta}_E$, $\boldsymbol{\beta}_o$, $\boldsymbol{\beta}_p$, $\boldsymbol{\sigma}_U^2$ and $\boldsymbol{\rho}$ using the probe vehicle travel time observations. The observations are multivariate normal within each trace and independent between traces. Hence, given all observed travel times $\mathbf{y}$, the log-likelihood function is analytical and given by

$$LL\left(\boldsymbol{\beta}_B, \boldsymbol{\beta}_E, \boldsymbol{\beta}_o, \boldsymbol{\beta}_p, \boldsymbol{\sigma}_U^2, \boldsymbol{\rho} \middle| \mathbf{y}\right) = -\frac{1}{2}N_R \log(2\pi) - \frac{1}{2}\sum_{k=1}^{N_K}(\mathbf{y}_k - \boldsymbol{\mu}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_k^{-1}(\mathbf{y}_k - \boldsymbol{\mu}_k) - \frac{1}{2}\sum_{k=1}^{N_K}\log|\boldsymbol{\Sigma}_k|. \tag{22}$$

We estimate the model parameters using standard numerical maximum likelihood techniques. Numerical estimation requires that the log-likelihood function (and its gradient vector, if analytical gradients are used) are evaluated for a sequence of parameter values until the maximum is found with sufficient accuracy. The computation times can be reduced considerably by pre-computing factors that do not depend on the parameter values outside of the optimization routine and keep as few operations as possible within the routine. For the mean structure, we note from (19) that the mean vector for trace k can be written as

$$\boldsymbol{\mu}_k = (\mathbf{1} + \mathbf{o}_k\boldsymbol{\beta}_o) \cdot (\mathbf{D}_k\mathbf{B}\ \ \mathbf{A}_k\mathbf{E})\left(\boldsymbol{\beta}_B^{\mathrm{T}}\ \ \boldsymbol{\beta}_E^{\mathrm{T}}\right)^{\mathrm{T}}. \tag{23}$$

The parameters for the mean segment travel time rates and the turn delays are thus merged into a single parameter vector. Here $(\mathbf{D}_k\mathbf{B}\ \ \mathbf{A}_k\mathbf{E})$ is an $N_k \times (N_B + N_E)$ matrix that is independent of the model parameters and may be computed and stored for all traces prior to the optimization. For the covariance structure, we note from (20) the model parameters are multipliers to constant matrices that may be precomputed and stored for every combination of parameters.

Matrix inversion is a costly operation, even if techniques such as LU or Cholesky factorization are used. In contrast to the SEM, the SMA model has the attractive property that it is not necessary to invert the full $N_L \times N_L$ link-level covariance matrix to compute the likelihood function, but only the $N_k \times N_k$ covariance matrices for the individual traces. Since the number of observations in a trace is typically much lower than the number of links in the network, this means that the SMA formulation requires significantly less computational effort to estimate. In addition to the theoretical considerations discussed in Section 2 this is another reason why we choose the SMA formulation for the model.

The sampling of the vehicle trajectories can be interpreted as a linear projection from the space of segment travel time rates and turn penalties to the space of observed travel times. The dimension of the observation space depends on the data: we expect the dimension to increase with the sampling frequency and with the number of observations, assuming that the vehicle trajectories are sampled at random locations. The network model represent another projection from the network components to the parameter space. Whether the parameters of the network model are identified through the data depends on the relative dimensions of the parameter space and the observation space: the higher the sampling frequency and the larger the number of observations, the larger the number of parameters that can be identified. This determines, for example, to what extent we can include fixed effects for specific segments or groups of segments in the model.

4.2 Network delimitations

One may often be interested in estimating travel times in a subnetwork, here called the *primary* network, which is smaller than the network spanned by the available GPS probes. If the primary network is small, for example a single street or road, the number of observations that traverse only segments in the primary network may be insufficient to estimate the travel times reliably. There may also be a bias since short traversed distances may be over-represented. Rather, we want to utilize all observations that to some extent traverse at least one of the segments in the primary network. We refer to these observations as *primary* observations.

With low frequency probe vehicle data, however, the primary observations will involve traversals of many segments and transitions also outside the primary network. We refer to these components, which depend on the used data, as the *secondary* network. If the primary network is small, the size of the secondary network can be many times greater. Once the secondary network has been identified, we can also utilize all observations that only traverse the secondary network, that is, do not extend the number of segments and transitions in the estimation any further. We refer to these observations as *secondary* observations, which may be many times greater in number than the primary observations. The concepts are illustrated in Figure 4.

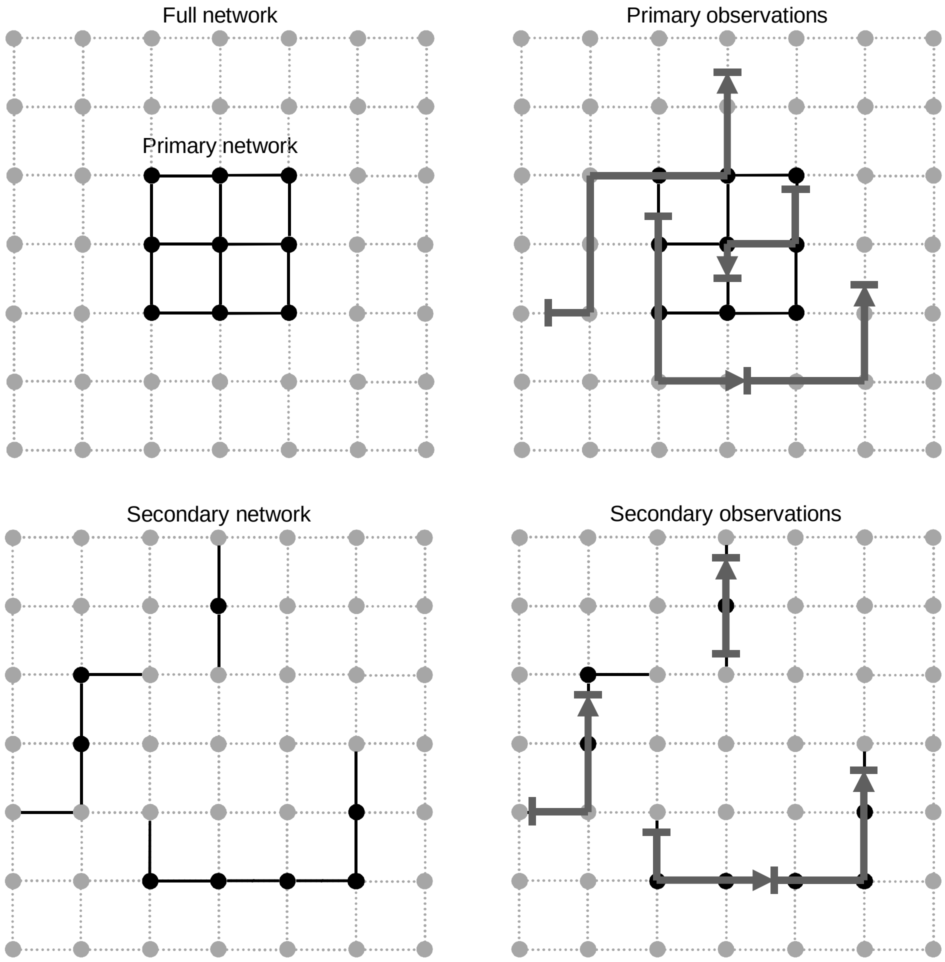

Figure 4: Illustration of the concepts of the primary and the secondary network.

