

Evaluating the Added-Value of Online Bus Arrival Prediction Schemes

Cats, Oded; Loutos, G

DOI

[10.1016/j.tra.2016.02.004](https://doi.org/10.1016/j.tra.2016.02.004)

Publication date

2016

Document Version

Accepted author manuscript

Published in

Transportation Research. Part A: Policy & Practice

Citation (APA)

Cats, O., & Loutos, G. (2016). Evaluating the Added-Value of Online Bus Arrival Prediction Schemes. *Transportation Research. Part A: Policy & Practice*, 86, 35-55. <https://doi.org/10.1016/j.tra.2016.02.004>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Evaluating the Added-Value of Online Bus Arrival Prediction Schemes

Oded Cats

Department of Transport and Planning, Delft University of Technology, The Netherlands

AND

Department of Transport Science, Royal Institute of Technology (KTH), Stockholm, Sweden

o.cats@tudelft.nl

Gerasimos Loutos (Corresponding author)¹

Department of Transport Science, Royal Institute of Technology (KTH), Stockholm, Sweden

gerasimos.loutos@vti.se

Cite as:

Cats O. and Loutos G. (2016). Evaluating the Added-Value of Online Bus Arrival Prediction Schemes. <i>Transportation Research Part A: Policy and Practice</i> , 86, 35-55.
--

Keywords: Real-time information, Reliability, Waiting times, Public transport, Travel time prediction

¹ Current contact details: Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden and Linköping University, Department of Science and Technology (ITN), Norrköping, Sweden. Tel. +46 13 20 40 56

Abstract

Online predictions of bus arrival times have the potential to reduce the uncertainty associated with bus operations. By better anticipating future conditions, online predictions can reduce perceived and actual passenger travel times as well as facilitate more proactive decision making by service providers. Even though considerable research efforts were devoted to the development of computationally expensive bus arrival prediction schemes, real-world real-time information (RTI) systems are typically based on very simple prediction rules. This paper narrows down the gap between the state-of-the-art and the state-of-the-practice in generating RTI for public transport systems by evaluating the added-value of schemes that integrate instantaneous data and dwell time predictions. The evaluation considers static information and a commonly deployed scheme as a benchmark. The RTI generation algorithms were applied and analyzed for a trunk bus network in Stockholm, Sweden. The schemes are assessed and compared based on their accuracy, reliability, robustness and potential waiting time savings. The impact of RTI on passengers waiting times are compared with those attained by service frequency and regularity improvements. A method which incorporates information on downstream travel conditions outperforms the commonly deployed scheme, leading to a 25% reduction in the mean absolute error. Furthermore, the incorporation of instantaneous travel times improves the prediction accuracy and reliability, and contributes to more robust predictions. The potential waiting time gains associated with the prediction scheme are equivalent to the gains expected when introducing a 60% increase in service frequency, and are not attainable by service regularity improvements.

1. Introduction

Service reliability is one of the main determinants of public transport level-of-service. Public transport services are subject to several sources of uncertainty which results with deviations from the planned service. This is particularly true for bus systems where the difficulty to predict bus arrival times leads to longer waiting times, passenger dissatisfaction and higher operational costs. Public transport systems are increasingly equipped with information and communication technologies such as automatic vehicle location (AVL) and automatic passenger counts (APC). These systems were first deployed to support fleet monitoring and frequency determination but later also facilitated the generation and dissemination of real-time information (RTI) (TCRP 2008). Information concerning next bus arrival is considered among the most important source of information and its accuracy is one of the main concerns among bus users (Caulfield and O'Mahony 2007, Rahman et al. 2013).

Even though considerable research efforts were devoted to the development of bus arrival prediction schemes, which involved the application of computationally-intensive statistical methods, there is lack of knowledge on the performance of real-world RTI systems. While it is often postulated that the accuracy and reliability of RTI are fundamental in realizing its potential benefits, there is limited knowledge on how RTI systems perform in practice. In order to gain a better understanding of the current state of the practice, an empirical analysis of the performance of a commonly deployed timetable-based prediction scheme was conducted by Cats and Loutos (2013). The analysis indicated that the provisioned information deteriorates significantly along the line and becomes increasingly unreliable for prediction horizons longer than several minutes. The results of the empirical study call for the development of alternative schemes for generating RTI. These schemes will address the identified shortcomings, while using the performance of the current system as a benchmark.

The primary objective of this study is to evaluate the added-value of alternative online bus arrival prediction schemes. The added-value is defined in this study in terms of passenger waiting time savings and the capability to foresee downstream vehicle trajectories and thus facilitate the deployment of proactive operations management. The performance of the prediction schemes is assessed in terms of their accuracy, reliability and robustness. The added-value of the proposed schemes is appraised by benchmarking their performance against static information and a commonly deployed prediction scheme. Furthermore, the potential time savings associated with RTI dissemination and improvements, as compared with more expensive measures to improve service frequency and reliability, are examined. We implement and evaluate alternative link-based online prediction schemes.

Two state-of-the-art schemes are designed to enrich the state-of-the-practice by (a) embedding data on downstream travel times and, (b) considering the underlying mechanism that drives the bunching process, namely the impact of headway on dwell times through passenger flows. The schemes were inspired by the mathematical model of Newell and Potts (1964) for bus service reliability, which first established the fundamental relations between bus travel time elements. By analyzing prediction schemes that are directly applicable, rely on existing information sources, account for the prevailing control strategy, and do not involve any estimation or calibration techniques prior to their implementation, this study narrows down the gap between the state-of-the-art and the state-of-the-practice in generating RTI for public transport systems.

Following the literature review (Section 2), the formulation of alternative prediction schemes (Section 3) is presented. The schemes are applied to a trunk bus network in Stockholm, Sweden (Section 4). An analysis and evaluation of their performance, in terms of their accuracy, reliability, robustness and added-value in terms of waiting time savings, as compared with improvements in service provision is provided in Section 5. This paper concludes with a discussion of performance implications (Section 6), practical considerations and suggestions for further research (Section 7).

2. Literature review

There is an extensive literature on public transport prediction models. Since this study is concerned with the evaluation of methods to generate RTI, prediction methods are reviewed in Section 2.1 with a focus on their data requirements and online application considerations. In Section 2.2, approaches and findings related to measuring the impacts of RTI dissemination and its quality on service performance are reviewed.

2.1 Public transport prediction models

Predicting future public transport states, such as vehicle arrival and departure times, and on-board crowding levels, requires the collection, integration and process of instantaneous and historical data.. Most research efforts were devoted to the development of bus arrival time predictions because of the importance of waiting times and the uncertainty associated with bus operations. Thus, vehicle arrival time constitutes the primary state to be obtained from prediction schemes. Previous studies applied various statistical and meta-heuristic methods for bus arrival predictions including: regression models (Patnaik et al. 2004, Chang et al. 2010), artificial neural networks (Jeong and Rilett 2005, Mazloumi et al. 2011), Kalman filter (Cathey and Dailey 2003; Shalaby and Farhan 2004, Chen et al. 2004, 2005), support vector machines (Yu et al. 2011), genetic algorithms (Fadaei Oshyani and Cats 2014) and statistical pattern recognition (Vu and Khan 2010).

Regression models were used to predict the remaining travel time as function of independent variables. Patnaik et al. (2004) estimated a series of linear regression models for bus arrival times as function of distance, number of stops, passenger volumes and weather conditions. However, in general, variables in public transport operations are highly inter-correlated and exercise complex non-linear relations. These properties can therefore hinder the applicability of regression models. Non-parametric regression models such as the k-nearest neighbors proposed by Chang et al. (2010) can be effective when applied to large datasets which are underlined by non-linear relations. However, their performance is undermined by long computational time because of their reliance on large amounts of historical data.

Machine learning methods can be advantageous over statistical methods in predicting future states of the public transport system due to their capability to utilize large amounts of data, to reveal complex patterns and to address noise in data streams. Artificial neural networks offer an effective algorithm to generate outcomes from complex non-linear systems. Moreover, their application does not require independence between prediction factors. Jeong and Rilett (2005) and Mazloumi et al. (2011) developed artificial neural networks for bus arrival predictions. While this method was reported to outperform other techniques, large amounts of data are needed to construct the network and the training phase requires extensive computational efforts.

The capability of Kalman filter to constantly update the state variable estimation proved instrumental in obtaining high computational speed. Previous studies used AVL (Chen et al. 2004, 2005), APC (Cathey and Dailey 2003) or both (Shalaby and Farhan 2004) when applying the Kalman filter to bus arrival predictions. The algorithm requires real-time feeds at every update while data fluctuations might cause difficulties in solving the time lag. For example, Vu and Khan (2010) applied a statistical pattern recognition technique to mine AVL and APC archives and found it more reliable than Kalman filter when handling unusual events of bus operation.

Support vector machine algorithms are based on statistical learning theory. The solution of support vector machine is always unique and globally optimal since the process of reaching the solution is similar to a linearly constrained quadratic programming problem. [Bin et al. \(2006\) demonstrated that support vector machine are capable of accurately predict bus arrival times by integrating multiple sources of information without requiring the explicit formulation of their relation.](#) Yu et al. (2011) tested different models - linear regression, k-nearest neighbors, artificial neural networks and support vector machine - for the prediction of arrivals at multiple bus routes at common stops and compared them against arrival data collected by manual surveys. The authors concluded that the support vector machine method outperformed the rest of the models. Hans et al. (2015) generated bus trajectory predictions using a micro-simulation model with random travel and dwell times. The link between bus headways and passenger flows allowed considering the non-linear relationships of travel and dwell times.

Most of the abovementioned studies considered solely vehicle-based performance metrics and often relied on simulated data. Fadaei Oshyani and Cats (2014) proposed a hybrid model for predicting vehicle trajectories as well as countdown displays. The model was tested on AVL data from several trunk lines. The hybrid model integrates schedule, instantaneous and historical data. The weights of each data source were optimized using a genetic algorithm. The results suggested that although the route-specific estimations obtained more accurate predictions, their added-value might not be justified in light of the additional complexity involved in its implementation.

The prediction schemes reported in the literature either involve the estimation and calibration of a set of parameters tailored to a specific application or require large amounts of historical data and long computational time. The optimization process could become computationally expensive and hinder their online implementation. While being very efficient, machine learning techniques do not provide a tractable formulation of the prognosis logic which relates predictions to the underlying service mechanisms.

2.2 Evaluating the impacts of public transport predictions

Online public transport predictions can reduce uncertainty as well as facilitate more adaptive decision making by system users as well as system operators. Previous studies provide evidence that the dissemination of RTI improves user satisfaction with the public transport service, reduces their perceived and actual waiting times and increases ridership. Several empirical studies found that passengers' perceived waiting times decreased substantially after RTI systems were deployed, although waiting times were still overestimated by passengers (Mishalani et al. 2006, Dziekan and Kottenhoff 2007, Chow et al. 2014). Moreover, waiting times were shorter among users that accessed RTI from their mobile phone due to either departure time choice (Watkins et al. 2011, Brakewood et al. 2014) or route choice effects (Cats et al. 2011). The abovementioned studies

provide behavioral evidence that reliable and timely arrival information can lead to actual as well as perceived waiting time savings.

In addition to experienced and perceived travel time savings, the introduction of public transport information systems improves passenger satisfaction and even increases ridership (Tang and Thakuriah 2012). According to a survey conducted by Caulfield and O'Mahony (2007), 92% of travelers regard RTI on the estimated time of arrival to be an important or a very important source of information. The dissemination of accurate and reliable information is therefore of great importance for the public transport industry.

The benefits of RTI may be hindered, fail to materialize or even become counterproductive in case of poor information quality. Gooze et al. (2013) found that RTI accuracy has significant impacts on ridership and satisfaction. Furthermore, 38% of public transport users reported an acceptable error of up to 3 minutes for RTI provision while additional 37% indicated a tolerance of up to 4-5 minutes. However, this tolerance may not be symmetric as passengers may be more sensitive towards underestimation of the remaining waiting time than overestimation because the former is associated with prolonging the expected waiting time. In contrast, passengers that coordinate their arrival based on RTI will be more sensitive towards arrivals that are earlier than the predicted arrival times since this may result with missed connections. Chow et al. (2014) suggested that inaccurate RTI could result with greater anxiety and dissatisfaction than the absence of information. They argue that information provision raises passenger expectations and hence increases the importance of service reliability. Following the same reasoning, public transport operators that constantly fail to deliver punctual service may decide to revoke timetables, believing that providing no information is better than providing poor information.

In addition to the dissemination through digital displays and travel journeys, online public transport predictions can support control interventions. Forecasted arrival times, and their implications on headways between successive buses and even passenger flows, are an integral component in control strategies designed to improve service regularity (Cats et al. 2012) and coordinate transfers (Dessouky et al. 2003). Thus, the efficiency of real-time control strategies depends on the quality of online predictions. From the operator's point of view, online predictions are also beneficial in terms of their potential to reduce the uncertainty associated with fleet and crew management. For example, forecasting downstream arrival times can be instrumental in monitoring the effects of delays on trip chaining and driver relief points.

