

1 **Travel Time Estimation for Urban Road Networks Using Low Frequency GPS Probes**

2

3

4

5 Erik Jenelius

6 Corresponding author

7 Department of Transport Science

8 KTH Royal Institute of Technology

9 Teknikringen 72, SE-100 44 Stockholm, Sweden

10 Tel: +46 8 790 8302

11 Fax: +46 8 21 2899

12 Email: erik.jenelius@abe.kth.se

13

14 Mahmood Rahmani

15 Department of Transport Science

16 KTH Royal Institute of Technology

17 Teknikringen 72, SE-100 44 Stockholm, Sweden

18 Tel: +46 8 790 8302

19 Fax: +46 8 21 2899

20 Email: mahmood.rahmani@abe.kth.se

21

22 Haris N. Koutsopoulos

23 Department of Transport Science

24 KTH Royal Institute of Technology

25 Teknikringen 72, SE-100 44 Stockholm, Sweden

26 Tel: +46 8 790 8302

27 Fax: +46 8 21 2899

28 Email: haris.koutsopoulos@abe.kth.se

29

30

31 29 November, 2011

1 ABSTRACT

2

3 The paper presents a model for urban road network travel time estimation using low frequency GPS

4 probes as observations. The model separates trip travel times into segment travel times and intersection

5 delays, which are modeled as stochastic variables. We combine fixed segment effects with geometric

6 attributes and trip conditions as explanatory variables to reduce the number of parameters to estimate

7 and to be able to predict travel times also in areas with very few observations. We are able to assess

8 the significance of the results by reporting standard errors. The approach also makes results

9 considerably easier to interpret, evaluate and use for qualitative forecasting. Further, the model allows

10 correlation between travel times on different network segments and presents a way to estimate the

11 correlation experienced by a driver traversing the segments sequentially on the trip. The model is

12 applied in a case study for the network of Stockholm, Sweden. We find that geometric attributes and

13 trip conditions have significant effects on travel times and that there is significant correlation between

14 segments.

1 INTRODUCTION

2
3 Many urban road transport systems today experience increasing congestion that threatens the
4 environment and the transport efficiency. To tackle these problems, knowledge about traffic conditions
5 is critical at many levels of traffic management and transport policy. Through information and
6 personalized advice, individuals and transporters can plan their trips more accurately and increase the
7 efficiency of the system. For traffic management, travel time information at the segment level can
8 reveal problematic locations where new or revised traffic control schemes may be introduced to
9 increase performance. For transport policy, network-wide travel time information provides input for
10 travel demand forecasting and impact assessments of policy instruments such as congestion charges.

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

21 Most recently, GPS devices, already installed for other purposes in vehicle fleets (e.g., taxis,
22 commercial vehicles, service vehicles, etc.) or smartphones, provide a new type of traffic sensor.
23 These opportunistic sensors have a great potential for provision of data for traffic management
24 applications. Unlike stationary sensors, they can collect travel time data for any part of the network
25 where equipped vehicles move. Unlike designated floating cars, they can continuously collect data for
26 any time and day that equipped vehicles are active. However, despite their advantages, an important
27 disadvantage for the widespread use of these data is that more advanced and sophisticated methods are
28 needed to process the data and generate useful information, compared to traditional sensors (1).

29 A number of limitations make the use of opportunistic FCD data challenging. First, the
30 penetration rate is still typically low, which means that the collected information is only a small
31 sample of the full traffic state of the system. Furthermore, there may be systematic differences between
32 the equipped vehicles and the overall population (2). Second, the accuracy and the frequency of the
33 position reports, while adequate for the original purpose of the GPS device, may be of low quality
34 when used for travel time estimation.

35 The literature on travel time estimation and forecasting using GPS sensors has grown in recent
36 years as the technology has become more available. Most papers, however, have dealt with high
37 frequency data, e.g., (3,4), which eliminates many of the challenges of interest here. Low sampling
38 frequency creates difficulties in inferring the true path of the vehicle between two position reports,
39 which may involve a considerable number of network segments (5,6). It also becomes difficult to
40 identify the fraction of the travel time spent on each individual segment (7,8). For short segments in
41 particular, there may be few observations available to estimate the travel time distribution under the
42 conditions of interest.

43 A probabilistic model of travel times through the arterial network based on low frequency taxi
44 GPS probes is presented in (9). The model takes into account that the path between two consecutive
45 position reports may contain multiple segments, but not that a reported position may be in the interior
46 of a segment. The authors formulate a maximum likelihood (ML) problem to estimate the segment
47 travel time distributions based on the set of observed route travel times. As the segment travel times
48 are typically not directly observed, a simulation based EM estimation algorithm is proposed. The
49 authors assume that the travel times on different segments are independent and briefly report
50 estimation results using normal and log-normal distributions.

51 A development of the approach in (9), more clearly aimed towards travel time forecasting, is
52 presented in (10). The model assumes that each segment can be in one of several possible states, each
53 with its own conditional, independent travel time distribution. The transitions between states among
54 neighboring segments are modeled as a Coupled Hidden Markov Chain model. The unobserved state
55 and transition probabilities and the travel time distribution parameters are estimated in a simulation

1 based EM approach. In an experimental study, they assume independent normal distributions for the
2 segments.

3 Another approach, using low frequency GPS data from ambulances, is presented in (11). In the
4 paper, path inference and travel time estimation are performed simultaneously using Bayesian data
5 augmentation techniques. The framework makes use of instantaneous speed information reported by
6 the vehicles. The travel times on the road links are assumed to be independent and log-normally
7 distributed, and the parameters are estimated using Markov chain Monte Carlo methods.

8 The previous approaches to travel time estimation using low frequency GPS data all assume
9 that the travel times on different road links are independent (9-11). In practice, however, congestion,
10 weather conditions etc. mean that travel times are often positively correlated across links (12,13). That
11 is, if on a given day the travel time on one segment is unusually high (or low), then the travel time on
12 nearby segments will likely also be unusually high (low). Spatial and temporal dependencies between
13 segment travel times have been considered for forecasting purposes using STARMA (Spatio-Temporal
14 Auto-Regressive Moving Average) and similar models (14,15).

15 This paper presents a statistical model for urban road network travel times using observations
16 from low frequency GPS probes. The purpose of the model is to predict the travel time for any route in
17 the network under specified conditions, considering both the mean travel time and the variability. It is
18 intended to be used for the arterial network as well as for highways. The basic idea is to increase the
19 reliability of the estimation by utilizing all observed vehicle trajectories that provide information about
20 some part of the route in the estimation, and not only such observations that cover the entire route (of
21 which there may even be none).

22 The presented approach provides a complement to previous models by partially using
23 explanatory variables to predict travel times. This may include attributes such as speed limit, number
24 of lanes, functional class, bus stops, traffic signals, stop signs, left turns, etc. The assumption is that
25 these properties have a significant impact on the travel times that is uniform across the network. The
26 statistical approach reduces the number parameters to estimate and allows us to predict travel times
27 also in areas with very few observations. We are able to assess the robustness of the results by
28 reporting standard errors. The approach also makes results considerably easier to interpret, evaluate
29 and use for qualitative forecasting. The model does not require data about the instantaneous speeds of
30 vehicles, nor about the traffic flow on the segments.

31 Further, the model allows correlation between travel times on different network segments and
32 presents a way to estimate this correlation using GPS data. The correlation between segments is
33 incorporated not only in each travel time observation, but in the entire sequence of observations from
34 the same vehicle trajectory. It is important to note that we are here considering the correlation
35 experienced by a driver traversing the segments sequentially on the trip, and not the correlation of
36 travel time rates at the same instant across the network. As far as we are aware, correlation in this
37 sense has not been previously considered in network travel time estimation.

38 The paper is organized as follows. The general model is presented in the next section,
39 followed by a description of a case study, sensitivity analysis and evaluation, and some concluding
40 remarks.

41 NETWORK MODEL

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

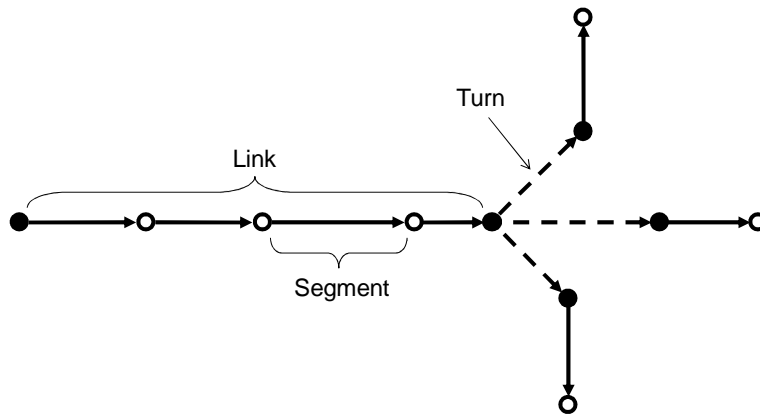
- 43 1. Running travel time along links,
- 44 2. Delay at intersections and traffic signals (turns).