4.3 Spatial clustering of network links

Sparse probe vehicle data may not have the resolution required to identify the travel time rate on all individual segments; indeed, this is the case in the application presented in Section 5. The use of explanatory variables can reduce the dimensionality of the problem. Still, it is very likely in practice that there are systematic differences in segment travel time rates between different parts of the network not captured by observed segment attributes. It is thus desirable to find a method for spatial clustering of segments that for a given dataset can provide a feasible compromise between the two extremes of fixed travel time rate effects for each segment and a single baseline travel time rate for all segments. Since applications would consist of anywhere from hundreds to hundreds of thousands of segments, the process furthermore needs to be automatic rather than manual.

This section describes one such automated process. The algorithm works at the link level, which ensures that all segments of the same link belong to the same group. The analyst may specify minimum, modal and maximum threshold values for the number of links in each group and the number of observations covering each group, respectively, which makes it possible to find a good balance between robust estimation and model resolution. Any constraint can be made non-binding by setting the corresponding threshold value sufficiently low or high. The algorithm can be used together with a manual grouping method for some part of the network; in the case study below, for example, we manually partition the primary network into link groups based on assumed similarity of traffic characteristics, while we use the automated partition method for the secondary network.

The algorithm is initialized with each link belonging to a separate group; links in manually defined groups are excluded. In each iteration, the group with the smallest number of observations is selected (in case of ties, an arbitrary group is chosen). For each link in the current group, it is checked whether it is connected with some link in another group through a common node. All such identified adjacent groups are added to a list. Going through the list in the order of increasing number of observations, the current group is merged with the first adjacent group for which any of two conditions hold:

1. The total number of links and the total number of observations in the two groups do not exceed the modal values,
2. The number of links or the number of observations in any of the groups is less than the minimum value, and the total number of links and the total number of observations in the two groups do not exceed the maximum values.

The algorithm stops when no groups can be merged further. The output of the algorithm can be summarized as a matrix $\mathbf{C}$, where element C_{lc} is equal to 1 if link l belongs to group c and zero otherwise. Since each link can only belong to one group we must have $\sum_c C_{lc} = 1$ for all l. The mapping of segments to groups is then obtained as the composite projection $\mathbf{SC}$.

5. CASE STUDY

5.1 Analysis description

We now present an application of the model described in Section 3 in the urban network of Stockholm, Sweden. The primary network consists of a route along one of the major inner city streets, the southern half of Birger Jarlsgatan, southbound direction, shown in Figure 5. The main route is chosen to coincide as closely as possible with a pair of automatic number plate recognition (ANPR) sensors mounted at each end of the route; see further Section 5.5.

The main route is about 1.4 km long and contains 28 links, divided into 36 segments in total, 26 intersections and 10 traffic signals; the speed limit is constant at 50 km/h. There is a busy commercial and entertainment center in the middle of the route with a nearby taxi stop, where we hypothesize that the mean travel time rates are higher than in adjacent parts, in particular for taxis. The route ends with a complicated signalized intersection in the south where delays can be significant. On the second half of the route there is a bus lane that taxis may use.

In the first part of the empirical study, we consider a few different specifications of the model structure given above. The focus here is not to derive the best model specification possible, but to demonstrate the structure of the model and the impact and significance of different explanatory factors on the observed travel times during a specific time interval (7:30-8:00 AM). We also evaluate the estimated travel time for the main route under different trip conditions. In the second part, we compare the estimated route travel time with an estimate using a simple weighted average procedure as well as observed travel times from the ANPR sensors, and we perform a sensitivity analysis regarding the filtering of the observations.

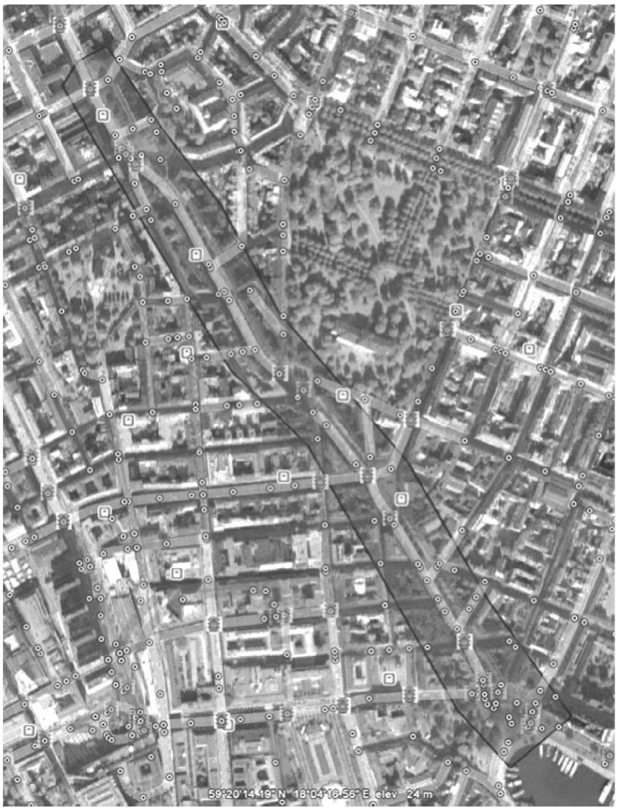

Figure 5: The case study area in Stockholm, Sweden. The shaded outlined area shows the main route, i.e., the primary network (Birger Jarlsgatan southbound). The secondary network extends even outside the shown area.