Previous studies provide important insights on the potential value of RTI and computationally expensive prediction schemes. None of the previous studies examined the impact of alternative prediction schemes on potential travel time savings. The contribution of this study to the literature is twofold: (1) test the performance of two state-of-the-art online bus arrival prediction schemes which: incorporate instantaneous data on vehicle positions, are not computationally expensive, and applied in a real-world case study; (2) quantify the potential waiting time savings associated with each prediction scheme and evaluate its added-value when compared with: static information, state of the practice RTI generation, and improvements in bus service frequency and reliability. The following section presents a modelling framework for generating and evaluating real-time bus arrival information.

3. Methodology

The bus arrival prediction schemes formulated and evaluated in this study rely on the availability of static timetables and real-time data about public transport vehicle positions. Vehicle positioning data could be either transmitted in a time- or event-based fashion. The former implies that a vehicle reports its location within fixed intervals (e.g. every 30 seconds, the time interval varies considerably between different systems), whilst the latter implies that positioning data is transmitted every time that the vehicle visits a stop or a driver relief point. The prediction methods described in this study could be applied to both time- and event-based AVL transmission. For simplicity reasons, the notations assume event-based generation, however the formulation can be easily adopted to accommodate time-based vehicle probes. In addition, historical time-dependent passenger demand profiles are given as input to the third prediction scheme. Online availability of data concerning passenger demand (APC or ticket validation) can potentially further improve this prediction scheme. The RTI generator yields predictions concerning the remaining time until the arrival of the next bus for each line at each stop across the network.

3.1 Modeling framework

Vehicle trajectory can be represented as a vector of time stamps along a list of locations, typically stops. The trajectories of an ordered set of bus trips - denoted as K_l for bus line $l \in L$, where L is the set of bus lines in the network, during a certain time interval - can be represented as a matrix. This matrix is denoted as π^a where each cell, $\pi_{k,s}^a$, is the actual time that bus trip k arrived at stop $s \in S_l$, and S_l is the list of stops on line l . Similarly, π^d represents the corresponding departure times ($\pi_{k,s}^d$, bus trip k departs from stop $s \in S_l$). These matrices are dynamically updated throughout service operations and they remain partially empty due to ongoing trips at any time instance τ .

A corresponding matrix denoted π^t contains the timetable trajectories for K_l ($\pi_{k,s}^t$, bus trip k scheduled arrival at stop $s \in S_l$). This time-dependent timetable database is seasonally constructed and contains the planned arrival times at each stop along the route. Many bus services use a subset of stops - known as time point stops (TPS) - for regulating departure times according to the timetable. Drivers are instructed to regulate the departure time from these stops based on a certain service criterion. The most commonly used schedule-based holding control implies that drivers should not depart from TPS prior to the scheduled departure time in order to improve service punctuality (Cats et al. 2012). In particular, dispatching from the origin stop is typically regulated based on the timetable. TPS are indicated in the timetable as a subset of the recorded locations for line l ($\hat{S}_l \subseteq S_l$). The output of the prediction scheme generated at time τ is the corresponding matrix of predicted bus arrivals, $\pi^p(\tau)$. In summary, each line has matrices of observed arrivals, observed departures, scheduled arrivals and predicted arrivals, π^a , π^d , π^t and π^p , respectively. The dimensions of these matrices are $|K_l| \times |S_l|$.

Regardless of the specific prediction scheme, the following algorithmic building-block is common for all the schemes presented in this paper. For each stop, the bus trip for which the prediction is made, k^p , has to be identified. A pointer to the latest trip, that has visited stop s , $k^p - 1$, is first made. Then, it is assumed that the bus which is closest to stop s at the time instance at which the prediction is made, τ , is expected to arrive first at this stop. A backward search is performed in order to identify the latest trip that had visited an upstream location m but has not visited stop s yet:

Algorithm 1: Identify reference trip and stop for arrival at stop s after prediction time τ

Initialization:

$$m = 0 ; k^p - 1 := \arg \max_{K_l} \{ \pi_{k,s}^a : \pi_{k,s}^a < \tau \}; k^p := \arg \max_{K_l} \{ \pi_{k,s}^s : \pi_{k,s}^s < \pi_{k^p-1,s}^s \}$$

Search:

For $\acute{s} = s - 1$ to $\acute{s} = 1$

If $\max_{K_l} \{ \pi_{k,\acute{s}}^a : \pi_{k,\acute{s}}^a < \tau \} > \pi_{k^p-1,s}^a$ then $k^p := \arg \max_{K_l} \{ \pi_{k,\acute{s}}^a : \pi_{k,\acute{s}}^a < \tau \}$ and $m := \acute{s}$
stop

End

In case that no relevant bus arrival has been recorded for an upstream stop, the scheme reserves to the next scheduled trip and the origin stop. Note that the bus arrival prediction always refers to the approaching bus and therefore the reference bus might alter in case that overtaking occurs between successive predictions. All of the temporal components presented hereafter refer to time τ .

Following the identification of the reference trip/vehicle and its latest recorded position, $\pi_{k^p,m}^a$, the forecasted arrival time at stop s , $\pi_{k^p,s}^p$, is determined by the remaining time required for traversing the segment between m and stop s , $\tilde{t}_{k^p}^{m \rightarrow s}$.

$$\pi_{k^p,s}^p = \pi_{k^p,m}^a + \tilde{t}_{k^p}^{m \rightarrow s} \quad (1)$$

Three schemes for predicting $\tilde{t}_{k^p}^{m \rightarrow s}$ are formulated and evaluated in this study: a scheme that reflects the state-of-the-practice and two state-of-the-art schemes. The criteria for formulating the latter were that they: (a) rely on data that is commonly available in real-time; (b) do not require computationally expensive techniques for parameter estimation and calibration; (c) can be directly embedded into the existing scheme; (d) are feasible in real-time large-scale applications.

First, a method which is commonly used in practice and is based on the timetable is described. Second, a method that incorporates information on link-level travel times experienced by preceding buses is formulated. Third, the latter method is extended to a successive forward scheme which involves the prediction of both running times and dwell times. Further details on their theoretical background and technical notes on the implementation of the prediction schemes are available in Newell and Potts (1964) and Loutos (2013), respectively. The three prediction schemes vary with respect to the input required and the computation procedure employed for obtaining $\tilde{t}_{k^p}^{m \rightarrow s}$. Table 1 summarizes the three prediction schemes that are detailed in sections 3.2-3.4.

Table 1: Summary of alternative prediction schemes

Prediction scheme	Algorithms	Input			
		Timetable, π^t	Current vehicle position, $\pi_{k^p, m}^a$	Downstream vehicle positions, π^a	Historical demand profile, $\lambda^s(\tau)$, and dwell time coefficients, α and β
Scheduled travel time (STT): commonly deployed in practice	2 (+1)	√	√		
Real-time travel time (RTTT): link-based travel times with a weighting function	4 (+1+3)	√	√	√	
Iterative forward running and dwell time (IF_RDT): separate prediction of running times (weighting function) and dwell times	7 (+1+5+6)	√	√	√	√

3.2 Scheduled travel time method

This scheme is provided by one of the three leading global ITS providers and is thus used by many public agencies worldwide, including the public transport authority in Stockholm, Sweden (SLL). This scheme is used as a benchmark in this study as it reflects the current state-of-the-practice. It was formulated based on direct discussions with the information technology department of the system provider and the technical department of SLL. The system provider confirmed that the scheme evaluated in this paper is currently used by most of its customers but its identity cannot be disclosed due to commercial competition reasons. A detailed empirical evaluation of the performance of this commonly deployed scheme is available in Cats and Loutos (2013).

This simple prediction method assumes that the remaining travel time is equal to the scheduled travel time. This is a conservative estimation that does not involve any further real-time data, beyond what is required for obtaining k^p and m (Algorithm 1). The prognosis is based on the following assumptions: (a) the travel time between bus current location and any downstream location is equal to the scheduled travel time, and; (b) buses never leave a TPS prior to their scheduled time. Formally, these prediction rules could be formulated as follows:

$$\pi_{k,j}^d - \pi_{k,i}^d = \pi_{k,j}^t - \pi_{k,i}^t \quad \forall i, j \in S_l, i < j \quad (2)$$

$$\pi_{k,j}^d = \pi_{k,j}^t \quad \forall j \in \hat{S}_l \quad (3)$$

The combination of these rules implies a delay conservation assumption. In other words, buses are assumed to maintain their schedule deviation, either earliness or lateness – unless the bus runs early and a TPS exists between its current location, m , and the relevant downstream location. In this latter case, by holding until the scheduled departure time at the TPS, the bus will be able to correct its schedule deviation and therefore arrive on-time at stops downstream of the TPS. Consequently, a bus that runs behind schedule is expected to sustain its current delay at all downstream stops. This scheduled travel time (STT) method is performed as follows:

Algorithm 2: Predict arrival time at stop s based on scheduled travel time

Initialization:

Obtain m and k^p from *Algorithm 1*

Predict:

If $m = 0$ OR $(\pi_{k^p,m}^a < \pi_{k^p,m}^t \text{ AND } \exists m \leq i < s, i \in \hat{S}_l)$ then $\pi_{k^p,s}^p = \pi_{k^p,s}^t$

Otherwise,

$$\pi_{k^p,s}^p = \pi_{k^p,m}^a + \pi_{k^p,s}^t - \pi_{k^p,m}^t$$

The condition in Algorithm 2 implies that if the reference trip has not started yet or in case it runs early and there is an intermediate TPS between m and s , then the predicted arrival time is simply the scheduled time. In all other cases, the predicted arrival time is calculated based on the scheduled remaining travel time (i.e. the difference between scheduled times at latest stop m and the stop for which the prediction is made, s). This prediction scheme does not take advantage of the abundance of real-time data which could potentially provide a better indication of the prevailing traffic conditions.

3.3 Real-time travel time method

The STT method implicitly assumes that $\tilde{t}_{k^p}^{m \rightarrow s}$, the remaining travel time required for the reference vehicle in order to arrive at a downstream stop, is fixed for a given schedule period, whereas in practice it is subject to within-day as well as day-to-day variations. Timetable design normally accounts for some of these differences by constructing different timetables for different seasons, days of the week (e.g. Monday-Friday, Saturday and Sunday) and time intervals (e.g. morning peak, off-peak, afternoon-peak, evenings). However, more nuanced variations (e.g. transition periods, heterogeneity within periods) may be neglected. More importantly, the static timetable refers to aggregate historical conditions rather than current prevailing conditions. The latter are made available through processing recently transmitted vehicle position probes. The incorporation of such information could potentially yield a more accurate estimation of future travel times.

The predicted travel time between each pair of stops is computed based on a weighted average of travel times experienced by δ preceding trips ($k^p - 1, k^p - 2, \dots, k^p - \delta$) that have traversed the respective segment. Hence, the set of preceding trips that is used for travel time prediction always utilizes the most recent available data for a given segment and may vary across segments. While various weighting functions could be deployed, the weighting function employed in this study is based on the weighted method proposed by Yu et al. (2011). This weighting method assumes that more recent travel times carry more valuable information for predicting future travel times. Consequently, increasing the number of preceding trips, δ , is expected to result with improved predictions albeit with diminishing returns. Travel times of preceding buses are thus assigned with a

weight inversed to the elapsed time from the upstream stop in order to account for the recency factor as formulated in Algorithm 3.