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

1 In this model vehicle speed can vary between segments but is assumed to be constant along
 2 each segment. The travel time on a segment s can always be decomposed as the length of the segment,
 3 ℓ_s , multiplied with the inverse speed or *travel time rate* X_s . The travel time rate may depend on
 4 observed and unobserved properties of the segment, such as speed limit, as well as conditions for the
 5 trip, such as weather conditions, congestion, characteristics of the driver and vehicle etc.

6 The second component of the trip travel time is intersection and traffic signal delay. We define
 7 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
 8 number of turns in the network. The turn t is assumed to give a travel time penalty Z_t which again
 9 may depend on the conditions for the trip. Systematic factors that would influence the magnitude of Z_t
 10 may include the type of traffic control in the intersection and whether it involves a left or a right turn.

11 Conceptually, turns can be treated as links having zero length, as illustrated in Figure 1. This
 12 means, for example, that the probability of observing a vehicle on a turn link is zero. While one could
 13 readily incorporate both types of components in a single set of variables, in this paper we will use
 14 separate sets of variables for link running times and turn penalties for clarity.



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

17
 18
 19 The segment travel time rates X_s and the turn penalties Z_t are modeled as stochastic variables.
 20 We are interested in both mean travel times and the variability around the mean values. It is
 21 convenient to introduce the $N_S \times 1$ vector \mathbf{X} with elements X_s and the $N_T \times 1$ vector \mathbf{Z} with elements
 22 Z_t . Assuming that the mean values are finite, we can write

$$\mathbf{X} = \mathbf{g} + \mathbf{v},$$

$$\mathbf{Z} = \mathbf{h} + \boldsymbol{\eta},$$

24 where \mathbf{g} and \mathbf{h} are vectors of mean values, and \mathbf{v} and $\boldsymbol{\eta}$ are vectors of stochastic terms with $E[\mathbf{v}] =$
 25 $E[\boldsymbol{\eta}] = \mathbf{0}$.

26
 27 The model assumes that the stochastic components follow a multivariate normal distribution.
 28 As shown below, this allows us to express the likelihood of the observations in closed form. We note
 29 that travel times and speeds are often found empirically to have more skewed distributions, which may
 30 make the assumption of normality here seem questionable. However, the better we are able to explain
 31 the mean travel times and the variance, the more the residuals should tend to white noise. In any case,
 32 we believe that the efficient estimation allowed through the normality assumption makes this approach
 33 a highly promising way forward. In the case study below we investigate the shape of the travel time
 34 distribution, and we discuss possible relaxations of the normality assumption in the concluding section
 35 of the paper.

37 Mean Structure

39 Segment Travel Time Rates

40 The mean vectors \mathbf{g} and \mathbf{h} are further expressed as functions of a number of factors with associated
 41 parameters to be estimated from data. The parametric structures should reflect the way that different
 42 factors affect travel times, while also allowing for convenient and efficient estimation. In the case

1 study presented below, we consider a few different model specifications. In the largest specifications,
 2 we divide the explanatory variables for \mathbf{g} into two different categories, which are hypothesized to
 3 impact the travel times in different ways as explained below.

- 4
- 5 1. The first category of attributes are static characteristics of the segments and the links to which
 6 they belong. This category could include regulatory factors such as speed limit and functional
 7 class, link length, congestion indicators and fixed effects for specific segments. The attributes
 8 are collected in the $N_S \times N_B$ design matrix \mathbf{B} .
- 9
- 10 2. The second category of attributes are observed conditions for the trip, which are assumed to
 11 act as a multiplier to all the link travel time rates determined by the variables in categories 1.
 12 This category could include temporal variations within the day, week and year, weather
 13 conditions, etc. For a given trip, the attributes are collected in the $1 \times N_O$ design vector \mathbf{o} .
- 14

15 Within each category we assume an additive structure, and between the categories we assume a
 16 multiplicative structure. The mean segment travel time rates can then be written as

$$17 \quad \mathbf{g} = (1 + \mathbf{o}\boldsymbol{\beta}_O) \cdot \mathbf{B}\boldsymbol{\beta}_B,$$

18 where $\boldsymbol{\beta}_B$ is an $N_B \times 1$ parameter vector for the baseline travel time rate attributes, $\boldsymbol{\beta}_O$ is an $N_O \times 1$
 19 parameter vector for the trip condition attributes.

20 *Turn penalties*

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

$$26 \quad \mathbf{h} = (1 + \mathbf{o}\boldsymbol{\beta}_O) \cdot \mathbf{E}\boldsymbol{\beta}_E.$$

27 The main contribution of this paper is not the particular structures assumed for the means, but
 28 rather the way we are able to handle correlation and identify the network model through the
 29 observations, as shown below. Hence, other structures may be adopted if they are found to better suit
 30 the data. For example, the model can be extended to also allow trip conditions to affect a specific
 31 segment or turn. This would be appropriate, e.g., if network-coded information is used about traffic
 32 incidents or construction works that do not cover the whole period of observations.

33 **Variance Structure**

34 *Segment Travel Time Rates*

35 The stochastic elements $\boldsymbol{\nu}$ and $\boldsymbol{\eta}$ represent the variability in travel times due to unobserved variability
 36 in traveler characteristics, traffic conditions and network component characteristics. We consider first
 37 the travel time rates. In the model this stochastic component is represented at the *link* level. That is, the
 38 travel time rates on different segments in a link are allowed to differ in means but are assumed to have
 39 the same variability around the mean, implying perfect correlation within the link. We consider links
 40 between intersections and traffic signals to be a more intuitive and convenient unit for representing
 41 travel time variability than segments, especially when considering the correlation between units.

42 Each link has an independent stochastic travel time rate component ϵ_l with variance σ_l^2 , which
 43 may be further decomposed (see below). Using vector notation, we have the $N_L \times 1$ vector of
 44 independent errors $\boldsymbol{\epsilon}_L$ with variances $\boldsymbol{\sigma}_L^2$. We also allow travel time rates to be correlated among
 45 nearby links due to common characteristics in unobserved congestion, geometric properties etc. Note
 46 that we are considering the correlation faced by a driver traversing the segments sequentially during a
 47 trip.

To capture this dependence we use a spatial moving average (SMA) error component specification adapted to an urban road network setting (16). The approach requires the specification of spatial *weights* to represent dependencies between links. In spatial econometrics, it is typically assumed that dependencies between units decay with the number of intermediate units and/or the distance separating them. For links in a road network, this intuitively corresponds to the number of turns and the network distance or travel time separating the links. However, dependencies between the travel times rates of road network links may also depend on traffic flows, network hierarchy etc., which may make the number of turns a less important characteristic. We are free to specify the weights in order to capture the mechanisms behind the correlation as well as possible.

We consider an arbitrary number N_ρ of dimensions or orders of link dependence. For each pair of links l and l' , we define $w_{ll'}^i$ to be the weight of the influence of l' on l in dimension i ; we require that $w_{ll}^i = 0$ for all l and i . Introducing the i :th dimension $N_L \times N_L$ weight matrix \mathbf{W}_i with elements $w_{ll'}^i$, the total error structure at the segment level is

$$\mathbf{v} = \mathbf{S} \left(\mathbf{I} + \sum_{i=1}^{N_\rho} \rho_i \mathbf{W}_i \right) \boldsymbol{\epsilon}_L,$$

where \mathbf{I} is the $N_L \times N_L$ identity matrix and ρ_i is a parameter to be estimated, representing the overall sign and magnitude of i :th dimension dependence, $i = 1, \dots, N_\rho$.