5.2 Data

The GPS probe vehicle data are obtained from the fleet dispatching system of a taxi company operating in total about 1500 vehicles in the Stockholm network. The data and the map-matching and path inference used to obtain the necessary estimation input are described in Rahmani and Koutsopoulos (2012). According to specifications from the manufacturers of the dispatching system, vehicles are sampled every 600 meters if occupied or every 400 meters if free, with a minimum threshold at 1 minute since the last sampling. In practice, the average sampling frequency in our data is about one report per 2 minutes and 780 meters, which is considerably lower than in most previously reported studies (e.g., one per minute in Hunter et al. (2009), Hofleitner et al. (2012), one per 200 meters in Westgate et al. (2011)).

In our baseline filtering procedure, observations covering less than 200 meters or more than 3 minutes are discarded as too unreliable in terms of map-matching and path inference. Indeed, the sampling rule specifications suggest that there should be no observations covering less than 400 meters, which lends additional uncertainty to short distance observations. We also remove observations with average speeds exceeding 90 km/h.

The digital network representation utilized contains information about various geometric attributes, including segment speed limit, functional class (a five-level hierarchical classification of the segments), traffic signals, and particular kinds of streets (ramps, tunnels, roundabouts, etc.). The network model is also used to identify intersections, left and right turns (defined as directional changes of more than 45 degrees), and one-way streets.

We use data for weekdays (Monday to Friday) and the time interval 7:30-8:00 AM between January 1, 2010 and December 31, 2011. For this two-year period, we have 63,680 observations in 44,844 traces after filtering. Of these, 10,604 observations are primary (that is, they cover the main route to some extent) and 53,076 are secondary (covering only the surrounding network). Across the primary and the secondary networks, the observations cover in total 1300 segments, 832 links and 1373 turns. On average, each observation covers 14.0 segments, 8.6 links and 7.7 turns. Figure 6 shows a histogram of the average vehicle speed for each observation, calculated as the traversed distance divided by the time between the reports. The speed distribution has mean 21.8 km/h, median 20.4 km/h and standard deviation 9.4 km/h.

Further, we have historical weather data for the same period as the probe vehicle data from a weather station in the area, reporting every 20-30 minutes. The data includes temperature, humidity, visibility distance and qualitative precipitation information (light/heavy rain/snow etc.).

There are many network attributes that we believe significantly impact travel times but are not available to us at this stage. This includes the number of lanes, locations of bus stops, pedestrian crossings, on-street parking, stop signs, traffic signal cycle times, nearby land use etc. Another important category of information would be the time and location for incidents and road works. If available, inclusion in the model is straightforward. All these attributes would likely contribute primarily to explain low speeds.

5.3 Model specification

5.3.1 Specification of link groups

In the most detailed model specification, each segment would have its own mean travel time rate parameter to be estimated. Our experiments revealed, however, that the resolution of the available probe vehicle data is not high enough to identify and reliably estimate all individual parameters. The situation is worst in the secondary network, where the number of observations covering each segment is generally lower. To overcome these challenges, we partition the network links into groups using a manual approach for the main route combined with the automated method described in Section 4.3 for the secondary network. First, the main route is divided into seven groups, each containing four links and 5.1 segments on average. The number of groups is selected so that fixed effects for the mean travel time rate of the segments in each group can be identified through the data, and the boundaries between groups are selected to capture the hypothesized heterogeneity of traffic conditions along the route.

Second, the secondary network is partitioned with the algorithm in Section 4.3 using the following threshold values which were found to produce good results with the given dataset. For the number of observations covering each group, the minimum value is set to 300 and the modal and maximum values are set to infinity. For the number of links in each group, the minimum, modal and maximum values are set to 10, 15 and 38, respectively. This produces 47 link groups for the secondary network, each containing 17.1 links and 26.9 segments on average. The distribution of the number of observations covering each link group is shown to the right in Figure 6. Together with the seven link groups for the primary network, we have 54 link groups in total.

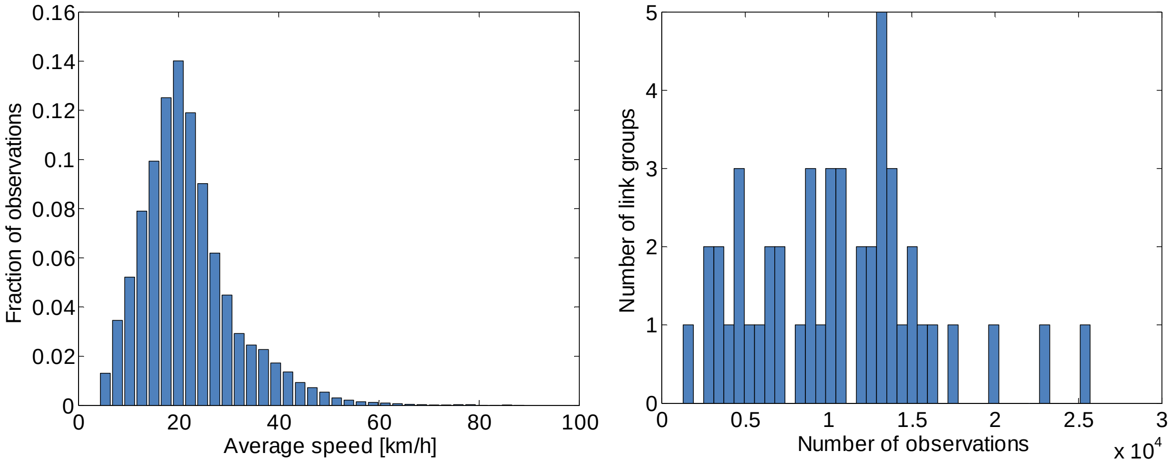

Figure 6: Left: Distribution of average speeds among the observations (7:30-8:00 AM). The speed distribution has mean 21.8 km/h, median 20.4 km/h and standard deviation 9.4 km/h. Right: Distribution of the number of observations covering each link group in the secondary network.

5.3.2 Specification of spatial weights

For the stochastic components of the travel time rates we specify an SMA structure as described in Section 2.3. We use a single weight matrix $\mathbf{W}$ and associated parameter ρ. The structure of $\mathbf{W}$ was chosen after extensive experimentation with alternative specifications. First of all, we investigated empirically the correlation between the travel time rates of consecutive observations within vehicle traces in the dataset. We found a statistically significant and positive correlation of about 6%. This suggests that there in general should be positive correlation between the network links.

The following structure turned out to produce meaningful results. Each link is assumed to be influenced by its nearest upstream neighbors in the network as well as the nearest upstream neighbors of those links (first and second-order neighbors, respectively), but only if there is at least one observation of a vehicle traversing these links in sequence. The latter requirement is added to prevent unintuitive effects such as a link being influenced by the opposite direction of the same street, which tend to give negative estimates of the spatial dependence. The second-order neighbors are given a weight of 0.25 relative to the first-order neighbors to represent the fact that influence should decay with distance. Finally, each row of $\mathbf{W}$ is normalized to sum to 1.