Algorithm 3: Compute remaining travel time between arrival at stops s_1 and s_2

// Identify reference trips for previous arrivals

For $j = 1$ to δ

$$k^p - j := \arg \max_{K_l^a} \{ \pi_{k,s_2}^a : \pi_{k,s_2}^a < \pi_{k^p-j+1,s_2}^a \}$$

End

// assign weights to reference trips

For $j = 1$ to δ

$$\Gamma_j^{s_2} = \frac{1 / (\pi_{k^p,s_1}^a - \pi_{k^p-j,s_1}^a)}{\sum_{j=1}^{\delta} 1 / (\pi_{k^p,s_1}^a - \pi_{k^p-j,s_1}^a)}$$

End

$$\tilde{t}_{k^p}^{s_1 \rightarrow s_2} = \sum_{j=1}^{\delta} \Gamma_j^{s_2} (\pi_{k^p-j,s_2}^a - \pi_{k^p-j,s_1}^a)$$

A prediction method which estimates the remaining travel time based on preceding buses and incorporates schedule regulation at TPS was formulated. The real-time travel time (RTTT) method is then applied as follows:

Algorithm 4: Predict arrival time at stop s based on recent vehicle travel times

Initialization:

Obtain m and k^p from Algorithm 1

// identify intermediate TPS:

$$\{ \hat{s} \in \hat{S}_l^{m \rightarrow s} : \hat{s} \in \hat{S}_l, m < \hat{s} < s \}$$

Predict:

If $\hat{S}_l^{m \rightarrow s} = \emptyset$ then obtain $\tilde{t}_{k^p}^{m \rightarrow s}$ from Algorithm 3

Otherwise,

// predict successively for intermediate TPS

For $s_i = \{ \hat{s}_1, \dots, \hat{s}_{|\hat{S}_l^{m \rightarrow s}|-1} \} \in \hat{S}_l^{m \rightarrow s}$

Obtain $\tilde{t}_{k^p}^{s_i \rightarrow s_{i+1}}$ from Algorithm 3

// find expected arrival time at intermediate TPS

$$\pi_{k^p,s_{i+1}}^p = \max(\pi_{k^p,n}^p + \tilde{t}_{k^p}^{s_i \rightarrow s_{i+1}}, \pi_{k^p,s_{i+1}}^t)$$

End

$$\pi_{k^p,s}^p = \pi_{k^p,\hat{s}_{|\hat{S}_l^{m \rightarrow s}|-1}}^p + \tilde{t}_{k^p}^{\hat{s}_{|\hat{S}_l^{m \rightarrow s}|-1} \rightarrow s}$$

Unlike the scheduled travel time method, the inclusion of downstream travel time data in combination with schedule-based control implies an iterative prediction procedure. The prediction of the remaining travel time, $\tilde{t}_{k^p}^{m \rightarrow s}$, is therefore obtained by performing successive predictions for

intermediate TPS located between m and s . In case of an intermediate TPS, the schedule-based holding logic which underlies *Algorithm 2* (STT) is iteratively applied.

Similarly to the STT method, the input for the RTTT method is composed of a time-dependent timetable and real-time vehicle position probes. However, its implementation requires processing a larger number of vehicle positions and a successive prognosis rather than a single-step prediction. Note that unlike STT, in case the reference trip has not started yet ($m = 0$) the RTTT method will still result with a prediction which incorporated dynamic information. In the lack of any preceding trip for a certain segment (e.g. early morning trips), the scheme reserves to the timetable.

3.4 Iterative forward running and dwell time method

The aforementioned RTTT method proposes a way to utilize RTI from previous trajectories by referring to segment travel times. These travel times are composed of running times between stops and dwell times at stops. While running times of preceding buses are arguably indicative of downstream traffic conditions, this is not expected to hold true for dwell times. Passenger flows are the most important determinants of dwell time. In particular, the number of boarding passengers typically dominates the dwell time function. Furthermore, in the context of high-frequency lines passengers are assumed to arrive randomly at stops and hence their arrival process can be expressed as a Poisson process (e.g. Bowman and Turnquist 1981). The number of waiting passengers is therefore a random variable that can be approximated by a negative exponential distribution. The inter-arrival parameter corresponds to the headway between the bus concerned and the previous bus. The positive feedback loop between headway and dwell time is the underlying source for the bunching phenomenon (Cats et al. 2012).

In order to account for the potential prediction power of headways on downstream bus trajectory, the RTTT method is further elaborated by decoupling running times and dwell times and predicting them separately and successively. This method involves an iterative forward running and dwell time (IF_RDT) prediction, where the running times are predicted in a similar fashion to how travel times are predicted in the RTTT method. Algorithm 3 is modified in order to refer to running times only based on elapsed time between departure and arrival times at downstream stops, as follows:

Algorithm 5: Compute remaining travel time between departure from stop s_1 and arrival at stop s_2

// identify reference trips for previous arrivals

For $j = 1$ to δ

$$k^p - j := \arg \max_{K_l^a} \{ \pi_{k,s_2}^a : \pi_{k,s_2}^a < \pi_{k^p-j+1,s_2}^a \}$$

End

// assign weights to reference trips

For $j = 1$ to δ

$$\Gamma_j^{s_2} = \frac{1 / (\pi_{k^p,s_1}^d - \pi_{k^p-j,s_1}^d)}{\sum_{j=1}^{\delta} 1 / (\pi_{k^p,s_1}^d - \pi_{k^p-j,s_1}^d)}$$

End

$$\tilde{t}_{k^p}^{s_1 \rightarrow s_2} = \sum_{j=1}^{\delta} \Gamma_j^{s_2} (\pi_{k^p-j,s_2}^a - \pi_{k^p-j,s_1}^d)$$

As its name implies, this prediction scheme is composed of predicting successively and alternately running and dwell times. While algorithm 5 yields a prediction for the running time between departure and arrival, algorithm 6 predicts the dwell time at stop s between arrival and departure of trip k^p . The computation of the dwell time consists of embedding the predicted number of boarding passengers, $\tilde{q}_{k^p}^s$, into the dwell time function.

Ideally, real-time passenger demand information concerning $\tilde{q}_{k^p}^s$ is made available. This might be the case where the number of downstream passengers can be approximated based on gate counts. However, most systems do not have RTI concerning passenger demand and will therefore rely on historical time-dependent demand patterns. Dwell time predictions can be then made based on the relation between passenger arrival process and headways. First, the headway prior to the arrival of trip k^p at stop s , $h_{k^p,s}^p$, is derived from the predicted departure times of subsequent trips. Second, the number of boarding passengers is predicted based on the hourly Poisson arrival rate, $\lambda^s(\tau)$, and the time that elapsed since the previous bus arrival. Passenger arrival rates for various time intervals can be estimated based on aggregate historical passenger counts. Third, $\tilde{q}_{k^p}^s$ is embedded into the dwell time function, where α is the non-service lost time (e.g. door opening or closing time) and β is the boarding time per passenger. The values of these parameters should be estimated for the specific line or system as they depend on the number of doors, payment method, vehicle and stop characteristics. Dwell time function coefficients, α and β , are commonly available from the local public transport agency or operators.

Algorithm 6: Predict dwell time at stop s

```
// Predict headway
 $h_{k^p,s}^p = \pi_{k^p,s}^p - \pi_{k^p-1,s}^p$ 
// estimate number of boarding
 $\tilde{q}_{k^p}^s = \frac{\lambda^s(\tau)}{60} h_{k^p,s}^p$ 
// predict dwell time
 $\tilde{d}_{k^p}^s = \alpha + \beta \cdot \tilde{q}_{k^p}^s$ 
```

The predicted dwell time and running time are integrated into the prediction scheme as presented schematically in Figure 1 and formally in algorithm 7. A sequence of inter-dependent predictions are performed progressively along the segment between vehicle's latest recorded location, m , and the relevant downstream location, s . The subsequent predictions involve the prognosis of headways, dwell times, departure times, running times and arrival times at downstream stops. In addition to travel times experienced by previous vehicles, passenger time-dependent demand profiles per stop are given as input to this scheme.

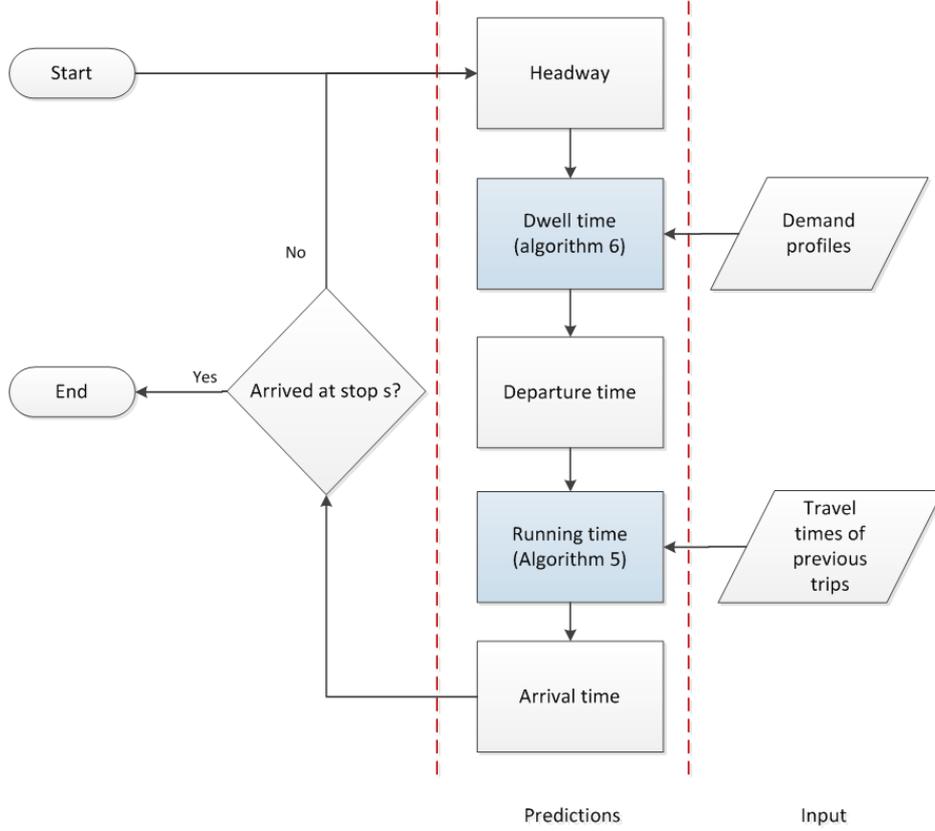


Figure 1: Iterative forward running and dwell time prediction scheme

The iterative prediction process shown in Figure 1 is formulated in algorithm 7. In case the next expected trip has not started yet then the scheduled departure time from the origin terminal is used as the initial time reference. The step-wise repetitive process updates $\pi_{k^p,s}^p$ by accumulating the travel time along the segment. The latter process depends on whether trip k^p has already departed from stop m or not. Similarly to the previous schemes, the prediction takes into consideration the schedule-based control at intermediate TPS.

Algorithm 7: Predict arrival time at stop s based on recent vehicle running times and dwell times

Initialization:

Obtain m and k^p from *Algorithm 1*, $\pi_{k^p,s}^p = \pi_{k^p,m}^a$

If $m = 0$ then $\pi_{k^p,s}^p(\tau) = \pi_{k^p,m}^t$

Predict:

For $s_1 = m$ to $s - 1$

 If $\pi_{k^p,s}^d \neq 0$ then

 Obtain $\tilde{d}_{k^p}^{s_1}$ from *Algorithm 6*

$\pi_{k^p,s}^p := \pi_{k^p,s}^p + \tilde{d}_{k^p}^{s_1}$

 If $s_1 \in \hat{S}_l$ AND $\pi_{k^p,s}^p < \pi_{k^p,s_1}^t$ then $\pi_{k^p,s}^p := \pi_{k^p,s_1}^t$

 End

 Obtain $\tilde{t}_{k^p}^{s_1 \rightarrow s_1+1}$ from *Algorithm 5*

$\pi_{k^p,s}^p := \pi_{k^p,s}^p + \tilde{t}_{k^p}^{s_1 \rightarrow s_1+1}$

End

3.5 Performance metrics

RTI performance is assessed by a series of metrics that are calculated ex-post and consider both passengers' and operators' perspectives. While the former is concerned with the sequence of arrivals of any vehicle on a certain line, the latter is interested in arrival times of specific vehicles. The prediction error for the arrival of trip k at stop s is therefore assessed by comparing the prediction generated at time τ against the corresponding actual arrival time of the same trip:

$$e_{k,s}^p(\tau) = \pi_{k,s}^a - \pi_{k,s}^p(\tau) \quad (4)$$

From passengers' perspective, however, no importance is attached to the specific trip identity, and the accuracy is determined by the difference between the provisioned RTI and the next arrival of line l at stop s , calculated as follows:

$$e_{l,s}^p(\tau) = \pi_{k^a,s}^a(\tau) - \pi_{l,s}^p(\tau) \quad (5)$$

Where k^a is the first trip to arrive at the stop, defined as $k^a = \arg \min_{K_l} \{\pi_{k,s}^a : \pi_{k,s}^a > \tau\}$.

This could be interpreted as the difference between the predicted and experienced waiting times for a passenger that arrived at stop s at time τ . Note that k^a might differ from k^p when an overtaking occurs between m and s .

The prediction error measures enable to identify the difference between predicted and observed arrival times. The prediction errors yielded by an unbiased prediction scheme will have a average value of zero. Moreover, the variability of prediction errors has to be minimized in order to obtain an accurate and reliable prediction scheme.