Similarly to the means, the variances σ_L^2 are expressed as functions of a number of factors with associated parameters to be estimated from data. First, the baseline variances are assumed to consist of a number of additive components, for example fixed link effects, indicators for congestion, etc. These components are represented by the $N_L \times N_U$ design matrix \mathbf{U} . Further, the observed travel conditions for the trip are assumed to impact the travel time variances as well as the means. Relevant attributes could include weekday, time of year, weather conditions, etc. The attributes are collected in the $1 \times N_p$ design vector \mathbf{p} . Thus,

$$\sigma_L^2 = (1 + \mathbf{p}\boldsymbol{\beta}_p)^2 \cdot \mathbf{U}\boldsymbol{\sigma}_U^2,$$

where $\boldsymbol{\beta}_p$ is an $N_p \times 1$ parameter vector and $\boldsymbol{\sigma}_U^2$ is an $N_U \times 1$ parameter vector. If we introduce the diagonal variance matrix $\boldsymbol{\Omega}_0 = \text{diag}(\mathbf{U}\boldsymbol{\sigma}_U^2)$, the *segment-level* covariance matrix $\boldsymbol{\Omega}$ can be obtained as

$$\begin{aligned} \boldsymbol{\Omega} &= E[(\mathbf{S}\boldsymbol{\epsilon}_L)(\mathbf{S}\boldsymbol{\epsilon}_L)^T] \\ &= (1 + \mathbf{p}\boldsymbol{\beta}_p)^2 \\ &\quad \cdot \left(\mathbf{S}\boldsymbol{\Omega}_0\mathbf{S}^T + \sum_{i=1}^{N_\rho} \rho_i (\mathbf{S}\mathbf{W}_i\boldsymbol{\Omega}_0\mathbf{S}^T + \mathbf{S}\boldsymbol{\Omega}_0\mathbf{W}_i^T\mathbf{S}^T) + \sum_{i=1}^{N_\rho} \sum_{j=1}^{N_\rho} \rho_i \rho_j \mathbf{S}\mathbf{W}_i\boldsymbol{\Omega}_0\mathbf{W}_j^T\mathbf{S}^T \right). \end{aligned}$$

Turn penalties

Turn penalties may be correlated through for example the regular cycles of traffic signals, so that stopping at one signal increases or reduces the probability of stopping at subsequent signals. We capture such spatial dependencies by considering a SMA error model for the turn penalties much in the same way as for the link travel time rates. To begin with, each turn has an independent stochastic delay component; in vector notation, we have the $N_T \times 1$ vector of independent errors $\boldsymbol{\epsilon}_T$ with variances σ_T^2 . For each dimension of dependence i , we introduce the $N_T \times N_T$ weight matrix \mathbf{Q}_i and parameter λ_i analogous to \mathbf{W}_i and ρ_i above, respectively. With N_λ dimensions in total, the error structure is thus

$$\boldsymbol{\eta} = \left(\mathbf{I} + \sum_{i=1}^{N_\lambda} \lambda_i \mathbf{Q}_i \right) \boldsymbol{\epsilon}_T.$$

1 Like for the link travel time rates, the variances σ_T^2 are expressed as functions of a number of
 2 factors with associated parameters to be estimated from data. The baseline variances are assumed to
 3 consist of a number of additive components, for example fixed effects, indicators for traffic signals,
 4 etc., collected in the $N_T \times N_V$ design matrix \mathbf{V} . The observed travel conditions are assumed to impact
 5 the variance of the turn penalties in the same way as the segment travel time rates. Thus,

$$6 \quad \sigma_T^2 = (1 + \mathbf{p}\boldsymbol{\beta}_P)^2 \cdot \mathbf{V}\boldsymbol{\sigma}_V^2,$$

7
 8 where $\boldsymbol{\sigma}_V^2$ is an $N_V \times 1$ parameter vector. Introducing the diagonal variance matrix $\boldsymbol{\Pi}_0 = \text{diag}(\mathbf{V}\boldsymbol{\sigma}_V^2)$,
 9 the turn covariance matrix $\boldsymbol{\Pi}$ is

$$10 \quad \boldsymbol{\Pi} = E[\boldsymbol{\epsilon}_T(\boldsymbol{\epsilon}_T)^T] = (1 + \mathbf{p}\boldsymbol{\beta}_P)^2 \cdot \left(\boldsymbol{\Pi}_0 + \sum_{i=1}^{N_\lambda} \lambda_i (\mathbf{Q}_i \boldsymbol{\Pi}_0 + \boldsymbol{\Pi}_0 \mathbf{Q}_i^T) + \sum_{i=1}^{N_\lambda} \sum_{j=1}^{N_\lambda} \lambda_i \lambda_j \mathbf{Q}_i \boldsymbol{\Pi}_0 \mathbf{Q}_j^T \right).$$

11 12 13 OBSERVATION MODEL

14
 15 The travel time measurements considered in this paper consist of sparsely sampled vehicle trajectories
 16 through the network obtained from GPS devices or similar sensor technologies. In general, GPS
 17 location measurements are associated with errors. Here we assume that the most likely network
 18 location corresponding to each GPS measurement, as well as the path (i.e., the sequence of network
 19 segments) taken between each pair of consecutive measurements, have been determined by a map-
 20 matching and path inference process (5). A basic observation then consists of

- 21 1. a vehicle identification number,
- 22 2. a pair of time stamps τ_1, τ_2 ,
- 23 3. a path representing the trajectory of the vehicle between the two time stamps, involving a
 24 sequence of segments (s_1, \dots, s_n) , a corresponding sequence of turns (t_1, \dots, t_m) , and two
 25 offsets δ_1, δ_2 specifying the vehicle locations at τ_1, τ_2 in relation to the upstream nodes of the
 26 first and last segments of the path, respectively.

27
 28
 29 The observed travel time $y = \tau_2 - \tau_1$ is a linear combination of the travel time rates of the traversed
 30 segments, each weighted by the distance travelled on the segment, and the delays of the performed
 31 turns. We exploit this fact to derive the distribution of y when perceived as a stochastic variable. With
 32 ℓ_s denoting the length of segment s , we have

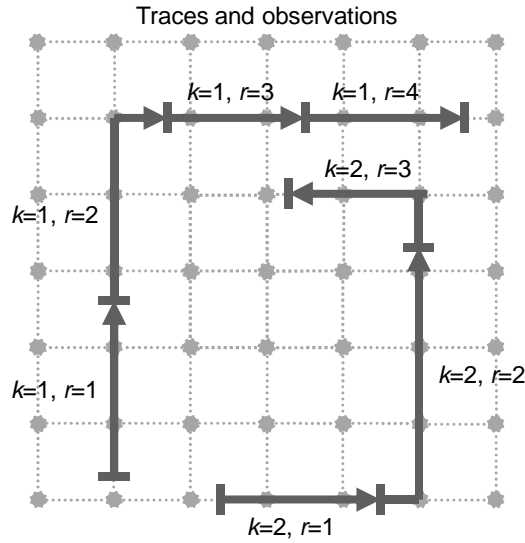
$$33 \quad y = \begin{cases} (\delta_2 - \delta_1)X_{s_1} & \text{if } n = 1, \\ (\ell_{s_1} - \delta_1)X_{s_1} + \delta_2 X_{s_2} & \text{if } n = 2 \text{ and } m = 0, \\ (\ell_{s_1} - \delta_1)X_{s_1} + \delta_2 X_{s_2} + Z_{t_1} & \text{if } n = 2 \text{ and } m = 1, \\ (\ell_{s_1} - \delta_1)X_{s_1} + \sum_{i=2}^{n-1} \ell_{s_i} X_{s_i} + \delta_2 X_{s_n} + \sum_{i=1}^m Z_{t_i} & \text{otherwise.} \end{cases}$$

34
 35 We can express this relation more compactly using vector notation. First we define d_s as the
 36 distance traversed on segment s and a_t as equal to 1 if turn t was undertaken and 0 otherwise. After
 37 defining \mathbf{d} as the $1 \times N_S$ vector with elements d_s , and \mathbf{a} as the $1 \times N_T$ vector with elements a_t , we
 38 can write

$$39 \quad y = \mathbf{d}\mathbf{X} + \mathbf{a}\mathbf{Z}.$$

40
 41 Further, we define a *trace* to be a contiguous sequence of observations from the same vehicle
 42 as it moves through the network. Figure 2 illustrates the relationship between observations and traces.
 43 Observations from the same trace are correlated through the link and turn correlation structures, while

1 we treat observations from different traces as independent. We assume here that there are no
 2 significant systematic differences between vehicles/drivers. Including a random effect at the vehicle
 3 level in the model to capture such differences is straightforward, however.
 4



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

6 We introduce the index $k = 1, \dots, N_K$ to represent a certain vehicle trace, where N_K is the
 7 number of traces in the data, and the index $r = 1, \dots, N_k$ to represent a certain observation in the trace,
 8 where N_k is the number of observations in the trace. The total number of observations in the data is
 9 $N_R = \sum_{k=1}^{N_K} N_k$. We define d_{krs} as the distance traversed on segments s and a_{krt} as equal to 1 if turn t
 10 was undertaken and 0 otherwise. We then define \mathbf{d}_{kr} as the $1 \times N_S$ vector with elements d_{krs} , and \mathbf{a}_{kr}
 11 as the $1 \times N_T$ vector with elements a_{krt} .