5.3.3 Model 1: Link groups

In order to investigate the effect and potential of different types of variables, we consider four different specifications. In the first specification (Model 1), the mean segment travel time rates are explained only with fixed effects at the link group level (54 groups in total), created as explained in Section 5.3.1. The mapping of every segment to its assigned group thus form the design matrix $\mathbf{B} = \mathbf{SC}$. We do not include explanatory variables for segment attributes, trip conditions or turn penalties. For the stochastic components we use, for the sake of simplicity, a formulation in which the variance parameter is common for all links.

5.3.4 Model 2: Segment characteristics

In the second specification (Model 2), we add explanatory variables for segment attributes and turn penalties, but no variables for trip conditions. This is a relevant level of specification for strategic applications, where typical travel times across different travel conditions are of interest in modeling and forecasting. We add the following variables, based on their relevance in reality, our access to data and their significance in the estimation.

Segment travel time rate attributes (matrix **B**):

- Dummy variables for every combination of segment speed limit (in km/h) and segment functional class (abbreviated FC) in the data. The FC classification ranges from 1 to 5, where 1 are highways and 5 are the most peripheral side streets. The combination $(50, 3)$ is used as reference level; the other existing combinations of speed limit and functional class are $(10, 5)$, $(30, 3)$, $(30, 4)$, $(30, 5)$, $(50, 1)$, $(50, 2)$, $(50, 4)$, and $(50, 5)$.
- One dummy variable for the case that the segment is on a one-way street.
- A dummy variable for the existence of a taxi stop on the segment. Our hypothesis is that the mean travel time rates for taxis on such segments are higher due to many stops.
- Dummy variables for the length of the link to which the segment belongs: length < 50 m and length ≥ 200 m, the reference level is length $\in [50,200)$ m. Our hypothesis is that deceleration and acceleration makes mean travel time rates higher on short segments.

Turn penalties (matrix **D**):

- Three dummy variables for the cases that the turn involves a signalized left turn, right turn and straight through movement, respectively.
- Two dummy variables for the cases that the turn involves a non-signalized left turn and a right turn, respectively.

5.3.5 Model 3: Trip conditions

In the third specification (Model 3), we also add explanatory variables for the trip conditions influencing each observation. For both the travel time means (vector $\mathbf{o}$) and the variances (vector $\mathbf{p}$), we consider the following variables:

- A dummy variable for the late half of the time interval, i.e., 7:45-8:00 (not July or public holiday). This allows us to estimate travel time variations within the day.
- One dummy variable for each weekday from Tuesday to Friday (reference level is Monday).
- A dummy variable for public holidays on weekdays (Christmas, Easter, etc.)
- Dummy variables for the summer (June-August), fall (September-November) and winter (December-February) seasons. The reference level is spring (March-May).
- A dummy variable for the month of July, the main vacation period in Sweden. The total effect of July is obtained by adding the effect of the summer season (June-August).
- A dummy variable for the year 2011 (reference level is 2010).
- Recent snow: Number of hours of consecutive snowfall reports preceding the trip.
- Recent rain: Number of hours of consecutive rainfall reports preceding the trip.
- A dummy variable for the taxi being free, as opposed to occupied or assigned to a customer.

5.3.6 Model 4: Independent links

The fourth specification (Model 4) is identical to Model 3 except that all links are assumed to be independent; in other words, the dependence parameter ρ is omitted. By comparing the performance of Model 3 and Model 4, we can evaluate the contribution of the correlation structure in the model.

5.4 Estimation results

The model specifications were estimated using the maximum likelihood estimation routine in the Statistics Toolbox for MATLAB and a trust-region reflective Newton optimization algorithm (MATLAB, 2009). Gradients of the log-likelihood function were evaluated analytically while the Hessian used to calculate standard errors was calculated numerically.

Estimation results are shown in Table 1. To begin with, the segment group fixed effects are all significant with T-statistics greater than 5 in all four model specifications. This demonstrates that the automated grouping method is capable of handling the identification problems associated with the low sampling frequency, even in the outer parts of the secondary network with few observations per segment. Also, the explanatory power of Model 1, captured by the log-likelihood and Akaike's information criterion (AIC), is considerably better than a model with a single mean travel time rate parameter and a single travel time rate variance parameter (the log-likelihood for this model is 181,049, AIC is -362,095).

The fit of the model increases greatly when segment and turn attributes are added (compare Model 2 vs. Model 1). This result is encouraging as it suggests that the low sampling frequency and identification power of the observations can be counter-balanced to some extent by making use of attributes of the network components. As expected, speed limits and functional classes have a strong impact on travel time rates. The directions and magnitudes are also intuitive: all else equal, a lower speed limit increases the travel time rate, as does a higher functional class (i.e., a lower hierarchical level). The only exception is the combination of 50 km/h speed limit and functional class 1, where the effect is not significant due to few observations.

Travel time rates are significantly higher on one-way streets, which may be due to more side friction compared to two-way streets. They are also considerably higher on segments with taxi stops, no doubt due to many taxis stopping there to wait for customers. When we predict travel times for personal trips, we can control for this bias in the taxi data by setting the taxi stop variable to zero for all links. Further, travel time rates are higher on shorter links and lower on longer links, reflecting the effect of acceleration and retardation as we hypothesized.

A traffic signal gives a delay that is different depending on the direction of movement: about ten second for left turns, seven seconds for right turns and three seconds when continuing straight ahead. A non-signalized left turn or right turn gives a delay of about 5 seconds, shorter than for signalized turns. Recall that these are average delays for all vehicles passing the traffic signals and intersections, whether they need to break or not. However, the delays are probably somewhat underestimated (and segment running travel times correspondingly overestimated), since the delay caused by the traffic signal or intersection may be distributed along the preceding link(s) due to queues etc., so that a vehicle sending a report just before reaching a turn as specified in the network model may already have experienced some of the associated delay.

Adding explanatory variables for the conditions of the trip improves the fit of the model further (compare Model 3 against Model 2). Mean travel time rates are significantly higher in the late half 7:45-8:00 of the period (+5%), suggesting a build-up of the morning peak compared to 7:30-7:45. There is also variation across the week: travel time rates are lowest on Mondays and highest on Tuesdays (2% higher than on Mondays), after which they decay towards the weekend. Travel time rates are considerably lower during public holidays when they drop almost 15%. This result is strongly significant even though the public holidays on weekdays are quite few. The summer season and in particular the main vacation month July are also associated with big reductions in travel time rates, whereas the fall and winter seasons as defined here are not significantly different from the spring. Interestingly, there is a small but significant increase in travel time rates from 2010 to 2011 of about 1.3%. This may be a sign of a long-term increase of congestion in the inner city of Stockholm.