The performance of static information concerning arrivals is used as a benchmark for assessing the added-value yielded from RTI provision. Static information accuracy from operators' and passengers' perspectives - $e_{k,s}^t$ and $e_{l,s}^t(\tau)$, respectively - was formulated similarly by substituting the RTI prediction with the corresponding timetable term, as follows:

$$e_{k,s}^t = \pi_{k,s}^a - \pi_{k,s}^t \quad (6)$$

$$e_{l,s}^t(\tau) = \pi_{k^a,s}^a(\tau) - \pi_{k^t,s}^t(\tau) \quad (7)$$

Where k^t is the first trip scheduled to arrive at the stop, defined as $k^t = \arg \min_{K_l} \{\pi_{k,s}^t : \pi_{k,s}^t > \tau\}$.

Note that k^a might differ from k^t in case the first arriving bus was scheduled to arrive earlier than the RTI generation time (i.e. passenger arrival time at the stop).

Furthermore, the extent to which timetables and RTI are effective in assisting passengers to shift their expectations closer to the actual waiting time is assessed. The actual waiting time of a passenger arriving at stop s at time τ with the intention to board line l is:

$$w_{l,s}^a(\tau) = \pi_{k^a,s}^a - \tau \quad (8)$$

While the expected waiting time implied by the timetable and RTI are, respectively:

$$w_{l,s}^t(\tau) = \pi_{k^t,s}^t(\tau) - \tau \quad (9)$$

$$w_{l,s}^p(\tau) = \pi_{l,s}^p(\tau) - \tau \quad (10)$$

The mean absolute error provides an aggregate measure of performance which could be compared across lines, systems, time periods and prediction scheme. It is calculated as follows:

$$MAE^t = \overline{|w_{l,s}^a - w_{l,s}^t|} = \frac{\sum_{l,s,\tau} |e_{l,s}^t(\tau)|}{\sum_{l \in L} |S_l| \cdot |\tau|} \quad (11)$$

$$MAE^p = \overline{|w_{l,s}^a - w_{l,s}^p|} = \frac{\sum_{l,s,\tau} |e_{l,s}^p(\tau)|}{\sum_{l \in L} |S_l| \cdot |\tau|} \quad (12)$$

Where $|S_l|$ and $|\tau|$ denote the number of stops on line l and the number of RTI generation instances considered in the analysis (e.g. analysing half a minute generation-intervals for the peak hour will result with 120 instances), respectively. This measure corresponds to the average time difference obtained by a certain information source when compared with the observed bus arrival times.

The abovementioned performance metrics enable the overall evaluation of an online prediction in terms of its accuracy and reliability. In addition, the performance of the prediction scheme should be robust with respect to various operational conditions such as traffic conditions, demand levels and timetables. Hence, the performance metrics are analyzed spatially (along the line) and temporally (different days of the week and times of the day) to evaluate the extent to which their performance fluctuates under different circumstances.

4. Application

The case study network and the respective vehicle positioning and passenger counts datasets are presented in the following sections (4.1-4.3). Thereafter, Section 4.4 provides details on the implementation of the prediction schemes which yielded RTI that is subject to detailed analysis in Section 5.

4.1 Network description

The trunk bus network of Stockholm's inner-city was selected as the case study network. This network was selected due to the importance of RTI provision for frequency-based services. The case study bus network serves high passenger demand and is subject to significant variations in travel conditions (Cats et al. 2012). In recent years, digital real-time arrival information panels were installed at all stops along the case study routes. Moreover, a significant share of public transport users rely on journey planner smartphone apps, suggesting that information reliability is an important determinant of passenger waiting times in the case study area.

The system consists of four bus lines which constitute the backbone of Stockholm's inner-city bus network (Figure 2). These lines are characterized by high frequency (planned headway of 4-7 minutes between 7:00 and 19:00), articulated vehicles, designated lanes on the main arterial streets and traffic signal priority. The case study network includes more than 200 stops along more than 80 route-km. Each route has two to four TPS located at key public transport transfer locations. These

lines account for 60% of the total ridership in this area with approximately 120,000 boarding passengers per day between 7:00-19:00.

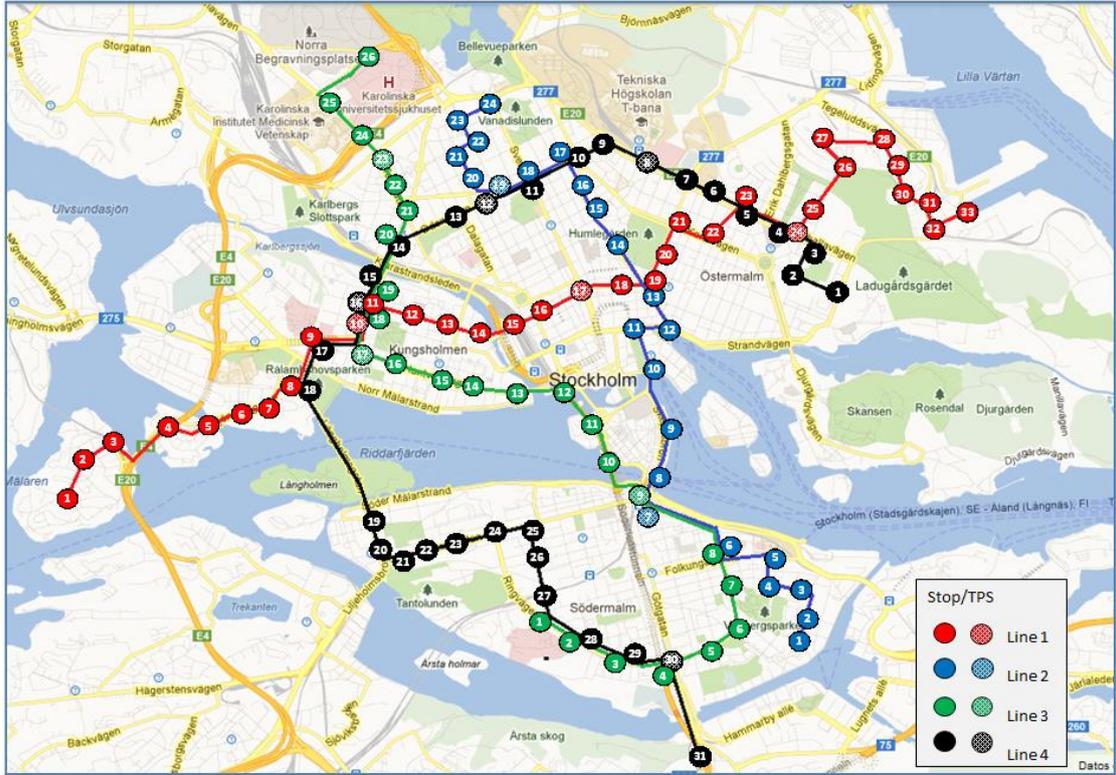


Figure 2: Stockholm's inner-city trunk lines routes

4.2 Vehicle positioning data

The entire bus fleet in Stockholm is equipped with an AVL system. All buses record and communicate their location every 15 seconds. The system is used for several purposes including radio communication, real-time monitoring and control of vehicles, fleet management strategies and the generation of real-time passenger information. The AVL database used in this study consists of a vehicle position record for a bus trip visit at each bus stop. This event-based database is constructed by computing the arrival and departure time of each vehicle trip from each stop along its route. Each stop visit record contains information regarding trip ID, stop ID, actual and scheduled arrival and departure times and the respective scheduled time.

The performance of the RTI generation method was analysed based on detailed and comprehensive AVL data. These data were provided by SLL, the Transport Administration of Stockholm County Council. The study period consists of records from 15/11/2011-15/12/2011 and 9/1/2012-19/1/2012 in order to exclude the seasonal holidays. All records for trips undertaken between 07:00-19:00 were considered for further analysis resulting with a dataset of more than one million records. A statistical analysis of the AVL data confirms that bus running times and dwell times vary considerably for the case study trunk lines. The coefficient of variation of running times is 0.25 and on some links it is as high as 0.95 during the analysis period. The coefficient of variation of dwell times is 0.6.

4.3 Passengers demand profiles

An additional dataset of APC records is used in this study. APC equipment is installed on approximately 15% of the trunk bus fleet in Stockholm and it is only available offline. These limitations hinder the possibility to apply a prediction scheme that incorporates real-time passenger counts data. Nevertheless, the APC dataset facilitates the analysis of passenger time-dependent demand - number of boarding, alighting and on-board passengers - along the line.

Load profiles were constructed in this study based on APC data collected between 1/9/2011-1/12/2011. Since only a sample of the case study fleet is equipped with APC devices, passenger counts were aggregated by the authority into four time of day periods: AM peak (06:00-09:00), off-peak (09:00-15:00), PM peak (15:00-18:00), and evening (18:00-21:00). Demand aggregation yields robust estimates per stop and route while discarding within-period variations that can have large impacts on bus dwell times. The dwell time function parameters – constant delay (α) and boarding time per passenger (β) – were specified based on values estimated by the regional authority based on local data, 3 and 2 seconds, respectively.

4.4 Implementation Details

The three prediction schemes described in Section 3 were applied to the case study network. The STT method involves algorithms 1 and 2, the RTTT method requires the implementation of algorithms 1,3 and 4, and IF_RDT is based on the combination of algorithms 1,5,6 and 7 (Table 1).

Vehicle positioning data and the corresponding timetable were first transformed into a matrix format resulting with π^a , π^d and π^t . The AVL data were sorted based on the scheduled trip departure time. Note that some lines consist of several trip patterns implying that the data contains also partial trips which traverse only a subset of the stops along the respective line. The prediction schemes address it explicitly by ensuring that the arrival prediction refers to a trip that is destined to serve the respective downstream stop.

The generation of RTI was implemented to yield the predicted time to arrival of each bus line at each stop. The output of the prediction scheme is equivalent to a snapshot of the countdown RTI provisioned at every stop across the study network with one minute interval throughout the study period. The RTI generator results with the predicted bus arrival time at each stop and line combination, $\pi_{i,s}^p(\tau)$, for every minute so that $\tau = (7:00, 7:01, \dots, 19:00)$.

The prediction scheme was implemented in MATLAB. The RTI generation performed at time τ follows then the steps outlined in the methodology and thus uses only bus positioning data collected prior to time τ . Note that while the input data – vehicle positioning, timetables and demand profiles – is attached to a specific trip and its availability follows an event-based manner, predictions are generated in a time-based fashion and refer to stops in order to reflect passenger perspective. As explained in Section 3, prediction methods' implementation addresses explicitly overtaking. The reference trip, k^p , may therefore refer to different trips, when generating predictions at stop s on successive time instances and with no intermediate vehicle arrivals.

The output produced by the implemented RTI generator enables the computation of the performance metrics defined in Section 3.5. The performance of the RTI generator is based on a cross-network comparison of the prediction provisioned by the RTI system and the corresponding

realized arrival times. This implies the calculation of the passengers' and operators' prediction error metrics across the network with one minute sampling. The RTTT and IF_RDT methods were implemented with $\delta = \{1,2,3,4,5,10\}$ in order to test the performance of the prediction scheme when utilizing different number of preceding buses.

The execution time of RTI generation algorithms concerning the next arrival from each line to all stops required less than a second for the case study network. A single full-network RTI generation (e.g. at time τ) was performed in 0.17 second in the case of STT, whereas RTTT took 0.30 seconds. The significantly greater complexity of IF_RDT results with a considerably longer computational time of 0.81 second. However, all of these execution times make their real-time implementation feasible. Note that all of the computations were executed on a personal computer with a CPU of 2.1 GHz, RAM of 8 GB, where the RTI generation code can be further optimized.

The implementation facilitates the temporal and spatial evaluation of the RTI generator. 'SLL minute', named after the public transport agency, is notoriously known and used by the public and popular media in Stockholm as a particularly 'long' minute because the actual waiting time exceeds the provisioned RTI. The design of this study enables to examine the extent to which the coined term is empirically justified for the current system.

5. Results

The first part of this section presents the prediction error accuracy and reliability. Prediction error is investigated from global performance and from prediction time horizon perspectives. The added-value of RTI was measured for passengers' waiting times and for operators' capability to monitor their fleet. The analysis continues with temporal and spatial analysis of the prediction error. The results section concludes with comparing the impacts of RTI with alternative measures to reduce passenger waiting time.

5.1 Prediction error accuracy and reliability

5.1.1 Global performance

The performance of RTI generation schemes was first analyzed by investigating the prediction accuracy – the difference between actual and predicted time to arrival – across the network for the entire study period. A sensitivity analysis with respect to the number of preceding trips embedded into the prediction scheme, δ , suggested that the marginal benefit in terms of prediction accuracy and reliability from increasing the number of buses diminishes sharply for $\delta > 5$, as was hypothesised in section 3.3. The number of preceding trips was therefore set to five in all subsequent results.