12 The trace k and observation r define the trip conditions as captured by the vector \mathbf{o}_{kr} , which
 13 affects the mean travel time rates \mathbf{g}_{kr} and delays \mathbf{h}_{kr} . The mean travel time for the trip defined by kr
 14 is then
 15

$$16 \mu_{kr} = E[y_{kr}] = \mathbf{d}_{kr} \mathbf{g}_{kr} + \mathbf{a}_{kr} \mathbf{h}_{kr} = (1 + \mathbf{o}_{kr} \boldsymbol{\beta}_O) \cdot (\mathbf{d}_{kr} \mathbf{B} \boldsymbol{\beta}_B + \mathbf{a}_{kr} \mathbf{E} \boldsymbol{\beta}_E).$$

17 For the travel time variance, the trip conditions are expressed at the level of the trace rather than the
 18 individual observation. That is, the vector of attributes \mathbf{p}_k may vary with k but not with r , which then
 19 also holds for the covariance matrices $\boldsymbol{\Omega}_k$ and $\boldsymbol{\Pi}_k$. The covariance between observations r and r' (or
 20 variance, in case $r = r'$), both in trace k , is
 21

$$22 \sigma_{krr'} = \text{Cov}[y_{kr}, y_{kr'}] = \mathbf{d}_{kr} \boldsymbol{\Omega}_k \mathbf{d}_{kr'}^T + \mathbf{a}_{kr} \boldsymbol{\Pi}_k \mathbf{a}_{kr'}^T =$$

$$23 (1 + \mathbf{p}_k \boldsymbol{\beta}_P)^2 \cdot$$

$$24 (\mathbf{d}_{kr} \mathbf{S} \boldsymbol{\Omega}_0 \mathbf{S}^T \mathbf{d}_{kr'}^T + \sum_{i=1}^{N_\rho} \rho_i \mathbf{d}_{kr} \mathbf{S} (\mathbf{W}_i \boldsymbol{\Omega}_0 + \boldsymbol{\Omega}_0 \mathbf{W}_i^T) \mathbf{S}^T \mathbf{d}_{kr'}^T + \sum_{i=1}^{N_\rho} \sum_{j=1}^{N_\rho} \rho_i \rho_j \mathbf{d}_{kr} \mathbf{S} \mathbf{W}_i \boldsymbol{\Omega}_0 \mathbf{W}_j^T \mathbf{S}^T \mathbf{d}_{kr'}^T +$$

$$25 \mathbf{a}_{kr} \boldsymbol{\Pi}_0 \mathbf{a}_{kr'}^T + \sum_{i=1}^{N_\lambda} \lambda_i \mathbf{a}_{kr} (\mathbf{Q}_i \boldsymbol{\Pi}_0 + \boldsymbol{\Pi}_0 \mathbf{Q}_i^T) \mathbf{a}_{kr'}^T + \sum_{i=1}^{N_\lambda} \sum_{j=1}^{N_\lambda} \lambda_i \lambda_j \mathbf{a}_{kr} \mathbf{Q}_i \boldsymbol{\Pi}_0 \mathbf{Q}_j^T \mathbf{a}_{kr'}^T).$$

26 For observation r in trace k and observation r' in trace k' , meanwhile, we have $\text{Cov}[y_{kr}, y_{k'r'}] = 0$.

27 Under the assumption of travel time rates and delays distributed multivariate normal, the travel
 28 time observations are also multivariate normal, with means and covariances given above. It is
 29 therefore possible to formulate a closed form joint likelihood function for the travel times, and the
 30 model parameters can be estimated using standard numerical ML techniques. For the likelihood
 31 formulation, we introduce the $N_k \times 1$ observation vector \mathbf{y}_k with elements y_{kr} , the $N_k \times 1$ mean
 32 vector $\boldsymbol{\mu}_k(\boldsymbol{\beta}_B, \boldsymbol{\beta}_C, \boldsymbol{\beta}_E, \boldsymbol{\beta}_O)$ with elements μ_{kr} , and the $N_k \times N_k$ covariance matrix $\boldsymbol{\Sigma}_k(\boldsymbol{\beta}_P, \sigma_U^2, \sigma_V^2, \boldsymbol{\rho}, \boldsymbol{\lambda})$
 33 with elements $\sigma_{krr'}$. Given all observed travel times \mathbf{y} , the log likelihood function is
 34

$$LL(\boldsymbol{\beta}_B, \boldsymbol{\beta}_C, \boldsymbol{\beta}_E, \boldsymbol{\beta}_O, \boldsymbol{\beta}_P, \sigma_U^2, \sigma_V^2, \boldsymbol{\rho}, \boldsymbol{\lambda} | \mathbf{y}) = -\frac{1}{2} N_R \log(2\pi) - \frac{1}{2} \sum_{k=1}^{N_K} (\mathbf{y}_k - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{y}_k - \boldsymbol{\mu}_k) - \frac{1}{2} \sum_{k=1}^{N_K} \log |\boldsymbol{\Sigma}_k|.$$

Estimation Aspects

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, 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 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 3.

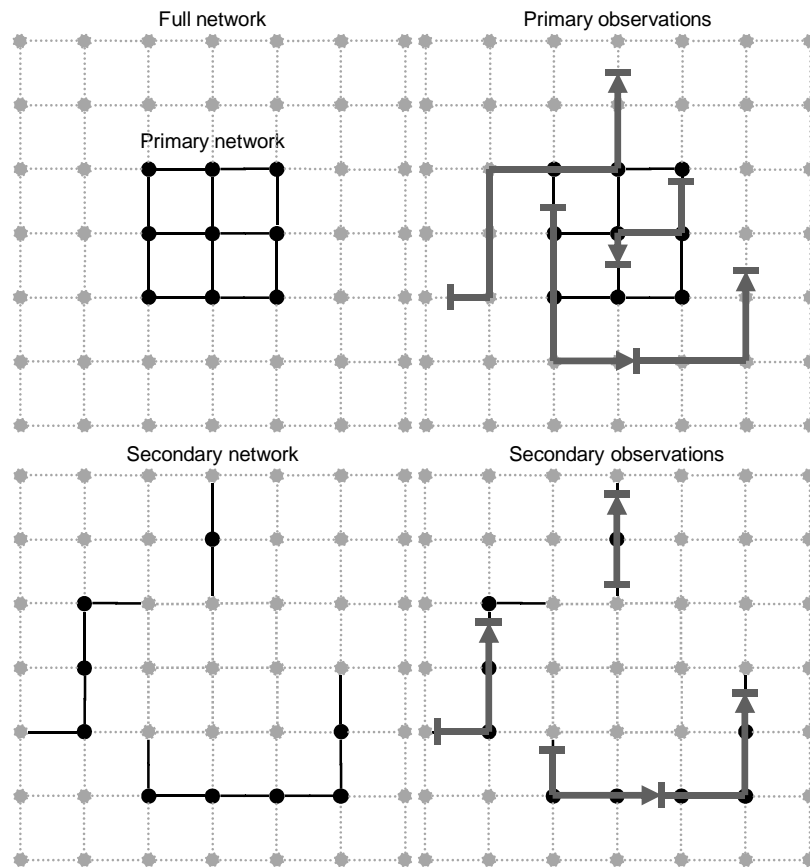


Figure 3: Illustration of the concepts of the primary and the secondary network.

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

1 frequency and, if the vehicle trajectories are sampled at random locations, with the number of
2 observations. The network model represent another projection from the network components to the
3 parameter space. Whether the parameters of the network model are identifiable through the data
4 depends on the relative dimensions of the parameter space and the observation space: the higher the
5 sampling frequency and the larger the number of observations, the larger the number of parameters
6 that can be identified. This determines, for example, to what extent we can include fixed effects for
7 specific segments or groups of segments in the model.

8 *Automated Grouping of Network Segments*

9 This section describes an automated method for partitioning a network into connected groups of
10 adjacent links. The algorithm allows the analyst to specify minimum, typical and maximum threshold
11 values for the number of links in each group and the number of observations covering each group,
12 respectively. The assumption is that fixed effects for these groups in the estimation can identify
13 systematic differences in segment travel time rates between different parts of the network not captured
14 by attributes such as speed limits. The approach makes it possible to find a feasible compromise
15 between the two extremes of fixed travel time rate effects for each segment and a single baseline travel
16 time rate for all segments. It can easily be used together with a manual portioning method for some
17 part of the network. In the case study below, for example, we manually partition the primary network
18 into link groups based on assumed similarity of traffic characteristics, while we use the automated
19 partition method for the secondary network.

20 The algorithm is initialized with each link belonging to a separate group; links in manually
21 defined groups are excluded. In each iteration, the group with the smallest number of observations is
22 selected (in case of ties, choose an arbitrary group). For each link in the current group, it is checked
23 whether it is connected with some link in another group through a common node. All identified
24 adjacent groups are added to a list. Going through the list in the order of increasing number of
25 observations, the current group is merged with the first adjacent group for which any of two conditions
26 hold:
27

- 28 1. The total number of links and the total number of observations in the two groups do not
29 exceed the typical values,
- 30 2. The number of links or the number of observations in any of the groups is less than the
31 minimum value, and the total number of links and the total number of observations in the
32 two groups do not exceed the maximum values.

33 The algorithm stops when no groups can be merged further. Note that any constraint can be made non-
34 binding by setting the corresponding threshold value sufficiently low or high. Since the algorithm
35 works at the link level, it ensures that all segments of the same link belong to the same group.