Recent snowfall is found to significantly increase mean travel time rates. The estimate suggests that every four hours of consecutive snowfall before the trip increases travel time rates about 1%. This effect is expected since snow on the ground makes driving more difficult. Meanwhile, there is a weaker but opposite impact of rain on travel time rates. This is more unexpected but may be due to lower travel demand during rain. It is quite possible that further analysis of the weather data could lead to more refined insights into the impact of weather conditions (including for example combined effects of precipitation and temperature changes) on travel time rates.

We further find that travel time rates are significantly lower when the taxi is hired, which is likely due to a more determined driving behavior and that less time is spent cruising for customers at low speeds. When travel times are predicted for personal cars it may be appropriate to set this variable to zero, or estimate the model only on data from hired taxis.

Regarding the travel time variability, we find that the standard deviation is higher (+6%) in the late half of the period. Thus, both the average and the variability of the travel time rates increase, which is the expected effect of increasing congestion. Interestingly, the variability is greater than on Mondays for all other weekdays by roughly the same factor (around +5%). Public holidays, summer and vacation have lower standard deviations as expected due to less congestion. The standard deviation is significantly lower in 2011 (-6%), which is the opposite trend compared to the mean travel time rate. This result is intriguing and should, we believe, be investigated further.

The dependence parameter ρ for the SMA structure is positive and statistically significant in all models where it is included, which shows that there is positive correlation between network link travel time rates. For the main route, the estimated parameter value in Model 3 translates to travel time correlations between links up to about 10%, which is quite low. Also, treating links as independent reduces the fit of the model only moderately (Model 4 vs. Model 3). This would seem to suggest that correlations between network components are not vital to consider when estimating travel times. However, more research is needed to determine the characteristics of inter-link correlation and appropriate structures for the spatial weight matrices. Also, the variability introduced by the uncertainty in GPS positions and path choices may dilute the true dependencies in the observations, which means that the estimated correlation is biased downwards whereas the estimated variance is biased upwards.

Table 1: Travel time estimation results.

Parameters	Model 1		Model 2		Model 3		Model 4	
Mean segment travel time rates (**B**) [s/m]:								
54 segment groups	yes		yes		yes		yes	
Speed limit 10, functional class 5	-		0.3874	(0.0243)	0.3723	(0.0235)	0.3744	(0.0236)
Speed limit 30, functional class 3	-		0.0303	(0.0027)	0.0298	(0.0026)	0.0300	(0.0026)
Speed limit 30, functional class 4	-		0.0604	(0.0031)	0.0574	(0.0030)	0.0569	(0.0030)
Speed limit 30, functional class 5	-		0.0973	(0.0026)	0.0935	(0.0025)	0.0931	(0.0025)
Speed limit 50, functional class 1	-		0.0090	(0.0052)	0.0069	(0.0049)	0.0069	(0.0049)
Speed limit 50, functional class 2	-		-0.1183	(0.0212)	-0.1118	(0.0202)	-0.1104	(0.0200)
Speed limit 50, functional class 4	-		0.0741	(0.0028)	0.0717	(0.0027)	0.0717	(0.0027)
Speed limit 50, functional class 5	-		0.0437	(0.0064)	0.0421	(0.0061)	0.0416	(0.0061)
One way street	-		0.0124	(0.0022)	0.0134	(0.0021)	0.0135	(0.0021)
Taxi stop	-		0.0902	(0.0088)	0.0884	(0.0085)	0.0880	(0.0085)
Link length < 50 m	-		0.0159	(0.0053)	0.0158	(0.0050)	0.0176	(0.0050)
Link length ≥ 200 m	-		-0.0428	(0.0031)	-0.0417	(0.0029)	-0.0433	(0.0029)
Mean turn penalties (**D**) [s]:								
Traffic signal, left turn	-		10.5806	(0.5135)	10.1735	(0.4957)	9.9331	(0.4928)
Traffic signal, right turn	-		7.2148	(0.4485)	7.0411	(0.4310)	6.9168	(0.4302)
Traffic signal, straight through	-		3.0187	(0.2411)	2.9592	(0.2310)	2.8322	(0.2290)
Non-signalized left turn	-		5.5853	(0.4228)	5.5264	(0.4045)	5.4368	(0.4036)
Non-signalized right turn	-		5.0452	(0.4514)	4.8425	(0.4319)	4.7879	(0.4302)
Mean trip conditions (**o**) [%]:								
7:45-8:00 (not July, not holiday)	-		-		5.0771	(0.3312)	5.0627	(0.3298)
Tuesday	-		-		2.3204	(0.5075)	2.3110	(0.5054)
Wednesday	-		-		2.1867	(0.5159)	2.2062	(0.5136)
Thursday	-		-		1.2058	(0.4935)	1.2282	(0.4913)
Friday	-		-		0.4651	(0.4866)	0.4509	(0.4843)
Public holiday	-		-		-14.914	(1.3952)	-14.904	(1.3901)
Winter (December-February)	-		-		0.8166	(0.4692)	0.8321	(0.4672)
Summer (June-August)	-		-		-4.3173	(0.4569)	-4.3299	(0.4548)
Fall (September-November)	-		-		-0.5494	(0.4289)	-0.5329	(0.4270)
July (vacation month)	-		-		-6.0493	(0.6537)	-6.0117	(0.6509)
2011	-		-		1.3318	(0.3133)	1.3530	(0.3119)
Recent snow [% h^{-1}]	-		-		0.2493	(0.0509)	0.2501	(0.0508)
Recent rain [% h^{-1}]	-		-		-0.2273	(0.0964)	-0.2263	(0.0957)
Free taxi	-		-		3.9553	(0.3859)	3.9763	(0.3842)
Variance trip conditions (**p**) [%]:								
7:45-8:00 (not July, not holiday)	-		-		5.9000	(0.6664)	5.9505	(0.6678)
Tuesday	-		-		4.4767	(0.9681)	4.6134	(0.9706)
Wednesday	-		-		5.5488	(0.9933)	5.6092	(0.9955)
Thursday	-		-		5.2440	(0.9581)	5.3086	(0.9603)
Friday	-		-		5.8676	(0.9521)	5.8804	(0.9538)
Public holiday	-		-		-4.7278	(2.7427)	-4.6759	(2.7486)
Winter (December-February)	-		-		-2.3062	(0.8779)	-2.1807	(0.8799)
Summer (June-August)	-		-		-5.0317	(0.8631)	-4.9731	(0.8646)
Fall (September-November)	-		-		-2.4162	(0.8081)	-2.3277	(0.8096)
July (vacation month)	-		-		-4.5272	(1.2479)	-4.4151	(1.2513)
2011	-		-		-5.9791	(0.5807)	-5.9831	(0.5818)
Recent snow [% h^{-1}]	-		-		0.1327	(0.1010)	0.1412	(0.1014)
Recent rain [% h^{-1}]	-		-		-0.3269	(0.2026)	-0.3344	(0.2023)
Free taxi	-		-		8.9441	(0.6954)	9.1864	(0.6963)
Travel time rate correlation (ρ)	0.2519	(0.0313)	0.1794	(0.0292)	0.1350	(0.0288)	-	
Travel time rate variance (σ^2) [(s/m)^2]	0.0290	(0.0006)	0.0283	(0.0005)	0.0264	(0.0007)	0.0286	(0.0005)
Log likelihood	184,812		187,159		187,891		187,880	
AIC	-369,512		-374,172		-375,580		-375,559	
Observations	63680		63680		63680		63680	
Variables	53		73		101		100	