Figure 3 presents the distribution of prediction error, $e_{l,s}^p(\tau)$ (Eq. 5), for the three prediction schemes – STT, RTTT and IF_RDT. The null hypothesis that the prediction error follows a normal distribution was rejected for all prediction schemes at the 95% confidence level by the Kolmogorov-Smirnov, Jarque-Bera and Lilliefors tests. The t-location-scale and Logistic distributions have the best goodness-of-fit, followed by the normal distribution. In all cases, the distribution of prediction error is characterized by a high kurtosis, implying that prediction errors are more prone to outliers than a normal distribution will suggest, as reflected by the heavy tail in Figure 3. Furthermore, the prediction error of RTTT and IF_RDT follow a narrower distribution (e.g. the best fitted t-location-scale distribution has shape parameters of 38 and 42 seconds, respectively, compared with 51

seconds in the case of STT) implying more reliable information with fewer large discrepancies from the actual bus arrival times.

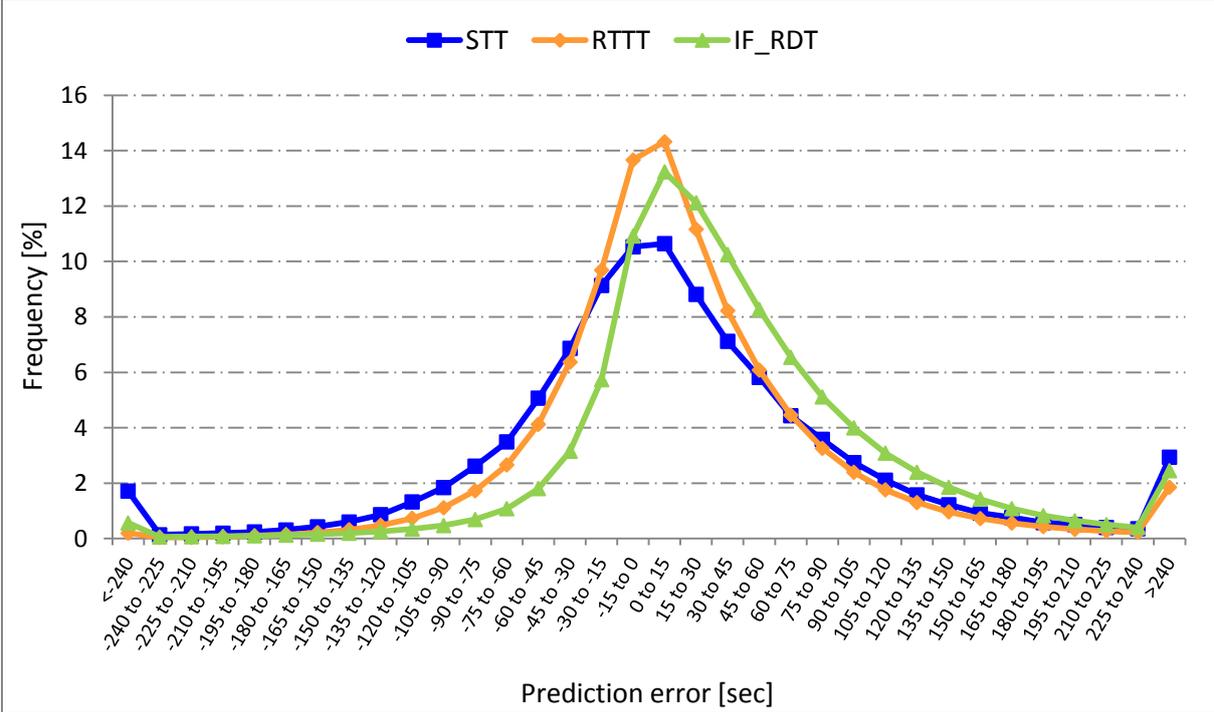


Figure 3: Real-time information prediction error

Table 2 summarizes descriptive statistics of the prediction error distributions. All prediction methods resulted with a systematic underestimation of the remaining time to arrival. The underestimation varies from a slight bias of 15 or 20 seconds for STT and RTTT, respectively, to a substantial systematic underestimation of 41 seconds in the case of IF_RDT.

The average prediction error indicates whether there is a consistent bias inherent to the prediction scheme. However, it is not a very informative indicator of scheme performance since it can be obtained by a highly dispersed but symmetric scheme because under- and over-predictions can be canceled out. The accuracy and reliability of RTI should therefore be assessed by the overall distribution of values. The standard deviation is highest for the currently deployed scheme. Moreover, even though STT results with the smallest systematic bias, it performs worst in terms of the share of high prediction errors. The share of predictions that result with an error exceeding one minute decreases from 36% for STT to 34% for IF_RDT and drops to 25% for RTTT. The latter is particularly effective in reducing the share of very large errors of more than two and four minutes as is clearly visible in Figure 3. This indicates that RTTT scheme improves the prediction success rate for various tolerance ranges. For both STT and RTTT, two thirds of the errors are attributed to underestimation, while this share increases to 80% for IF_RDT. It should be noted that positive and negative prediction errors may not be evaluated equally by passengers. On one hand, underestimation may be more noticeable as it induces longer waiting time than expected and results with dissatisfaction and stress. On the other hand, an overestimation may result with passengers missing a bus because they rely on the provisioned time (e.g. coordinate their arrival or perform

activities in proximity to the stop) and hence induce substantially longer waiting time and dissatisfaction.

Table 2: Descriptive statistics of alternative prediction schemes

	STT	RTTT	IF_RDT
Average [sec]	15	20	41
Standard deviation [sec]	120	93	111
$\geq \pm 1$ min [%]	36	25	34
$\geq \pm 2$ min [%]	14	8	13
$\geq \pm 4$ min [%]	5	2	3
Success rate for short-term predictions [%]	73%	83%	80%
Success rate for long-term predictions [%]	84%	91%	85%

The analysis of prediction errors facilitates the quantitative investigation of the ‘SLL minute’ and assessing it empirically. The prediction error of STT, the currently deployed scheme, was assessed by accounting for the respective waiting time. This analysis obtained that the average excessive waiting time per projected minute is 6.2% for the case study network. In absolute terms, the average ‘SLL minute’ lasts in fact 63.7 seconds. If the proposed schemes were applied the ‘SLL minute’ would have even been prolonged by 1 to 6 seconds. However, it is postulated that the reputation of ‘SLL minute’ arises from experiences of very large prediction errors which occur more often with STT (see Section 5.2.1).

5.1.2 Impact of prediction horizon

The performance of alternative prediction schemes may also depend on the prediction horizon – the elapsed time between the time that the prediction is generated (τ) and the time that the respective event actually takes place ($\pi_{k^p,s}^a$). Chow et al. (2014) mention that there is an industry-standard based on the share of predictions that are accurate within a pre-defined tolerance, depending on the prediction horizon. They refer to a ± 1 min tolerance for predictions shorter than 5min and ± 3 min for longer prediction horizons. Chow et al. report that more than 90% of predictions are considered accurate based on this criterion for the Boston subway system. The currently deployed system, STT, yields a 73% and 84% success rate for short- and long-term predictions, respectively (Table 2). This falls short of the industry-standard reported above. Notwithstanding, it is expected that RTI concerning bus arrival times would be less accurate than the corresponding information for rail-bound traffic. The success rates increase significantly for short-term predictions under both RTTT and IF-RDT and in the case of the former also for long-term predictions, when compared with the current scheme.

As reflected in the analysis of success rates, the relative prediction error depends greatly on the remaining waiting time. The relation between prediction error and prediction horizon – the elapsed

time between RTI provision and the next bus arrival time - is presented in Figure 4. The prediction horizon is equivalent to the actual waiting time, $w_{t,s}^a(\tau)$ (Eq. 8). The bars in Figure 4 show the relative difference between the actual waiting time and the waiting time projected by the different RTI schemes (Eq. 10) as a function of the former. For very short waiting times the STT scheme overestimates the remaining waiting time followed by an increasing underestimation for waiting times longer than 2 minutes. In particular, the underestimation reaches almost 10% for waiting times longer than 8 minutes. This implies that a RTI projection of 9 minutes would on average result with a waiting time of almost 10 minutes. In contrast, RTTT and IF_RDT result with no mean bias for short-term predictions and then increasingly underestimate the remaining waiting time to a greater extent than the STT method does. As discussed above, IF_RDT yields an underestimation of more than 10% for waiting times longer than half a minute.

A measure of RTI variability, the standard deviation of the prediction error, is also presented in Figure 4 as a function of the remaining waiting time. The curve indicates that the RTI prognosis is less reliable for longer prediction horizon for all prediction schemes. This is expected as longer prediction horizons undermine the possibility to estimate the prevailing traffic conditions. Since the standard deviation of RTI prediction error increases linearly with the waiting time, the reliability does not increase in relative terms (e.g. coefficient of variation). Notwithstanding, the RTTT and IF_RDT prediction methods are found to be more reliable than the STT method throughout the entire range of remaining waiting times. In other words, the error produced by the RTTT and IF_RDT methods is much more predictable. This trend is observed both in absolute (standard deviation) and relative (coefficient of variation) measures of variability. Next, we will examine whether these methods also yield better predictions from passengers' point of view.

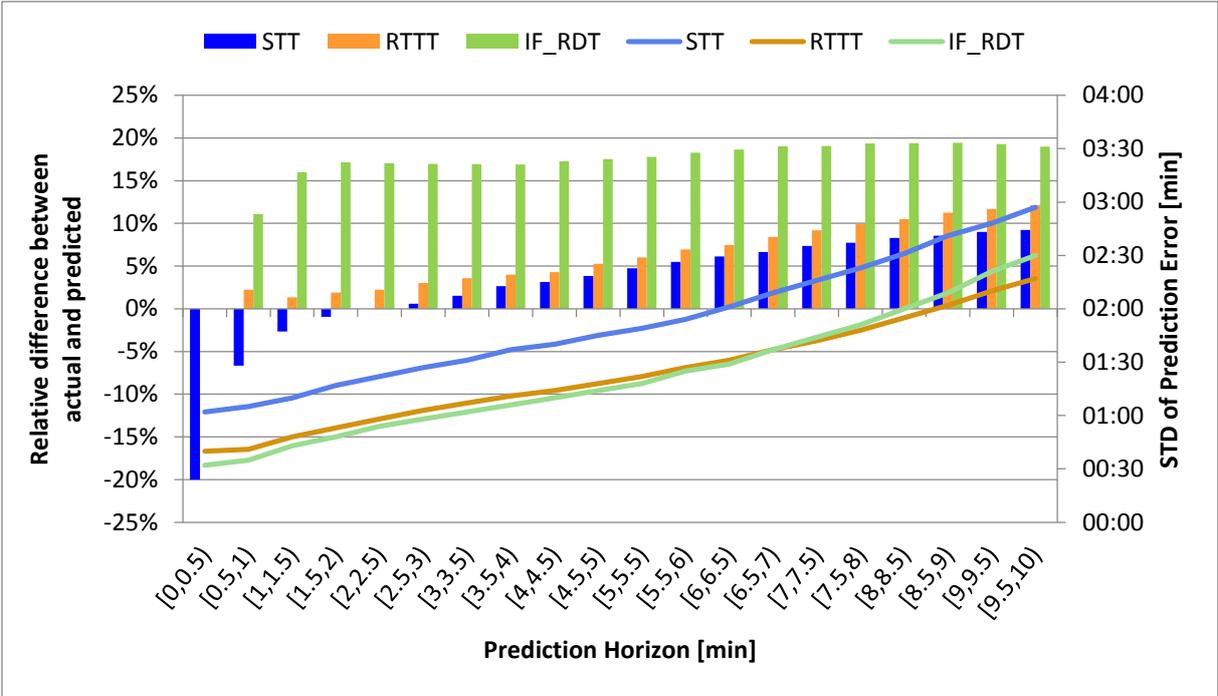


Figure 4: Real-time information accuracy and reliability as function of the remaining time until the next bus arrival

5.2 Added-value of real-time information

5.2.1 Passengers' waiting times

The added-value of RTI provision could be ultimately assessed against the benchmark of expectations that could be derived from the static timetable. Alternative prediction methods may vary with respect to how well they correspond to the actual waiting times and hence their potential to shift passengers' expectations closer to the realized experience. A discrepancy between user expectations and experience is the prime determinant of dissatisfaction (TCRP 2003). Moreover, more accurate RTI can result with lower perceived waiting time and even support better travel decisions that will result with actual waiting time savings. In order to investigate the added-value of RTI, the distributions of expected waiting time based on the timetable and RTI (Eq. 9 and 10, respectively) for the different prediction schemes were constructed. Figure 5 illustrates the aforementioned distributions along with the distribution of the actual time to arrival (Eq. 8). Assuming that passengers arrive randomly and homogeneously at stops, the average actual waiting time for the study network during the whole study period equals 4 minutes and 12 seconds, where 80% of the passengers wait less than 5 minutes. Since the RTI generation procedure followed a uniform temporal distribution, the number of observations is linearly proportional to the headway. It is worthwhile to note that the average actual waiting time is 57% longer than the value that would have been obtained from a perfectly regular bus arrival.