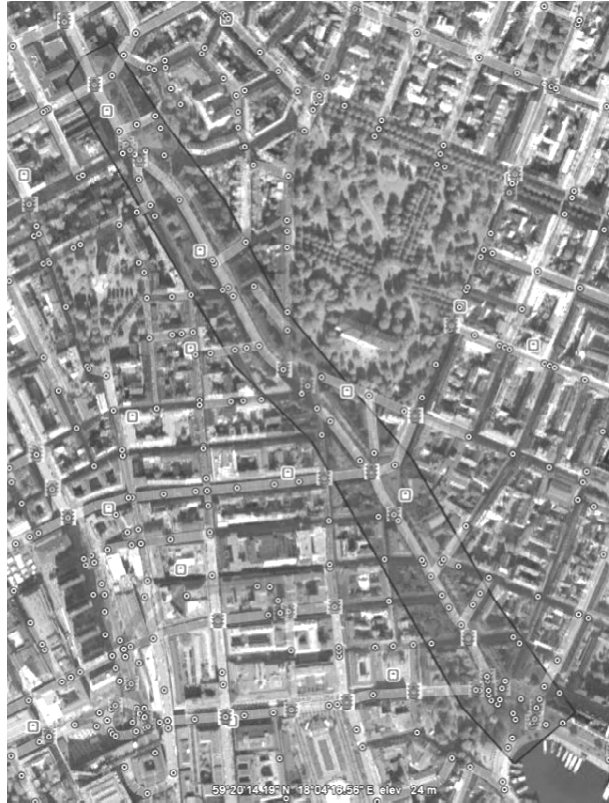
36 **CASE STUDY**

37 **Analysis Description**

38 We now describe an application of the model in the urban network of Stockholm, Sweden. The
39 primary network consists of only one route and stretches about 1.4 km along one of the main inner city
40 streets, the southern half of Birger Jarlsgatan, southbound direction, shown in Figure 4. The study
41 route is chosen to coincide as closely as possible with a pair of automatic number plate recognition
42 (ANPR) sensors mounted at each end of the route.

43 In the first part of the empirical study, we consider a few different specifications of the model
44 structure given above. The focus here is not to derive the best model possible, but to demonstrate the
45 concepts of the model and impact and significance of different explanatory factors on the observed
46 travel times during a specific time interval (7:30-8:00 AM). In the second part, we evaluate the
47 estimated travel time for the main route based on a rich model specification for different trip
48 conditions, and we perform a sensitivity analysis regarding the filtering of the observations. We then
49 compare the estimated route travel times with observed travel times from the ANPR sensors across
50 different periods of the day (6:00 AM to 10:00 PM).

1



2
3 **Figure 4: The case study area in Stockholm, Sweden. The shaded outlined area shows the main route, i.e.,**
4 **the primary network.**

5
6 **Data**

7
8 The GPS probe data are obtained from the fleet dispatching system of a taxi company operating in
9 total ca. 1500 vehicles in the Stockholm network. The data and the map-matching and path inference
10 used to obtain the necessary estimation input are described in (5). According to the technical
11 specifications, vehicles are sampled every 600 meters if occupied or every 400 meters if free, with a
12 minimum threshold at 1 minute since the last sampling. In practice, the average sampling frequency in
13 our data is about one sampling per 2 minutes and 780 meters, which is considerably lower than in most
14 previously reported studies (e.g., one per minute in (9,10), one per 200 meters in (11)).

15 In our baseline filtering procedure, observations covering less than 200 meters or more than 3
16 minutes are discarded as too unreliable in terms of map-matching and path inference (we do not
17 currently control for the time-based truncation in the estimation). In fact, the sampling rule
18 specifications suggest that there should be no observations covering less than 400 meters, which adds
19 additional uncertainty to short distance observations. We also remove observations with average
20 speeds exceeding 90 km/h.

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

26 The primary network contains 28 links, divided into 36 segments in total, 26 intersections and
27 10 traffic signals; the speed limit is constant at 50 km/h.

28 We use data for weekdays (Monday to Friday excluding holidays) and the time interval 7:30-
29 8:00 AM between January 14, 2010 and September 15, 2011. For this period, we have 43,324
30 observations in 33,143 traces after filtering. Of these, 8029 observations are primary and 35,296 are
31 secondary. Across the primary and the secondary networks, the observations cover in total 1227
32 segments, 783 links and 1198 turns. On average, each observation covers 14.5 segments, 8.9 links and

7.9 turns. Figure 5 shows a histogram of the average speed of the vehicle for each observation. The speed distribution has mean 21.8 km/h, median 20.4 km/h and standard deviation 9.4 km/h.

We have travel time data available from the ANPR sensors at the individual vehicle level for the period August 15 2011-September 15 2011, which we use to evaluate the estimation results below. Further, we have historical weather 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.).

Model Specification

Specification of link groups

We partition the network links into groups using a manual approach for the primary network combined with the automated method described above for the secondary network. First, the primary network 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 presumed variability of traffic conditions along the route; for example, there is a commercial and entertainment center in the middle of the route with a nearby taxi stop, where we hypothesize that the mean travel time rate is higher than in adjacent groups.

Second, the secondary network is partitioned with the algorithm above using the following threshold values which were found to produce good results. For the number of observations covering each group, the minimum value is set to 300 and the typical and maximum values are set to infinity. For the number of links in each group, the minimum, typical and maximum values are set to 10, 15 and 35, respectively. This produces 43 link groups for the secondary network, each containing 18.1 links and 27.7 segments on average. The distribution of the number of observations covering each link group is shown to the right in Figure 5. Together with the seven link groups for the primary network, we have 50 link groups in total.

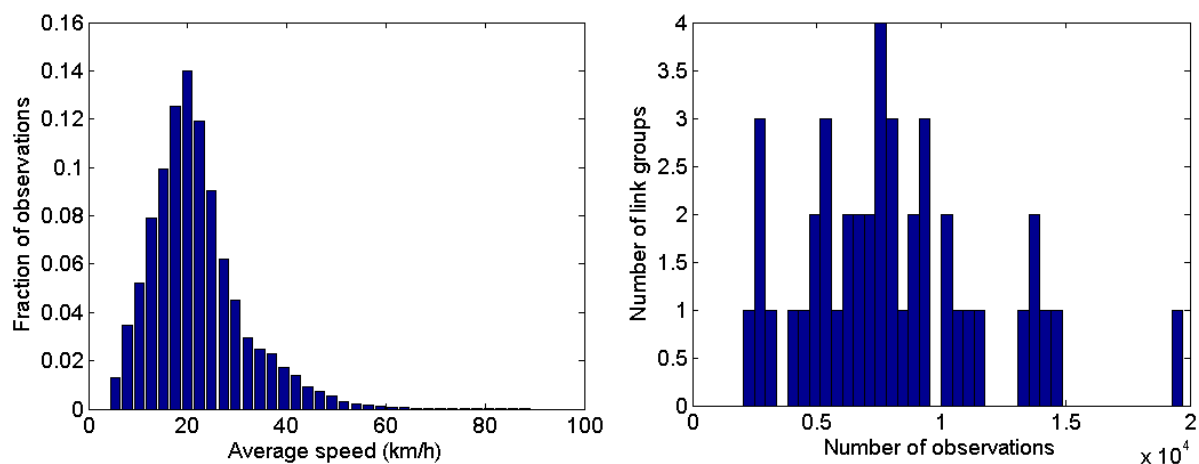


Figure 5: Left: Distribution of average speeds among the observations (7:30-8:00 AM). The speed distribution has mean 21.8 km/h, median 20.4 km/h and standard deviation 9.4 km/h. Right: The number of observations covering each link group in the secondary network.

Model 1

In order to demonstrate the performance of the model, we consider four different specifications, summarized in Table 1. In the first specification (Model 1), the mean segment travel time rates are explained only with fixed effects at the link group level (50 groups in total) which form the design matrix **B**. We do not include explanatory variables for segment or link attributes, trip conditions or turn penalties.

For the variance structure (**U**), we hypothesize that the secondary network is more heterogeneous in its characteristics, since it is larger and extends in all directions. Therefore, we allow the variance of primary and secondary network links to be different.

Further, we specify an SMA error structure for the link travel time rates. We hypothesize first that, *ceteris paribus*, dependence decreases with distance, in terms of both travel distance and the number of turns. Second, since correlation should largely follow traffic flows, we assume that the influence from a downstream link depends on the fraction of the flow on the upstream link that also traverses the downstream link. To estimate the traffic flow across the links, we use the taxi observations, assuming that they are approximately representative of the overall traffic flow. Otherwise, we are aware that some endogeneity may be introduced in the estimation. Based on the hypothesis and some experimentation, we introduce two weight matrices \mathbf{W}_1 and \mathbf{W}_2 with parameters ρ_1 and ρ_2 . The element $(\mathbf{W}_1)_{ll'}$ is proportional to the fraction of the flow on link l that also traverses the downstream link l' ; it is zero if l' is not directly downstream of l . The element $(\mathbf{W}_2)_{ll'}$ is defined in the same way but for links that are further away from l ; it is zero if l' is immediately downstream of l . For practical reasons we set $(\mathbf{W}_2)_{ll'}$ to zero if the two links are more than five turns apart in the network. Each matrix is normalized so that the largest element is equal to 1.