The travel time on the main route is estimated by applying the estimated parameters to a hypothetical observation traversing the route from beginning to end. Figure 7 illustrates how the estimated mean and variability vary under different combinations of trip conditions using the specification of Model 3. The reference, baseline conditions are those of the period 7:30-7:45 on a non-holiday Monday in spring 2010 with no recent rain or snow when the vehicle is hired. As the figure shows, the estimated mean travel time may vary up to a minute depending on the travel conditions considered here. Also, the variations in travel time may be very small but still statistically significant, such as the increase in travel time between 2010 and 2011.

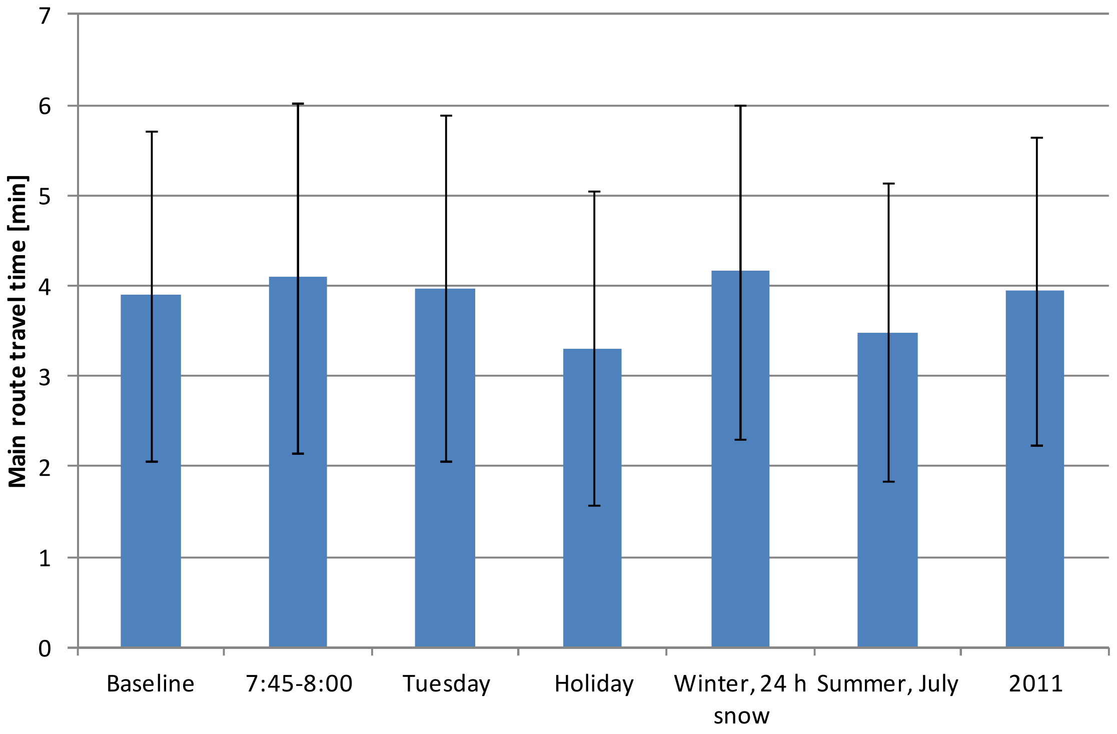

Figure 7: Estimated impact of trip conditions on the travel time for the main route (Birger Jarlsgatan southbound). The baseline conditions are 7:30-7:45 a.m., Monday, no holiday, Spring, 2010, no recent snow, no recent rain, hired taxi. The filled bars show the mean travel time, the error bars show 95% confidence intervals.

5.5 Comparison of estimates and filtering sensitivity analysis

Located at both ends of the main route are two ANPR cameras that record license plate numbers and time stamps when vehicles pass. The travel time of a vehicle on the route is calculated by taking the difference between the time stamps as the vehicle passes the two ANPR sensors. Note that the cameras capture all vehicles, including personal cars, taxis, trucks, buses etc. We have travel time data available from the ANPR sensors at the individual vehicle level for the period August 15 2011-April 12 2012.

For a number of reasons, the ANPR data cannot be regarded as "ground truth" data for the taxi probe vehicle data. Hence, comparison of the estimated travel time against the ANPR data should not be considered as validation in any strict sense. First, the ANPR data is noisy and a considerable number of observations has to be filtered out as outliers due to vehicles stopping along the route, taking detours, mismatched license plates, etc. This makes the statistics more uncertain. Second, it is not known exactly at which points the ANPR sensors register the vehicles; it may even vary between vehicles depending on when the cameras are able to read the license plates. This means that there may be systematic differences between the length of the taxi route and the ANPR route. Third and most important, it is generally acknowledged that taxis behave systematically different from overall traffic: they may drive faster than the average vehicle speed when occupied and far from the destination, whereas they stop more frequently and drive slower than average when picking up or dropping off customers.

Using probe vehicle data from the same period as the available ANPR data (August 15 2011-April 12 2012), we estimate a Model 2 specification using only hired taxis. This gives the main route travel time an estimated mean of 4.02 minutes and a standard deviation of 0.90 minutes. The traffic signals and intersections along the route contribute with 0.91 minutes (55 seconds) of delay, while the remaining 3.11 minutes is the running time on the links.

The distribution of travel times obtained from the ANPR data is shown in Figure 8. As can be seen, the distribution is unimodal and slightly skewed to the left. It can be reasonably approximated with a normal distribution, which lends support to our assumption of normally distributed travel times, at least over a certain route length. The population has mean 3.52 minutes, median 3.47 minutes and standard deviation 0.63 minutes.