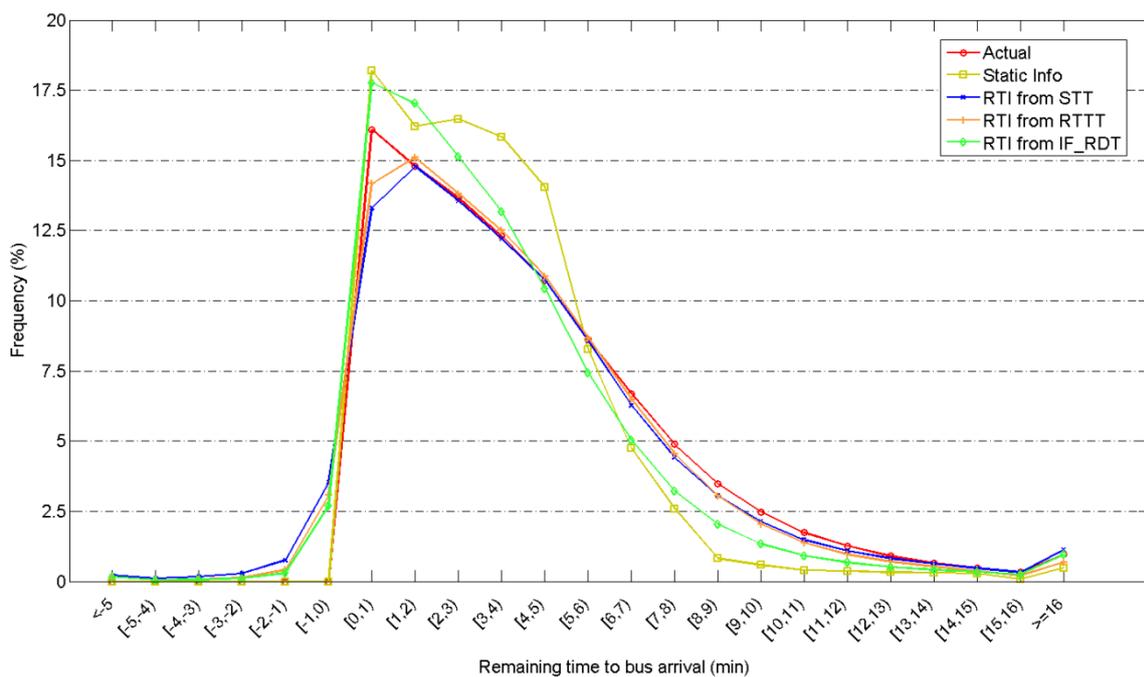


Figure 5: Information based on static timetable and alternative prediction schemes compared with the actual remaining time to arrival

Waiting time expectations derived from the timetable result in a considerable underestimation of waiting times. The comparison of static information and actual arrivals reveals an overestimation of the likelihood of waiting times shorter than 5 minutes. This is explained by the fact that the expected waiting time based on the timetable can be realized only in cases where the service is perfectly punctual.

Compared with static information, the distributions of expected waiting time based on RTI produced by the STT and RTTT methods follow closely the distribution of actual waiting time. Overall, the mean absolute error of the timetable is equal to $MAE_t = 146 \text{ sec}$ (Eq. 11). This discrepancy is reduced by more than a half to 68 and 63 seconds by the STT and IF_RDT schemes, respectively (Eq. 12). RTTT contributes to its further reduction to 51 seconds and it performs particularly well for waiting times shorter than 1 minute or within the range [5,8] minutes. RTI hence enables passengers to shift their expectations considerably closer to their experienced waiting time by 53% to 65% for STT and RTTT, respectively. This reduction refers to the objective improvement in projected waiting times. Interestingly, the reduction attained by the STT is of the same magnitude to the subjective improvement reported by Dziekan and Kottenhoff (2007) for the underground system in Stockholm. These savings can ultimately translate into welfare gains. If travelers are fully able to take advantage of the improved RTI accuracy by reducing accordingly their waiting times, then waiting time savings associated with the RTTT scheme amount to approximately 44,000 euro per day for the case study network, based on the value of time recommended by the Stockholm County council.

Note that a small share of the RTI predictions yields negative values. Negative values correspond to cases where stop's sign displays that the next bus should arrive 'Now' although the approaching bus has not reached the stop yet. The share of such cases is 5% under the current scheme and decreases to 3.9% and 3.4% for RTTT and IF_RDT, respectively. It is postulated that this phenomenon contributes significantly to the 'SLL minute' reputation since this underestimation is directly observed by passengers and can be a source of frustration.

5.2.2 Operators' vehicle monitoring

The added-value of RTI is of course not limited to passengers. Accurate and reliable predictions of bus arrival times are also beneficial for operators who can rely on this information for real-time fleet management considerations. Operators are primarily interested in predicting whether vehicle trajectories adhere to the planned schedule. The latter helps to assess the need to undertake mitigation strategies such as using reserve buses, deadheading, short-turning, holding for transfer coordination etc. The added-value of RTI can thus be translated into more efficient and effective resource allocation.

The implementation of the prediction schemes generates RTI at the stop-level rather than at the individual vehicle-level. The vehicle-level analysis was therefore conducted based on aggregate performance. Figure 6 presents the distribution of deviations between RTI prediction schemes and the timetable (Eq. 5 and 7) as well as the actual schedule adherence (Eq. 6). The distribution of the latter carries a long tail which corresponds to long delays. It is evident that STT greatly overestimate the share of on-time arrivals. This error stems from the underlying assumption that all drivers will comply with the schedule-based control strategy in case of early arrivals. In contrast, the deviations between RTTT and IF_RDT predictions and the timetable closely follow the empirical distribution of schedule deviations. The analysis suggests that the added-value of RTI, and in particular methods that include real-time data, are greater for operators than for travelers, when compared with the timetable. This difference stems from the fact that passengers are indifferent towards whether a certain bus is the earlier bus running late or the later bus arriving early. Whereas, bus operators are not only interested in the inter-arrival distribution but also in the order of occurrences and how well do they match the planned vehicle scheduling.

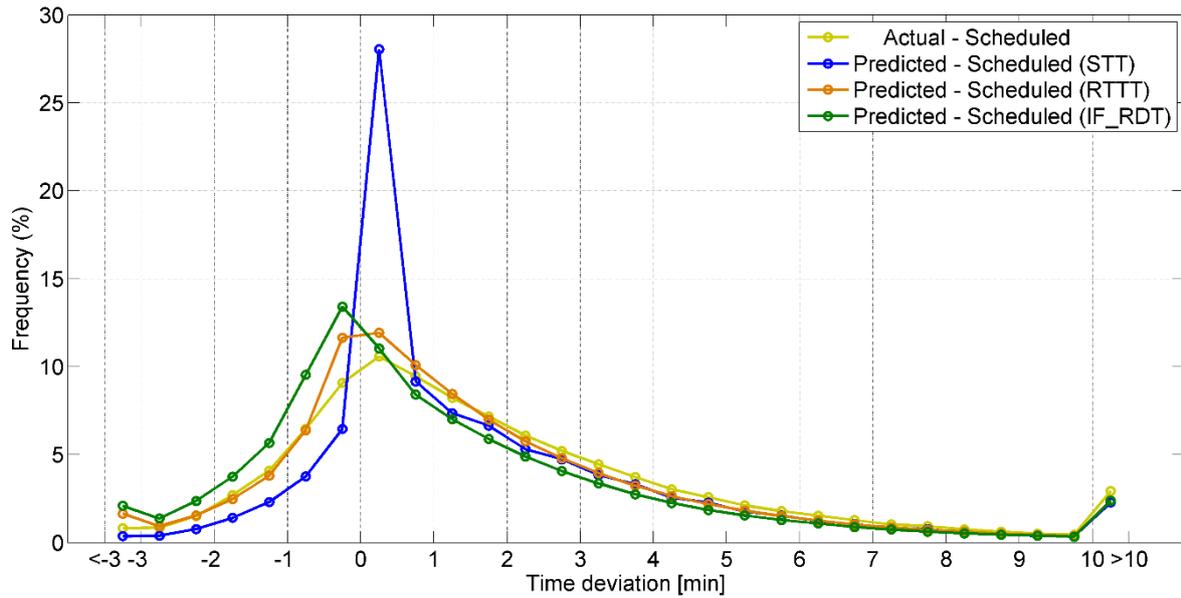


Figure 6: Information based on static timetable and alternative prediction schemes compared with the actual vehicle arrival time

5.3 Temporal analysis of prediction error

Temporal heterogeneity of traffic, timetables and passenger demand may result with temporal variations with respect to prediction schemes' performance. Moreover, prediction schemes may vary in their capacity to capture variations in system conditions which may have consequences on their robustness. This was investigated by segmenting prediction errors by day of the week and time of day periods.

Table 3 presents the average, standard deviation and MAE of the prediction error by day of the week for the three different methods. For all methods, while the underestimation of waiting times prevails on all days of the week, its extent varies considerably. The RTI predictions were most accurate and reliable on Mondays and Tuesdays. However, the prediction accuracy and reliability deteriorate along the week with the worst performance obtained on Saturdays. Sundays, in contrast, performed like an average weekday. It is presumed that these differences are highly influenced by the extent to which timetables reflect the prevailing traffic conditions. The case study network has a common timetable for all weekdays and separate timetables for Saturdays and Sundays.

Table 3: Prediction error of alternative prediction schemes by day of the week (mm:ss)

		MON	TUE	WED	THU	FRI	SAT	SUN
	Average	00:05	00:10	00:17	00:21	00:22	00:32	00:20
STT	StD	01:51	01:46	02:18	02:25	02:38	02:52	01:50
	MAE	01:05	01:05	01:05	01:17	01:18	01:40	01:06
RTTT	Average	00:15	00:17	00:23	00:26	00:30	00:31	00:18

	StD	01:29	01:28	01:52	02:00	02:14	02:05	01:23
	MAE	00:46	00:47	00:53	01:02	01:03	01:05	00:50
	Average	00:39	00:42	00:48	00:33	00:53	00:58	00:41
IF_RDT	StD	01:36	01:31	01:53	02:56	02:16	02:08	01:26
	MAE	00:55	00:56	01:04	01:23	01:11	01:18	00:58

Further analysis regarding the temporal variations with respect to time of day was carried out. Prediction errors were calculated separately for trips that started at the AM peak (7:00-9:00), Off-Peak (9:00-15:30) and PM peak (15:30-19:00) periods. The latter is associated with less accurate and less reliable RTI provision (Table 4) for all prediction methods. In contrast, the AM peak performs surprisingly well. This may be due to its more homogeneous conditions. For all time-of-day categories the IF_RDT method is found less accurate and slightly less reliable than the RTTT method while obtaining better MAE than STT.

Table 4: Prediction error of alternative prediction schemes by time of day (mm:ss)

		AM	Off-Peak	PM
STT	Average	00:11	00:10	00:15
	StD	01:35	01:49	02:12
	MAE	01:02	01:06	01:15
RTTT	Average	00:20	00:17	00:21
	StD	01:27	01:24	01:48
	MAE	00:49	00:48	00:54
IF_RDT	Average	00:41	00:41	00:42
	StD	01:32	01:52	02:05
	MAE	00:59	01:03	01:06

5.4 Spatial analysis of prediction error

Similarly to temporal heterogeneity, spatial variations in prediction performance were investigated in order to identify recurring patterns. The evolution of prediction errors were analyzed for each of the eight routes and were found to exercise common trends which are hereby exemplified through the eastbound route of Line 1. Figure 7 presents how the performance of the prediction methods evolves along the route. This route consists of 33 stops, of which 3 serve as TPS (Fridhemsplan, Hötorget and Värtavägen, marked with a star next to their name). The horizontal curves show the mean prediction

error per stop for each prediction scheme, while the vertical bars represent the corresponding standard deviations.

The average error of the STT scheme fluctuates greatly along the line with the value varying between ± 1.5 minutes. It is clearly evident that the fluctuations occur directly after a TPS. STT obtains the most accurate predictions at TPS locations followed by an immediate substantial increase in the prediction error, which is then reduced gradually until reaching the next TPS. It is argued that this pattern is determined by driving patterns that are related to the control scheme (Cats et al. 2012). In contrast, the average error of RTTT and IF_RDT does not fluctuate much along the line and remains within the range of [0,1]. These schemes also seem to capture better the schedule recovery effect, as TPS do not have a substantial effect on the prediction error. The negative average error on the last stretch of the route following the last TPS is attributed to driver behavior patterns as they wish to prolong their break at the end terminal. While this phenomenon is not reflected in the timetable, the RTTT and IF_RDT methods that incorporate travel times of preceding buses are able to reproduce this correctly.