Model 2

In the second specification (Model 2), we add explanatory variables for link attributes and turn penalties, but no variables for trip conditions. This is relevant for applications at a strategic level, where typical travel times across different travel conditions are of interest. 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 \mathbf{B}):

- Dummy variables for every combination of segment speed limit (in km/h) and segment functional class (abbreviated FC, ranging from 1 to 5, where 1 is the highest level of road) in the data: (10, 5), (30, 3), (30, 4), (30, 5), (50, 1), (50, 2), (50, 4), and (50, 5). The reference level is (50, 3).
- One dummy variable for each of the cases that the segment is part of a roundabout, ramp, tunnel, or 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 \mathbf{D}):

- A dummy variable for the case that the turn involves a traffic signal. Further, dummy variables for the cases that the signalized turn is from a FC 4 segment to a FC 3 segment, from a FC 5 segment to a FC 3 segment, and from a FC 5 segment to a FC 4 segment, respectively. The increases in functional classes are used as proxies for low priority in the signal cycle.
- Dummy variables for the cases that the turn is non-signalized and from a FC 4 segment to a FC 3 segment, from a FC 5 segment to a FC 3 segment, and from a FC 5 segment to a FC 4 segment, respectively. The increases in functional classes are used as proxies for low priority in the intersection.

For the travel time rate variance structure (\mathbf{U}), we add dummy variables for link length < 50 m, speed limit ≤ 30 , functional class ≥ 4 and the link containing a taxi stop. Our hypothesis is that each of these cases contributes to a higher link travel time rate variance.

For the turn penalty variance structure (\mathbf{V}), we allow the variance of primary and secondary network turns to be different (a turn belongs to the primary network if both the origin and the destination links belong to the primary network). The variance is allowed to be non-zero only for turns which may have non-zero means according to the matrix \mathbf{D} .

We also specify an SMA error structure for the turn penalties. We assume that two turn penalties are correlated only when they both involve traffic signals. Thus, we introduce a weight matrix \mathbf{Q} where element $(\mathbf{Q})_{tt'}$ is equal to 1 if both turns t and t' involve traffic signals and 0

1 otherwise. For practical reasons we set $(\mathbf{Q})_{tt'}$ to zero if the two traffic signals are more than four links
 2 apart in the network.

3 4 *Model 3*

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

- 7
- 8 • Dummy variable for the late half of the interval, i.e., 7:45-8:00 (not July).
- 9 • One dummy variable for each weekday from Tuesday to Friday (reference level is Monday).
- 10 • Dummy variables for the seasons Jun-Aug, Sep-Nov and Dec-Feb (reference level is Mar-
- 11 May).
- 12 • Dummy variable for the month of July (the main vacation period).
- 13 • Recent snow: Number of hours of consecutive snowfall reports (qualitatively reported as
- 14 “snow” or “rain-snow”) preceding the trip.
- 15 • Rain: Dummy variable for the most recent weather report being “rain”.
- 16 • Dummy variable for the taxi being free.

17 18 *Model 4*

19 The fourth specification (Model 4) is identical to Model 3 except that all links and turns are assumed
 20 to be independent. By comparing the performance of Model 3 and Model 4, we can evaluate the
 21 contribution of the correlation structure in the model.

22 23 **Estimation Results**

24
25 The model was estimated using the ML estimation routine in the Statistics Toolbox for MATLAB (17)
 26 and a trust-region reflective Newton optimization algorithm. Gradients were evaluated analytically
 27 while the Hessian used to calculate standard errors was calculated numerically.

28 Estimation results are shown in Table 1. There are several interesting observations to be made.
 29 To begin with, the segment group fixed effects are all significant with T statistics greater than 4 in all
 30 four models. This demonstrates that the automated partition method is capable of handling the
 31 identification problems associated with the low sampling frequency, even in the outer parts of the
 32 secondary network. Also, the fit of Model 1, captured by the log-likelihood and Akaike’s information
 33 criterion (AIC), is considerably better than a model with a single mean travel time rate parameter and
 34 a single travel time rate variance parameter (the log-likelihood for this model is 121,371, AIC is -
 35 242,738).

36 The fit of the model increases greatly when segment and turn attributes are added (Model 2
 37 vs. Model 1). This encouraging result suggests that the low sampling frequency and identification
 38 power of the observations can be compensated for to some extent by making use of attributes of the
 39 network components. As expected, speed limits and functional classes have a strong impact on travel
 40 time rates. Further, travel time rates are higher on shorter links. A taxi stop on the segment reduces the
 41 mean speed by about 30 km/h.

42 A traffic signal gives a delay that is particularly large if the trip involves a turn from a FC 4
 43 segment to a FC 3 segment, or from a FC 5 to a FC 3 or FC 4 segment (ca. 15 seconds in the latter
 44 case). A non-signalized turn also gives a delay, in particular from a FC 4 segment to a FC 3 segment
 45 (18 seconds).

46 Adding explanatory variables for the conditions of the trip improves the fit of the model
 47 further (Model 3 vs. Model 2). Mean travel time rates are significantly higher in the late half of the
 48 period (5%), suggesting a build-up of the morning peak. There is also variation across the week with
 49 travel time rates peaking on Wednesdays, 2% higher than on Mondays. Travel time rates are
 50 considerably lower during the summer and fall than during spring, in particular during the main
 51 vacation month July (10% lower compared to spring). The travel time variance also varies
 52 significantly at the different time scales.

53 Recent snowfall is found to significantly increase mean travel time rates. The estimate
 54 suggests that every four hours of consecutive snow before the trip increases travel time rates about 1%.

1 Meanwhile, we find no significant impact of rain on travel time rates. We further find that travel time
 2 rates are significantly lower when the taxi is hired, which may be due to a more determined driving
 3 behavior and that less time is spent cruising for customers at low speeds.

4 Treating links and traffic signal penalties as independent reduces the fit of the model
 5 considerably (Model 4 vs. Model 3). This suggests that correlations between network components
 6 should be considered when estimating travel times. In Models 1-3 the parameters for the correlation
 7 structures are strongly significant and negative throughout. We find the negative values of the
 8 parameters unexpected, at least for the link travel time rates. While we would expect travel times on
 9 adjacent links to be positively correlated, the results here suggest that the correlation is instead
 10 negative. These results are robust to different model specifications, filtering rules and time intervals
 11 across the day. The characteristics of inter-link correlation will be further investigated in future work.
 12