Hence, the model estimates of the mean and standard deviation using the probe vehicle data are high compared to the ANPR data. The higher mean may be caused by a combination of the three issues mentioned above: assumptions in the filtering (of both ANPR and probe vehicle data), mismatching definitions of the start and end points of the main route, and systematic differences between taxis and the overall traffic. The higher standard deviation may be further caused by uncertainty in GPS locations and path choices, which inflate the estimate. Also, the specification of the variance used here is simple with only a single variance parameter for all links; it is possible that refining the estimate by adding link groups and explanatory variables for the variance may reduce the variance for the main route travel time.

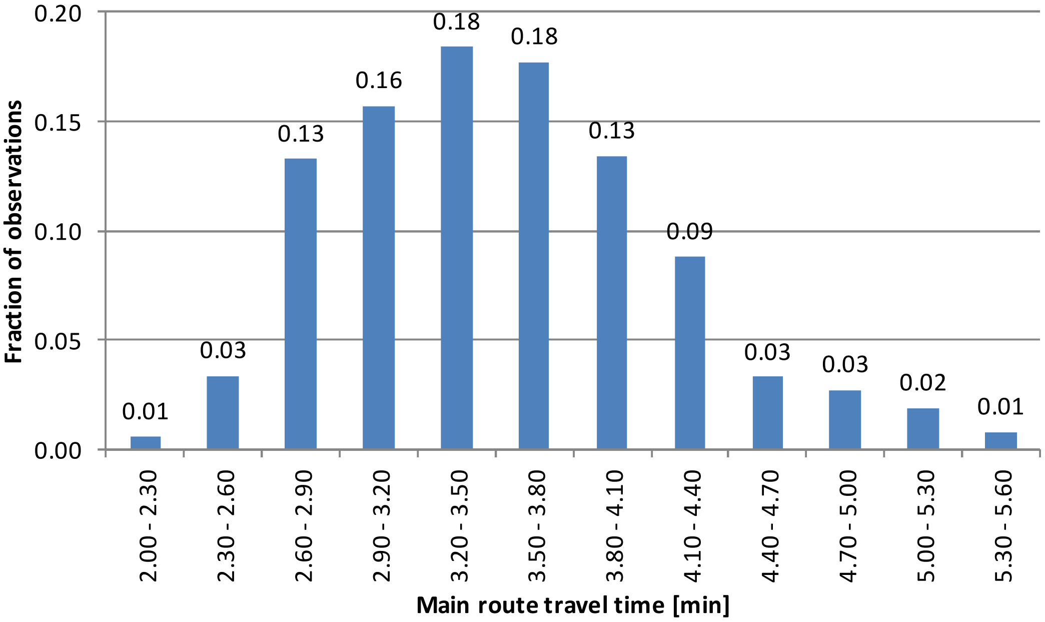

Figure 8: Distribution of travel times for the main route (Birger Jarlsgatan southbound) obtained from ANPR data. The distribution has mean 3.52 minutes, median 3.47 minutes and standard deviation 0.63 minutes.

We have compared the estimated main route travel time using the model presented in this paper with the travel time obtained from a simpler algorithmic procedure, described in Rahmani and Koutsopoulos (2012); a similar approach is used by Miller et al. (2010). In the simple procedure, the travel time of each probe vehicle observation is explicitly assigned to the traversed segments proportionally to the free-flow segment travel time rates (i.e., segment length divided by speed limit). This gives a set of travel time observations for each segment, which are weighted proportionally to the portion of the segment that is covered by the observation. A weighted average travel time is then calculated for each segment, and the main route travel time is calculated by summing the travel times of the constituent links.

The weighted average procedure is fast and simple, but the crude method of assigning travel times to links means that local traffic conditions will be spread out and bias estimates for all traversed links. A traffic light, for example, will lead to higher estimates for all links that are traversed in the observations passing the traffic light.

Using the weighted average procedure gives an estimated main route travel time of 3.71 minutes, in-between the ANPR travel time and the model estimate. This gives some support to the finding that the taxis travel slower than the overall traffic along the route, but the limitations of the model makes any further conclusions difficult.

We have investigated how sensitive the estimated travel times are to the filtering of the data. To recapitulate, in the filtering procedure used above we discard observations covering less than 200 meters or more than 3 minutes or with average speeds greater than 90 km/h since the map-matching and the path inference are judged to be too uncertain and to eliminate outliers. Removing the lower limit for the distance covered increases the estimated mean travel time for the main route by about 3%, while the estimated standard deviation is inflated by over 200%. Thus, while the mean travel time is robust to these uncertain observations, the variability around the mean is not, and filtering appears to be necessary. Removing the upper speed limit of 90 km/h has virtually no impact on the estimates, since the number of affected observations is very low. Lowering the upper limit on the time between consecutive probes from 3 minutes to 2 minutes, finally, reduces the estimated travel time mean and standard deviation by about 12%. This suggests that discarding long travel times due to uncertain path inference may lead to downward biased estimates of travel times. This is not a characteristic of the presented model but a general problem with using low frequency samplings of vehicle trajectories in travel time estimation.

6. DISCUSSION AND CONCLUSION

The paper has presented a statistical model for urban road network travel time estimation based on low frequency GPS probe vehicle data. Network travel times are modeled as running travel times on links and turn delay at traffic signals and intersections. To analyze the effects of network characteristics and trip conditions (time of day, season, weather conditions, etc.), the mean and variance of the link travel times and the turn delays are expressed as functions of explanatory variables in combination with fixed effects for groups of segments. This approach also allows the model to handle the low resolution of the sparse probe data. The network model allows correlation between travel times on different network links according to a spatial moving average structure. The observation model presents a way to estimate the parameters of the network model through low frequency samplings of vehicle traces, including the correlation as experienced by a driver traversing the links sequentially.

The model was applied to a part of the arterial network in Stockholm, Sweden. We found that attributes such as speed limits, functional classes, one-way streets and signalized or non-signalized left turns and right turns have significant effects on travel times. Trip conditions such as time-of-day, weekday, public holiday, vacation, year, recent snowfall and recent rainfall also have significant effects on both the mean and in some cases the variability of travel times. Further, there is significant positive correlation between links.

The case study highlights the potential of using sparse probe vehicle data to monitor the urban road transport system and identify changes in travel times and speeds based on a statistical foundation. Even small variations in travel times between days, seasons and years, can be identified with statistical precision, which suggests that opportunistic, sparse probe vehicle data can provide a cost-effective way of assessing the impacts of management actions, policy instruments and investments on traffic conditions. The increasing availability of such data also opens up the possibility for new types of control strategies; for example, congestion charges could be tied directly to a congestion index that is calculated from the estimated travel times and updated at suitable intervals.