The reliability of the STT prediction scheme deteriorates considerably along the route while the variations in prediction errors decrease along the line for RTTT and IF_RDT. The negative trend of the STT could be explained by the propagation of uncertainty associated with traffic conditions, dwell time and driver behavior along the route, which STT fails to capture due to its reliance on the timetable. This is reflected in the deterioration of service regularity along these routes, in line with previous studies (Chen et al. 2009 and Cats et al. 2012). In contrast, the alternative methods embed real-time data which results with predictions that adapt to changing downstream conditions. The standard deviation of the three methods is of the same magnitude for the origin terminal. This could be expected since these predictions will often collapse to the timetable. Interestingly, the reliability of IF_RDT is first worse than that of RTTT and STT. However, since the first TPS (Fridhemsplan) and onwards, the STT is the least reliable method while RTTT consistently improves its reliability and is persistently the most reliable scheme. It should be stressed that similar patterns to the ones reported in this section were observed for all case study routes.

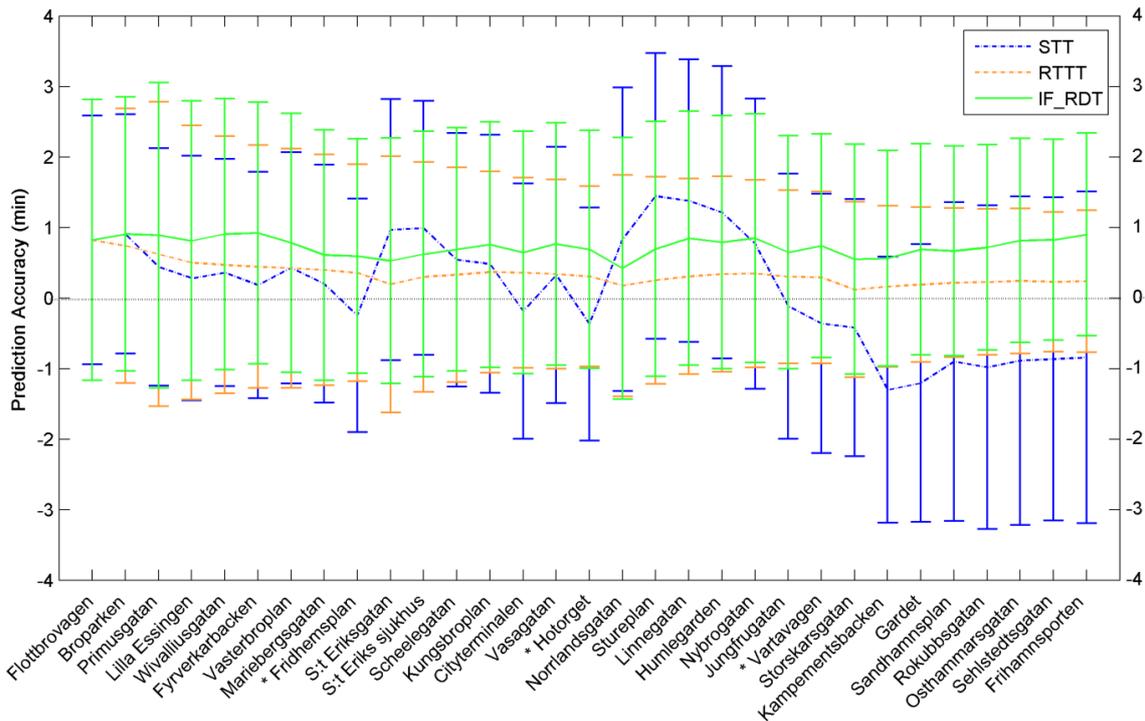


Figure 7: Prediction accuracy and reliability of alternative prediction errors along the eastbound route of Line 1

5.5 Service reliability and prediction performance

The added-value of RTI depends on the underlying service reliability. On one hand, in case of a perfectly punctual service, a static timetable will suffice as RTI will not entail any additional value. On the other hand, a highly irregular service is expected to be less predictable and hence plausibly limit the added-value of RTI. In order to explore this relationship, the evolution of service and RTI reliability along the line were investigated in Figure 8. In order to exemplify the observed patterns, the eastbound route of Line 1 is also used here for illustration. In addition to the mean absolute error of each RTI prediction scheme (Eq. 12) along the route, Figure 8 displays the corresponding indicators of service punctuality – the timetable mean absolute error (Eq. 11) - and service regularity – the headway coefficient of variation. Both service punctuality and service regularity tend to deteriorate along the line due to the propagation of uncertainty and the bunching effect. Control at TPS plays an important role in the evolution of reliability, as well as, the existence of partial trips which start or terminate at these stops. The performance of STT deteriorates simultaneously with schedule adherence since it is based on the assumption that current delays will be maintained further downstream in contrast to evidence. In contrast, the performance of RTTT and IF_RDT is proved immune to the deterioration of service punctuality and regularity and hence provide an increasing added-value at downstream stops where the static timetable is increasingly unreliable.

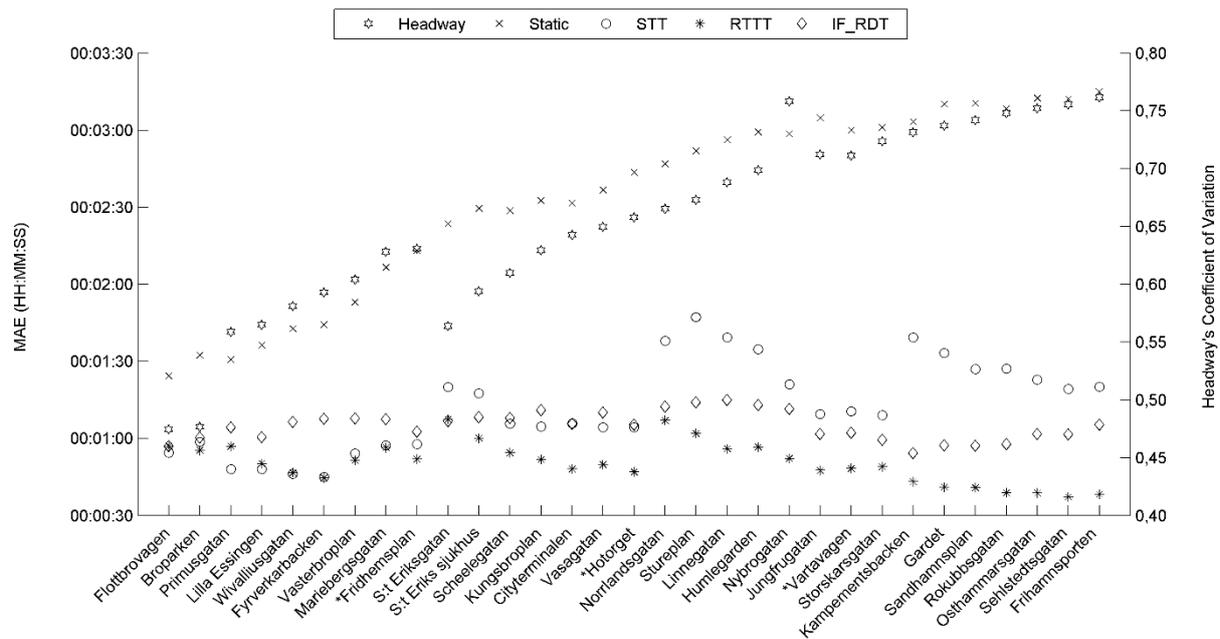


Figure 8: Mean absolute error of static and real-time information along the eastbound route of Line 1 and the corresponding coefficient of variation of headways

5.6 Comparing the impact of real-time information, service frequency and regularity improvements on passengers waiting times

The added-value of RTI could be further assessed by computing the improvement in service headways – the determinant of passenger waiting times in high-frequency networks - that is necessary in order to bring about an equivalent reduction in expected waiting times (see Section 5.2.1). A similar calculation was conducted by Mishalani et al. (2006) with respect to perceived waiting times. Figure 9 presents the combination of service frequency and regularity that is required for obtaining various expected waiting times. The functions are based on the assumption that passengers arrive randomly at stops. The equal waiting time curves displayed in Figure 9 correspond to waiting time ranging between 150 and 300 seconds with 50 seconds intervals. The performance of the case study network is marked with a star. The reduction in expected waiting times attained by replacing static information with the STT prediction scheme could alternatively be achieved by increasing service frequency by 45%. Such an increase implies 20 additional vehicle trips per hour for the case study network. Alternatively, service regularity has to be improved dramatically, reducing the average coefficient of headway variations from 0.76 to 0.30. Such a reduction requires substantial bus preferential treatments. Furthermore, the expected waiting time gains obtained by RTTT correspond to a 60% increase in service frequency or 27 additional trips per hour for the case study network. In fact, such a decrease in waiting times is not attainable through service regularity improvements with the existing service frequency (see Figure 9).

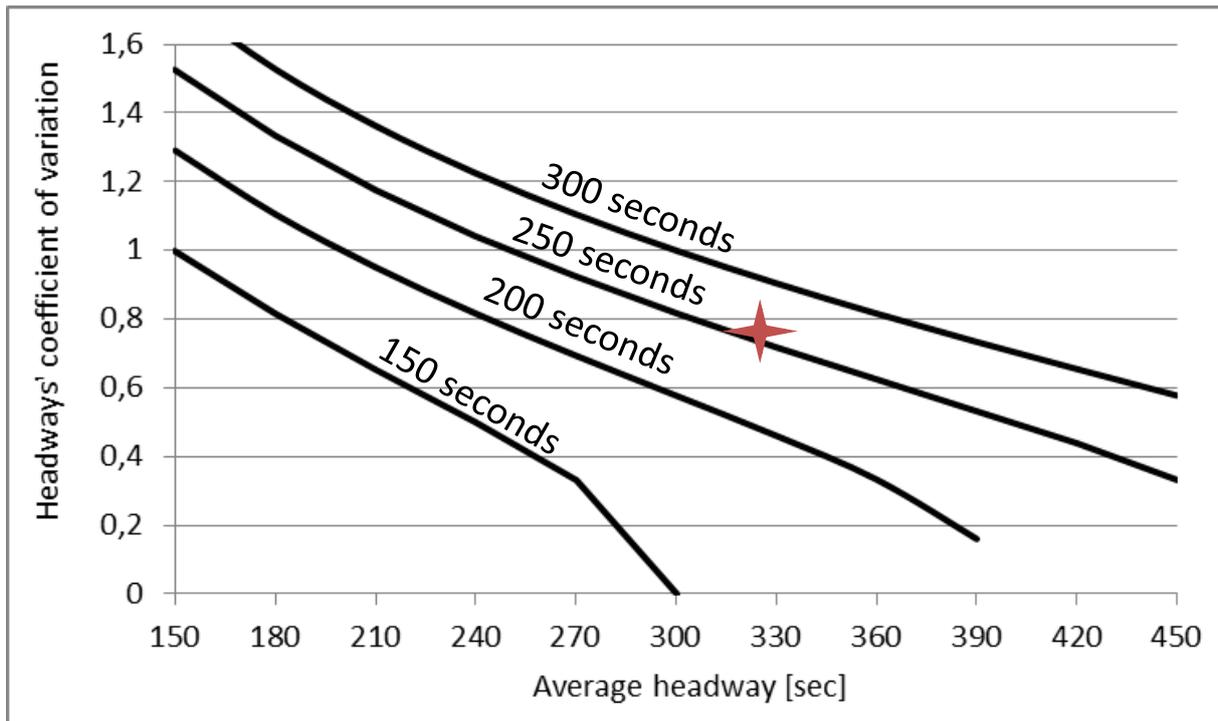


Figure 9: Equal-waiting time curves obtained by combinations of service frequency and regularity; the performance of the case study network is marked with a star

6. Discussion

This paper formulated three alternative prediction schemes for bus arrival times. The base scheme, STT, was constructed based on a commonly deployed RTI provision system which considers the interplay between the timetable, current delay and the schedule control strategy. The two additional schemes, RTTT and IF_RDT incorporate real-time data from a number of preceding vehicles. The former performs successive predictions between TPS while the latter involves iterative predictions of headways and dwell times at intermediate stops. The performance of the prediction schemes was evaluated by analyzing their predictions of the next bus arrival at all stops of the trunk bus system in Stockholm inner-city during a period of 45 days with one minute intervals. The execution times of all of the prediction schemes enables their online implementation.

All of the prediction schemes result with a substantial added-value for passengers compared with the information that can be derived from static timetables. The incorporation of travel times experienced by preceding vehicles in RTTT improves the prediction accuracy and reliability, and reduces the mean absolute error by 25% compared with the predictions generated by STT. The RTTT scheme also proves to be the most robust prediction scheme. It performs well on different days of the week, times of day, and maintains a consistently accurate and reliable prediction along the entire route. In addition, the predictions yielded by RTTT and IF_RDT outperformed the STT scheme in reproducing fleet punctuality performance which can support operational decisions.