13 **Table 1: Estimation results (7:30-8:00 AM).**

14 Parameters	Model 1	Model 2	Model 3	Model 4
15 Mean segment TT rates (B) [h/km]:				
16 50 segment groups	yes	yes	yes	yes
17 Speed limit 10, functional class 5	-	0.0556 (5.88)	0.0479 (5.27)	0.0483 (5.54)
18 Speed limit 30, functional class 3	-	0.0150 (15.8)	0.0143 (15.6)	0.0144 (15.5)
19 Speed limit 30, functional class 4	-	0.0194 (20.7)	0.0177 (19.4)	0.0180 (19.2)
20 Speed limit 30, functional class 5	-	0.0231 (26.1)	0.0218 (25.3)	0.0223 (25.5)
21 Speed limit 50, functional class 1	-	-0.0054 (-1.43)	-0.0060 (-1.66)	-0.0053 (-1.48)
22 Speed limit 50, functional class 2	-	0.0063 (1.51)	0.0059 (1.48)	0.0049 (1.16)
23 Speed limit 50, functional class 4	-	0.0143 (12.9)	0.0140 (13.0)	0.0129 (11.7)
24 Speed limit 50, functional class 5	-	0.0108 (7.25)	0.0093 (6.51)	0.0094 (6.45)
25 Roundabout	-	-0.0673 (-12.0)	-0.0649 (-12.0)	-0.0600 (-10.8)
26 Ramp	-	0.0051 (1.17)	0.0050 (1.21)	0.0042 (1.01)
27 Tunnel	-	0.0118 (4.87)	0.0115 (4.95)	0.0103 (4.30)
28 One way	-	-0.0029 (-4.50)	-0.0026 (-4.19)	-0.0026 (-4.11)
29 Taxi stop	-	0.0440 (14.9)	0.0411 (11.2)	0.0418 (14.7)
30 Link length < 50 m	-	0.0059 (3.13)	0.0047 (2.60)	0.0037 (2.01)
31 Link length ≥ 200 m	-	-0.0116 (-10.9)	-0.0117 (-11.4)	-0.0096 (-9.40)
32 Mean turn penalties (D) [s]:				
33 Traffic signal	-	1.4135 (4.67)	1.4541 (5.01)	1.6292 (5.51)
34 From FC 4 to FC 3, traffic signal	-	6.2882 (8.44)	5.7215 (7.97)	6.6961 (9.19)
35 From FC 5 to FC 3, traffic signal	-	8.9121 (10.1)	8.5582 (10.1)	9.7153 (11.3)
36 From FC 5 to FC 4, traffic signal	-	13.2707 (11.3)	13.0008 (11.5)	13.0312 (11.4)
37 From FC 4 to FC 3, no traffic signal	-	18.3608 (9.38)	16.3390 (8.70)	17.5524 (8.87)
38 From FC 5 to FC 3, no traffic signal	-	0.0838 (0.08)	-0.2435 (-0.24)	-0.1356 (-0.14)
39 From FC 5 to FC 4, no traffic signal	-	3.5022 (2.79)	3.4445 (2.84)	3.7054 (3.01)
40 Trip conditions mean TT (o):				
41 7:45-8:00 (not July)	-	-	0.0505 (12.7)	0.0505 (12.1)
42 Tuesday	-	-	0.0193 (3.20)	0.0197 (3.11)
43 Wednesday	-	-	0.0204 (3.39)	0.0195 (3.08)
44 Thursday	-	-	0.0134 (2.27)	0.0129 (2.08)
45 Friday	-	-	0.0019 (0.33)	0.0029 (0.47)
46 Winter (Dec-Feb)	-	-	0.0027 (0.49)	0.0034 (0.59)
47 Summer (Jun-Aug)	-	-	-0.0459 (-9.31)	-0.0463 (-8.94)
48 Fall (Sep-Nov)	-	-	-0.0252 (-4.55)	-0.0237 (-4.05)
49 July	-	-	-0.0570 (-8.40)	-0.0592 (-8.35)
50 Recent snow [h ⁻¹]	-	-	0.0027 (3.24)	0.0026 (3.05)
51 Rain	-	-	0.0104 (0.43)	0.0067 (0.27)
52 Free	-	-	0.0584 (12.8)	0.0566 (11.8)
53 Trip conditions TT variance (p):				
54 7:45-8:00 (not July)	-	-	0.0445 (5.87)	0.0483 (6.36)
55 Tuesday	-	-	0.0462 (3.97)	0.0460 (4.42)
56 Wednesday	-	-	0.0400 (3.43)	0.0387 (3.47)
57 Thursday	-	-	0.0434 (3.80)	0.0407 (3.58)
58 Friday	-	-	0.0515 (4.53)	0.0507 (4.68)
59 Winter (Dec-Feb)	-	-	-0.0264 (-2.52)	-0.0266 (-2.96)
60 Summer (Jun-Aug)	-	-	-0.0631 (-6.75)	-0.0675 (-7.25)
61 Fall (Sep-Nov)	-	-	-0.0330 (-3.10)	-0.0322 (-3.23)
62 July	-	-	-0.0732 (-5.59)	-0.0826 (-6.35)
63 Recent snow [h ⁻¹]	-	-	-0.0017 (-1.10)	-0.0017 (-1.14)
64 Rain	-	-	-0.0007 (-0.39)	-0.0010 (-0.54)
65 Free	-	-	0.0068 (0.87)	0.0115 (1.47)
66 Correlation TT rates:				
67 Flow fraction nearest link (W ₁)	-0.726 (-30.4)	-0.4698 (-26.9)	-0.4681 (-27.2)	-
68 Flow fraction more distant links (W ₂)	-0.099 (-16.1)	-0.0739 (-19.1)	-0.0722 (-19.1)	-
69 Correlation turn penalties:				
70 Traffic signal (Q)	-	-0.2048 (-33.0)	-0.2018 (-33.1)	-

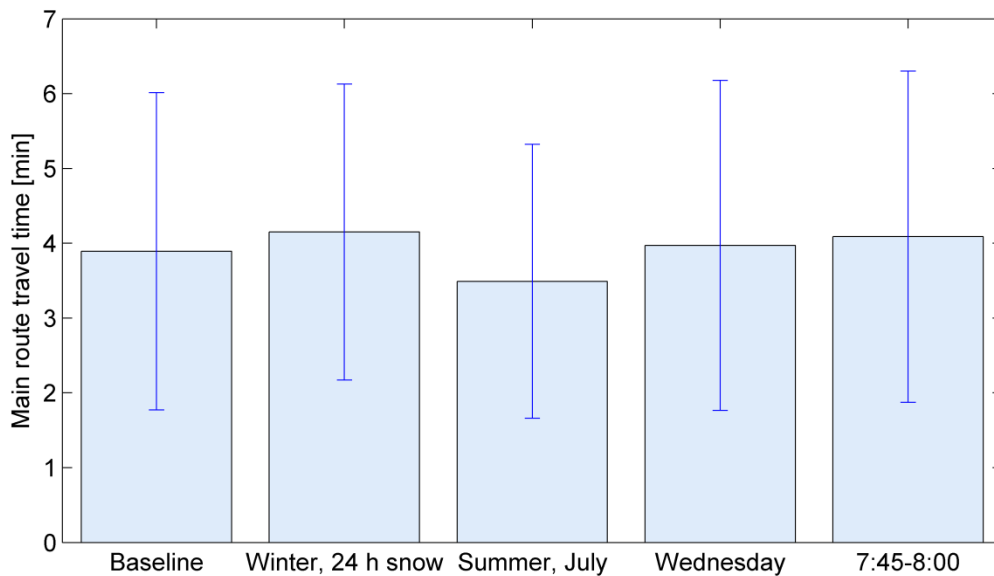
1	Variance TT rates (U) [10^{-2} (h/km) 2]:				
2	Primary network	0.0036	0.0267	0.0432	0.0255
3	Secondary network	0.0044	0.0477	0.0432	0.0452
4	Link length < 50 m	-	1.2573	1.2091	0.3854
5	Speed limit ≤ 30	-	0.1383	0.1322	0.0934
6	Functional class ≥ 4	-	0.1171	0.1102	0.0823
7	Taxi stop	-	0.1662	0.1664	0.0736
8	Variance turn penalties (V) [10^4 s 2]:				
9	Primary network	-	0.0297	0.0285	0.0291
10	Secondary network	-	0.0245	0.0232	0.0234
11	Log likelihood	125,403	130,158	130,622	129,442
12	AIC	-250,699	-260,149	-261,031	-258,675
13	Observations	43324	43324	43324	43324
14	Parameters	54	83	107	104

15

16 Main Route Travel Time Estimation

17

18 We estimate the travel time on the main route by applying the estimated parameters to a hypothetical
 19 observation traversing the route. We first consider the estimated travel time for the period 7:30-8:00
 20 AM under different trip conditions using the specification of Model 3. Figure 6 illustrates how the
 21 estimated mean and standard deviation vary under different combinations of trip conditions; the
 22 reference conditions are those of the period 7:30-7:45 on a Monday in spring with no rain or recent
 23 snow when the vehicle is hired.
 24



25

26 **Figure 6: Estimated travel time for the main route (7:30-8:00 AM). The baseline conditions are 7:30-7:45,**
 27 **Monday, Spring, no recent snow, no rain, hired taxi. The filled bars show mean values, the error bars**
 28 **show the 2.5% to the 97.5% percentile ranges.**

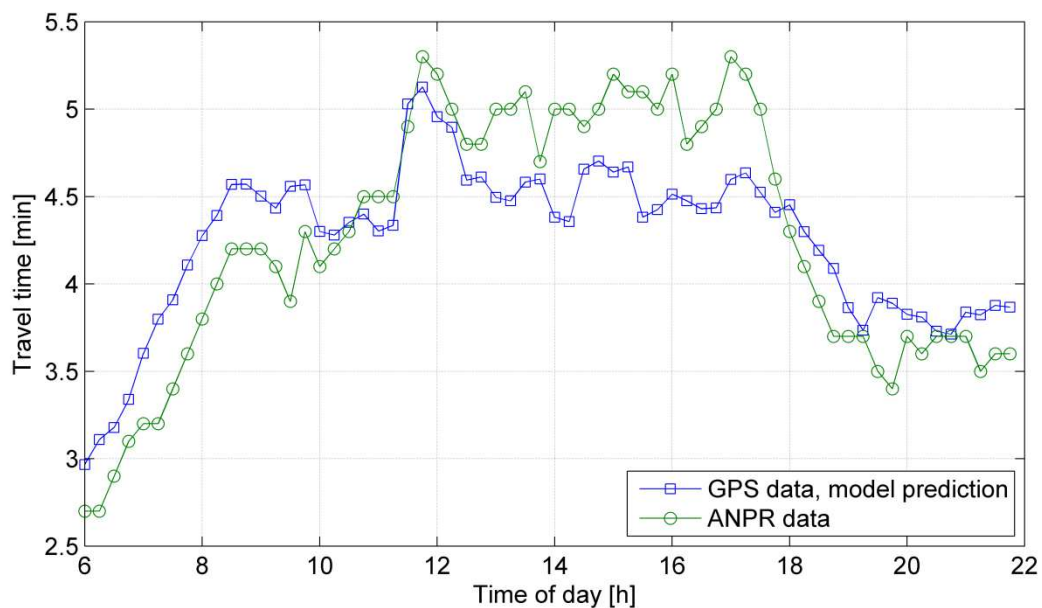
29

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

1 biased estimates of travel times. This is not a characteristic of the presented model but a general
 2 problem with using low frequency samplings of vehicle trajectories in travel time estimation.