The comparison with ANPR data recording travel times for all vehicles (personal cars, trucks, etc.), however, suggests that there may be systematic differences in traffic behavior between opportunistic probe vehicles, in this case taxis, and the overall vehicle fleet that we are ultimately interested in. These deviations may vary depending on the characteristics of the network and the source of data. In order to determine the best utilization of opportunistic probe vehicle data, the similarities and peculiarities compared to the overall traffic need to be investigated carefully. Further research is needed to determine the extent to which we may control for such deviations in the filtering of the data and the specification of the model.

On the methodological side, the case study also reveals the importance of specifying the spatial dependence weight matrices properly in order for estimated correlations between link travel times to be meaningful. This is still an undeveloped area of research where more analysis of the characteristics of inter-link correlations using probe vehicle data is needed.

The fact that the model represents segment travel times as a multivariate normal distribution means that there are good opportunities to extend the historical travel time estimation presented in this paper to online estimation and prediction using adaptations of well-established techniques such as Kalman filtering. This will be an area of further research. Another natural extension is statistical fusion with traffic data from other types of sensors, such as ANPR camera data for fixed routes and loop detectors. Further, the model can be developed to incorporate instantaneous speeds, which are reported by some probe vehicles, in addition to the travel times between consecutive reports.

Regarding the model structure, there are some assumptions that may be relaxed in future development of the model. For example, the assumption that the speed of a vehicle along a segment is constant could be generalized to the assumption that the speed profile along the segment follows a certain functional form. The parameters of this function can then be estimated along with the other model parameters based on the location of the GPS reports on the links. Further, the specification of the stochastic components of the travel times may be developed to better represent the noise arising from uncertainty in GPS locations and path choices, as well as unobserved day-to-day variations in travel conditions. Finally, it would be interesting to study the possibility of using mixtures of multivariate distributions to allow for more flexible travel time distributions while still maintaining some of the tractability of the current model. For example, a mixture model could perhaps better capture the discrete-continuous nature of traffic signal delay: either the vehicles have to break at the signal, causing some distribution of delay, or they can drive on undisturbed.

ACKNOWLEDGEMENTS

We would like to thank Tomas Julner from Trafik Stockholm for his support and provision of the data to the *iMobility* Lab at KTH, and Evanthia Kazagli for her assistance with processing the ANPR data. The work was funded by the Swedish Transport Authority through the Mobile Millennium Stockholm project.

REFERENCES

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

Bernard, M., J. Hackney, and K. W. Axhausen. Correlation of segment travel speeds. 6th Swiss Transport Research Conference, 2006.

de Fabritiis, C., Ragona, R. and Valenti, G. (2008) Traffic estimation and prediction based on real time floating car data. *11th International IEEE Conference on Intelligent Transportation Systems*, pp. 197–203.

Cheng, T., Haworth, J. and Wang, J. (2011) Spatio-temporal autocorrelation of road network data. *Journal of Geographical Systems*, in press.

Hellinga, B., Izadpanah, P., Takada, H. and Fu, L. (2008) Decomposing travel times measured by probe-based traffic monitoring systems to individual road links. *Transportation Research Part C* 16, 768–782.

Hepple, L. W. (2004) Bayesian model choice in spatial econometrics. In LeSage, J. P. and Pace, R. K. (eds.), *Spatial and Spatiotemporal Econometrics*, Advances in Econometrics vol. 18, pp. 101-126, Elsevier Ltd.

Herring, R., Hofleitner, A., Amin, S., Abou Nasr, T., Abdel Khalek, A., Abbeel, P. and Bayen, A. (2010) Using mobile phones to forecast arterial traffic through statistical learning. TRB Annual Meeting 2010.

Hofleitner, A., Herring, R., Abbeel, P. and Bayen, A. (2012) Learning the dynamics of arterial traffic from probe data using a dynamic Bayesian network. *IEEE Transactions on Intelligent Transportation Systems*, in press.

Hunter, T., Herring, R., Abbeel, P. and Bayen, A. (2009) Path and travel time inference from GPS probe vehicle data. Neural Information Processing Systems Foundation (NIPS) Conference, Vancouver, Canada.

Leduc, G. (2008) Road traffic data: collection methods and applications. JRC Technical Notes, Working Papers on Energy, Transport and Climate Change, N.1.

LeSage, J. and Pace, R. K. (2009) *Introduction to Spatial Econometrics*. Chapman & Hall/CRC, Boca Raton, Florida.

MATLAB (2009) MATLAB version 7.8.0. Natick, Massachusetts: The MathWorks Inc., 2009.

Miller, J., Kim, S., Ali, M. and Menard, T. (2010) Determining time to traverse road sections based on mapping discrete GPS vehicle data to continuous flows. *IEEE Intelligent Vehicles Symposium*, pp. 615–620.

Min, X., Hu, J., Chen, Q., Zhang, T. and Zhang, Y. (2009) Short-term traffic flow forecasting of urban network based on dynamic STARIMA model. *12th International IEEE Conference on Intelligent Transportation Systems*, pp. 461-466.

Miwa, T., Sakai, T., and Morikawa, T. (2008) Route identification and travel time prediction using probe-car data. *International Journal of ITS Research* 2, 21-28.

Rahmani, M., and Koutsopoulos, H. N. (2012) Path inference of low-frequency GPS probes for urban networks. Technical report, KTH Royal Institute of Technology, Stockholm, Sweden.

van Aerde, M., Hellinga, B., Yu, L. and Rakha, H. (2003) Vehicle probes as real-time ATMS sources of dynamic O-D and travel time data. *Large Urban Systems: Proceedings of the Advanced Traffic Management Conference* 2003, pp 207-230.

Westgate, B. S., Woodard, D. B., Matteson, D. S., and Henderson, S. G. (2011) Travel time estimation for ambulances using Bayesian data augmentation. Submitted, Journal of the American Statistical Association.

Work, D. B., Tossavainen, O.-P., Blandin, S., Bayen, A. M., Iwuchukwu, T. and Tracton, K. (2008) An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. *Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 5062-5068.

Zheng, F. and van Zuylen, H. (2012) Urban link travel time estimation based on sparse probe data. *Transportation Research Part C*, in press.

Zou, L., Xu, J.-M. and Zhu, L.-X. (2005) Arterial speed studies with taxi equipped with global positioning receivers as probe vehicle. *Proceedings of the 2005 International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1343-1347.