The reliability and robustness – the degree to which performance metrics remain stable in time and space - of the most elaborative scheme, IF_RDT, underperformed compared with RTTT. Moreover, the RTTT scheme does not require parameter estimation. The sensitivity analysis suggests that the more downstream buses are used for the prediction the better, although with diminishing returns. In contrast, the IF_RDT method relies on the availability of dwell time coefficients. The performance of

IF_RDT was overshadowed by the systematic underestimation of dwell times which resulted with an average underestimation that ranges between 10% for short prediction horizons to 20% for longer horizons because of the accumulated effect produced by the iterative prognosis. This error may stem from two sources: the need to calibrate the dwell time function (parameters α and β), possibly for different time-of-day and weather conditions; and the need to capture more refined demand variations (e.g. 15 minutes intervals). Furthermore, real-time passenger counts or in their absence an online prediction of passenger flows, can be embedded into bus arrival predictions by integrating the method proposed recently by Ma et al. (2014).

7. Conclusions

The conclusions of this study and directions for further development can be grouped into three pillars: practical implications for bus operators, potential extensions of the prediction schemes and in the context of passengers' response to RTI.

First, the results of this study suggest that developing applications to personal mobile devices is an inexpensive dissemination channel that can obtain substantial time savings. These savings can yield gains equivalent to expensive measures to improve level-of-service improvements. The deployment of the simple timetable-based prediction yielded in the case study network a substantial improvement in bus arrival time predictions when compared the timetable. Implementing more advanced prediction schemes, such as those evaluated in this study, further improves the quality of RTI. In addition, better predictions can also be utilized by operators to move beyond reactive monitoring strategies to active and even proactive control and thus deploy more efficient and effective real-time decisions (Cats et al. 2012). In order to quantify the economic gains associated with the added-value of RTI to operators, further research into the effects of more accurate projections on control center decision making is needed.

Second, the prediction methods proposed in this paper can be further improved by introducing an optimization procedure for determining the number of preceding vehicles to be included in the prognosis (Fadaei Oshyani and Cats 2014) or testing alternative weighting functions (Esawey and Sayed 2010). Moreover, non-linearity in dwell time and driver behavior patterns could be embedded into the prediction scheme to better capture the complexity of public transport dynamics. There is empirical evidence that drivers adjust their speed in response to RTI on schedule adherence (Cats et al. 2012, Ji et al. 2014). This driving pattern could be potentially accounted for in future prediction schemes in line with Lin and Bertini (2004). Additional information on vehicle scheduling and exogenous variables will enable to capture the potential propagation of delays from one trip to the other caused by vehicle trip chaining. Using high-frequency penetration of vehicle position data will also allow to refine the sensitivity of the prediction scheme to instantaneous conditions such as bus traffic congestion and delays caused by bus stop capacity. Finally, simulation models can be embedded into prediction schemes in order to incorporate traffic dynamics and public transport operations into bus travel time predictions (Toledo et al. 2010, Hans et al. 2015).

The online transmission of passenger counts data is expected to become more prevalent and its incorporation will potentially facilitate more accurate dwell time predictions. Dessouky et al. (2003) demonstrated the benefits from real-time availability of passenger information when applying real-time transfer coordination. Moreover, it would facilitate the generation of RTI concerning on-board crowding conditions at downstream stops.

Third, the results of this study can also be considered in the context of passengers' response to RTI. Previous studies that investigated the impact of travel information accuracy on travelers' decisions have used simulation (Hickman and Wilson 1995), analytical models (Ettema and Timmermans 2006) or behavioral experiments (Ben-Elia et al. 2013). The results indicate that the potential waiting time gains obtained by the best performing prediction scheme (RTTT) correspond to a 60% increase in service frequency and are not attainable through service regularity improvements or a 33% improvement, when compared with the STT scheme. These findings underpin the potential benefits from improving the quality of online buses performance when compared with alternative measures to reduce passenger waiting times. Previous studies have demonstrated that information provision can yield waiting time benefits as passengers can better coordinate their arrival with the vehicle arrival (Dziekan and Kottenhoff 2007, Watkins et al. 2011, Brakewood et al. 2014). Future research should examine the inter-relation between service reliability and RTI reliability, for example by conducting an empirical study of different service types. The added-value of disseminating and improving RTI could be then evaluated for different service reliability categories.

Large discrepancies between the provisioned and actual arrival time can result with a diminishing confidence in the quality of RTI. Cats and Gkioulou (2015) demonstrated how passengers accumulated experience on both service and RTI reliability can evolve in a day-to-day learning modelling framework. The importance of RTI accuracy and reliability is expected to increase in the years to come as passengers increase their expectations in the age of information and operators need to execute more proactive operation strategies. Future research should assess the determinants of information acquisition and usage and the potential value of informing users on the uncertainty associated with the information provisioned.

Acknowledgements

The authors would like to thank the anonymous referees for their helpful review of an early version of the paper. The authors are grateful for the data and advice that were kindly provided by SLL, the Transport Administration of Stockholm County Council.

References

Ben-Elia E., Di Pace R., Bifulco G.N. and Shiftan Y. (2013). The impact of travel information's accuracy on route-choice. *Transportation Research Part C*, 26, 146-159.

[Bin Y., Zhongzhen Y. and Baozhen Y. \(2006\). Bus arrival time prediction using support vector machines. *Journal of Intelligent Transportation System*, 10 \(4\), 151-158.](#)

Bowman L.A. and Turnquist M.A. (1981). Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A*, 15 (6), 465-471.

Brakewood C., Barbeau S. and Watkins K., (2014). An experiment evaluating the impacts of real-time transit information on bus riders in Tampa, Florida. *Transportation Research Part A*, 69, 409-422.

Cathey F.W. and Dailey D.J (2003). A prescription for transit arrival/departure prediction using automatic vehicle location data. *Transportation Research Part C*, 11, 241-264.

Cats O. and Gkioulou Z. (2015⁴). Modelling the impacts of public transport reliability and travel information on passengers' waiting time uncertainty. *EURO Journal of Transportation and Logistics*, 1-24. DOI 10.1007/s13676-014-0070-4.

Cats O., Koutsopoulos H.N., Burghout W. and Toledo T. (2011). Effect of real-time transit information on dynamic passenger path choice. *Transportation Research Record*, 2217, 46-54.

Cats O., Larijani A.N., Ólafsdóttir A., Burghout W., Andreasson I. and Koutsopoulos H.N. (2012). Holding control strategies: A simulation-based evaluation and guidelines for implementation. *Transportation Research Record*, 2274, 100-108.

Cats O. and Loutos G. (2015). Real-time bus arrival information system – An empirical evaluation. *Journal of Intelligent Transportation Systems*. 1-14. DOI: 10.1080/15472450.2015.1011638.

Caulfield B. and O'Mahony (2007). An examination of the public transport information requirements of users. *IEEE Transactions on Intelligent Transport Systems*, 8 (1), 21-30.

Chang H., Park D., Lee S., Lee H. and Baek, S. (2010). Dynamic multi-interval bus travel time prediction using bus transit data. *Transportmetrica*, 6 (1), 19-38.

Chen M., Liu X. and Xia J. (2005) Dynamic Prediction Method with Schedule Recovery Impact for Bus Arrival Time. *Transportation Research Record*, 1923, 208-217.

Chen M., Liu X., Xia J. and Chien S.I. (2004). A dynamic bus-arrival time prediction model based on APC data. *Computer-Aided Civil and Infrastructure Engineering*, 19 (5), 364–376.

Chen G., Yang X., An J. and Zhang, D. (2011). Bus-arrival-time prediction models: Link-based and section-based. *Journal of Transportation Engineering*, 138 (1), 60-66.

Chen X., Yu L., Zhang Y. and Gou J. (2009). Analyzing urban bus service reliability at the stop, route and network levels. *Transportation Research Part A*, 43 (8), 722-734.

Chow W., Block-Schachter D. and Hickey S. (2014). Impacts of real-time passenger information signs in rail-stations at the Massachusetts Bay transportation authority. *Proceedings of the 93rd Transportation Research Board Annual Meeting*, Washington DC.

Dessouky M., Hall R., Zhang L. And Singh A. (2003). Real-time control of buses for schedule coordination at a terminal. *Transportation Research Part A*, 37 (2), 145-164.

Dziekan K. and Kottenhoff K. (2007). Dynamic at-stop real-time information displays for public transport: effects on customers. *Transportation Research Part A*, 41 (6), 489-501.

El Esawey M. and Sayed T. (2010). Travel time estimation in urban networks using neighbour links travel time data. *Transportation Research Board the 89th Annual Meeting*, Washington, D.C.

Ettema D. and Timmermans H. (2006). Costs of travel time uncertainty and benefits of travel time information: Conceptual model and numerical examples. *Transportation Research Part C*, 14, 335-350.

Fadaei Oshyani M. and Cats O. (2014). Real-time bus departure time predictions: Vehicle trajectory and countdown display analysis. *IEEE 17th international conference on Intelligent Transport Systems (ITSC)*, 2556-2561.

Gooze A., Watkins K.E. and Borning A. (2013). Benefits of real-time transit information and the impacts of data accuracy on the rider experience. *Transportation Research Board 92nd Annual Meeting*.

- Hans E., Chiabaut N. and Leclercq L. (2015). Investigating the irregularity of bus routes: highlighting how underlying assumptions of bus models impact the regularity results. *Journal of Advanced Transportation*, 49, 358-370.
- Hickman M.D., and N.H.M. Wilson (1995). Passenger travel time and path choice implications of real-time transit information. *Transportation Research Part C*, 3 (4), 211-226.
- Jeong R. and Rilett L.R. (2005). Prediction model of bus arrival time for real-time applications. *Transportation Research Record*, 1927, 195–204.
- Ji Y., He L. and Zhang M. (2014). Bus drivers' response to real-time schedule adherence and effects on transit reliability. *Transportation Research Record*, 2417 (3), 1-9.
- Lin, W.H. and Bertini, R.L. (2004). Modeling schedule recovery processes in transit operations for bus arrival time prediction. *Journal of Advanced Transportation*, 38 (3), 347-365.
- Loutos G. (2013). Development of Prediction Schemes for Real-time Bus Arrival Information. Master thesis, Royal Institute of Technology, KTH.
- Ma Z., Xing J., Mesbah M. and Ferreira L. (2014). Predicting short-term bus passenger demand using a pattern hybrid approach. *Transportation Research Part C*, 39, 148-163.
- Mazloumi E., Rose G., Currie G. and Sarvi M. (2011). An integrated framework to predict bus travel time and its variability using traffic flow data. *Journal of Intelligent Transportation Systems: Technology, Planning and Operations*, 15 (2), 75-90.
- Mishalani R.G., McCord M.M. and Wirtz J. (2006). Passenger wait time perceptions at bus stops: Empirical results and impacts on evaluating real-time bus arrival information. *Public Transportation*, 9 (2), 89-106.
- Newell G.F. and Potts R.B. (1964). Maintaining a bus schedule. *Proceedings of 2nd Australian Road Research Board*, 2, 388-393.
- Patnaik J., Chien S. and Bladikas A. (2004). Estimation of bus arrival times using APC data. *Public Transportation*, 7(1), 1-20.
- Rahman Md. M., Wirasinghe S.C. and Katan L. (2013), Users' views on current and future real-time bus information systems. *Journal of Advanced Transportation*, 47 (3), 336-354.
- Shalaby A. and Farhan A. (2004). Prediction model of bus arrival and departure times using AVL and APC data. *Public Transportation*, 7 (1), 41-61.
- Tang L. and Thakuriah P.V. (2012). Ridership effects of real-time bus information system: A case study in the City of Chicago. *Transportation Research Part C*, 22, 146-161.
- TCRP (2003). Transit capacity and quality of service manual, 2nd ed. Report 100. Transportation Research Board of the National Academies, Washington, D.C., 2003.
- TCRP (2008). AVL systems for bus transit: update. TCRP Report 73. Transportation Research Board, Washington, D.C.
- Toledo T., Cats O., Burghout W. and Koutsopoulos H.N. (2010). Mesoscopic Simulation for Transit Operations. *Transportation Research Part C*, 18 (6), 896-908.

Vu N.H. and Khan A.M. (2010). Bus running time prediction using a statistical pattern recognition technique. *Transportation Planning and Technology*, 33 (7), 625-642.

Watkins K.E., Ferris B., Borning A., Rutherford G.S. and Layton D. (2011). Where is my bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. *Transportation Research Part A*, 45 (8), 839-848.

Yu, B., Lam, W. H., and Tam, M. L. (2011). Bus arrival time prediction at bus stop with multiple routes. *Transportation Research Part C*, 19 (6), 1157–1170.