3 We also compare the model estimates with observed travel times from the ANPR sensors for
 4 the period August 15-September 15 2011. We perform the estimation for every 30 minute interval
 5 from 6:00-6:30 AM to 9:30-10:00 PM, using the specification of Model 3 as basis; the dummy
 6 variable for the interval 7:45-8:00 AM is replaced with a corresponding dummy for the late half of
 7 each 30 minute interval. The estimation requires that we specify values for the trip condition variables,
 8 which should match the average conditions for the period of the ANPR data as closely as possible. We
 9 therefore add a dummy variable to the vectors \mathbf{o} and \mathbf{p} that is one during the period August 15-
 10 September 15 2011 and zero otherwise. The other variables are set to approximate the average values
 11 for the ANPR data period: the dummy for the late half of the interval is set to 0.5, weekday dummies
 12 are set to 0.2, summer and fall dummies are set to 0.5, remaining trip condition variables are set to 0.

13 Figure 7 shows the estimated mean travel times on the main route compared with the sample
 14 mean travel times from the ANPR sensors. The curves share the same general shape in that the travel
 15 times increase in the morning, flatten out during the day and decrease in the evening. There are some
 16 deviations also, in that the estimated values are higher than the ANPR values in the morning and lower
 17 in the middle of the day. It is important to note that the ANPR data are not ground truth for the taxi
 18 data but represent the overall traffic including private cars, trucks etc., which likely differs from the
 19 taxi traffic. For example, there is a bus line along a part of the route, which may be used by taxis but is
 20 restricted for private cars. Also, there is an entertainment center in the middle of the route which is a
 21 frequent taxi origin and destination at night.
 22



23
 24 **Figure 7: Mean travel time for the main route for each 15 minute interval between 6:00 AM to 10:00 PM.**
 25 **Squares show model estimates based on the taxi GPS data, circles show sample mean values from the**
 26 **ANPR data.**

27
 28 Deviations could also arise from other sources. First, the ANPR data is noisy and a
 29 considerable number of observations had to be filtered out as likely outliers, which makes the final
 30 values more uncertain. Second, it is not known exactly where the ANPR sensors register the vehicles
 31 (it may even vary between vehicles), which means that there may be systematic and random
 32 differences between the taxi route and the ANPR route. Third, the model estimates use observations
 33 for a much longer period of time and a much larger network, which may lead to differences. Fourth,
 34 there is always the possibility that the model misses some aspect of the true traffic conditions.
 35

1 Discussion

2
3 One problem with using opportunistic FCD is that the behavior of the equipped vehicles may differ
4 from that of the population of interest (e.g., private vehicles or the entire vehicle population). Taxis,
5 for example, can be expected to make more stops and low-speed cruising in some situations than the
6 population on average. In other situations, they may drive faster than the population average. Ideally,
7 we would want to separate each observation into a representative travel time for the population and a
8 sample bias, or to discard observations with too much bias. Advanced analysis of the vehicle
9 trajectories may provide some support, but in general one may have to resort to filtering the
10 observations from outliers before the estimation, with the risk that these represent actual traffic
11 conditions for the population, and accept that the results may not be entirely representative of the
12 population. The analysis here suggests that the estimates are relatively sensitive to how the filtering is
13 done, and future research effort should be put on making the path inference reliable also for long time
14 intervals between reports in order to minimize bias problems.

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

21 CONCLUSION

22
23
24 We have presented a statistical model for urban road network travel time estimation based on low
25 frequency GPS data. We use geometric attributes and trip conditions as explanatory variables in
26 combination with fixed effects to reduce the number of parameters to estimate and for predicting travel
27 times also in areas with very few observations. The model allows correlation between travel times on
28 different network links and presents a way to estimate the correlation as experienced by a driver
29 traversing the links sequentially.

30 The model was applied to a small example network in Stockholm, Sweden. We found that
31 geometric attributes and trip conditions have significant effects on travel times and that there is
32 significant correlation between links. Further work will be focused on the characteristics of correlation
33 between link travel times, on refining the automated partitioning of the network using statistical data
34 clustering methods, and on developing the network model to reflect the characteristics of uncongested
35 and congested traffic. We will also develop and adapt the model towards online prediction in
36 combination with real-time observations and traffic data from other types of sensors.

37 ACKNOWLEDGEMENTS

38
39
40 We would like to thank Tomas Julner from Trafik Stockholm for his support and provision of the data
41 to the *iMobility* Lab at KTH.

42 REFERENCES

- 43
44
45 1. Leduc, G. Road traffic data: collection methods and applications. JRC Technical Notes,
46 Working Papers on Energy, Transport and Climate Change, N.1, 2008.
47 2. van Aerde, M., B. Hellinga, L. Yu, and H. Rakha. Vehicle probes as real-time ATMS sources
48 of dynamic O-D and travel time data, *Large Urban Systems: Proceedings of the Advanced*
49 *Traffic Management Conference*, 2003, pp 207-230.
50 3. Zou, L., J.-M. Xu, and L.-X. Zhu. Arterial speed studies with taxi equipped with global
51 positioning receivers as probe vehicle. *Proceedings of the 2005 International Conference on*
52 *Wireless Communications, Networking and Mobile Computing*, 2005, pp. 1343-1347.
53 4. Work, D. B., O.-P. Tossavainen, S. Blandin, A. M. Bayen, T. Iwuchukwu, and K. Tracton. An
54 ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile

- 1 devices. *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 5062-
2 5068.
- 3 5. Rahmani, M., and H. N. Koutsopoulos. Path inference of low-frequency GPS probes for urban
4 networks. Working paper.
- 5 6. Miwa, T., T. Sakai, and T. Morikawa. Route identification and travel time prediction using
6 probe-car data. *International Journal of ITS Research*, vol. 2, 2008, pp. 21-28.
- 7 7. Miller, J., S. Kim, M. Ali, and T. Menard. Determining time to traverse road sections based on
8 mapping discrete GPS vehicle data to continuous flows. *IEEE Intelligent Vehicles Symposium*,
9 2010, pp. 615–620.
- 10 8. Hellenga, B., P. Izadpanah, H. Takada, and L. Fu. Decomposing travel times measured by
11 probe-based traffic monitoring systems to individual road links. *Transportation Research Part*
12 *C: Emerging Technologies*, vol. 16, 2008, pp.768–782.
- 13 9. Hunter, T., R. Herring, P. Abbeel, and A. Bayen. Path and travel time inference from GPS
14 probe vehicle data. Neural Information Processing Systems Foundation (NIPS) Conference,
15 Vancouver, Canada, 2009.
- 16 10. Herring, R., A. Hofleitner, P. Abbeel, and A. Bayen. Estimating arterial traffic conditions
17 using sparse probe data. *13th International IEEE Conference on Intelligent Transportation*
18 *Systems*, 2010, pp. 929-936.
- 19 11. Westgate, B. S., D. B. Woodard, D. S. Matteson, and S. G. Henderson. Travel time estimation
20 for ambulances using Bayesian data augmentation. Submitted, Journal of the American
21 Statistical Association, 2011.
- 22 12. Bernard, M., J. Hackney, and K. W. Axhausen. Correlation of segment travel speeds. 6th
23 Swiss Transport Research Conference, 2006.
- 24 13. de Fabritiis, C., R. Ragona, and G. Valenti. Traffic estimation and prediction based on real
25 time floating car data. *11th International IEEE Conference on Intelligent Transportation*
26 *Systems*, 2008, pp. 197–203.
- 27 14. Min, X., J. Hu, Q. Chen, T. Zhang, and Y. Zhang. Short-term traffic flow forecasting of urban
28 network based on dynamic STARIMA model. *12th International IEEE Conference on*
29 *Intelligent Transportation Systems*, 2009, pp. 461-466.
- 30 15. Herring, R., A. Hofleitner, S. Amin, T. Abou Nasr, A. Abdel Khalek, P. Abbeel, and A.
31 Bayen. Using mobile phones to forecast arterial traffic through statistical learning. TRB
32 Annual Meeting, 2010.
- 33 16. Cheng, T., J. Haworth, and J. Wang. Spatio-temporal autocorrelation of road network data.
34 *Journal of Geographical Systems*, in press, 2011.
- 35 17. MATLAB version 7.8.0. Natick, Massachusetts: The MathWorks Inc., 2009.
- 36