CrossMark

RESEARCH PAPER

Bunching and Headway Adherence Approach to Public Transport with GPS

**Young-Ji Byon[1] · Cristián E. Cortés[2] · Young-Seon Jeong[3] ·
Francisco Javier Martínez[2] · Marcela A. Munizaga[2] · Mauricio Zúñiga[2]**

Abstract Emerging GPS devices enable new data collection opportunities for transit performance monitoring. In addition to the fact that GPS devices can replace labor-intensive survey techniques, they also collect traffic information throughout transit routes that the traditional fixed loop detectors cannot. If public transit vehicles are equipped with GPS devices, it is possible to monitor the performance of public transit services throughout the routes and alert the associated authorities at significantly lower costs about potential problems for corrective actions to be taken. Transantiago, a major public transit service provider in Santiago, Chile, has recently installed GPS sensors on all its vehicles which provides an excellent venue on which an innovative transit monitoring methodology can be modeled and applied. This paper first conducts current-status analysis on distributions of headways throughout a route in Santiago by processing extensive raw GPS data from transit vehicles. Then, unique transit headway adherence indices are developed with respect to the expected passenger waiting time and are presented in forms of two-dimensional tempo-spatial graphs. The analysis of real-life data collected from bus GPS probes in Santiago, Chile indicates that GPS devices in transit buses can effectively provide the proposed performance measures throughout the route on a daily basis.

Keywords Transit performance · Transit headway · Bunching · Transit schedule · GPS · Bus probes

✉ Young-Seon Jeong
young.jeong@jnu.ac.kr

Young-Ji Byon
youngji.byon@kustar.ac.ae

Cristián E. Cortés
ccortes@ing.uchile.cl

Francisco Javier Martínez
fmartine@ing.uchile.cl

Marcela A. Munizaga
mamuniza@ing.uchile.cl

Mauricio Zúñiga
mzuniga@gmail.com

[1] Department of Civil Infrastructure and Environmental Engineering, Khalifa University, PO BOX 127788, Abu Dhabi, United Arab Emirates

[2] Department of Civil Engineering, Universidad de Chile, Blanco Encalada 2002, Santiago, Chile

[3] Department of Industrial Engineering, Chonnam National University, Gwangju, Republic of Korea

1 Introduction and Background

Traditionally, traffic monitoring-related applications often have required labor-intensive data collection processes in the forms of surveys. With the emergence of GPS devices, transportation engineers have started exploring the use of the newly available technology for various branches of transportation applications in efforts to replace the traditional means of data collection processes. Byon et al. [1] dispatch GPS units with dedicated probe vehicles to occasionally monitor the traffic conditions hoping to replace traditional probing methods typically with a dedicated driver and a data-logging employee. Herrera et al. [2] evaluate the traffic data obtained from GPS-enabled mobile phones. Work et al. [3] develop a Kalman filtering-based approach for highway traffic estimations using GPS-enabled mobile devices. Byon et al. [4], and Byon and Liang [5] utilize the GPS embedded cell phones as traffic probes for gathering traffic conditions from cell

Springer

phones that are voluntarily hovering over road networks eliminating the need of having expensive dedicated probes. However, in order for such approaches to be feasible, the mode of transportation of the cell phone users needs to be correctly identified as "auto" mode which refers to drivers or passengers of private vehicles. Instead of replacing the traditional methods from GPS entirely, Kwon et al. [6] and Byon et al. [7] propose using GPS devices as aids complementing loop detectors. Cathey and Dailey [8] and Kumar et al. [9] develop algorithms that relate the speed of transit vehicles and general traffic for utilizing transit vehicles as traffic probes. The main idea is to utilize voluntarily moving transit vehicles for traffic monitoring purposes at no additional operating costs. However, at the time of their research, GPS was not widely available and they had to use fixed sensors on the road and interpolate the speed values between sensors for their estimations. Song et al. [10] install GPS devices on Bus Rapid Transit (BRT) vehicles and attempt to use them as traffic probes. Above-mentioned studies primarily focus on applying the GPS technology for monitoring general traffic conditions.

For the public transit sector, Agarwal and Goel [11] apply the GPS technology as the main tool for the Automatic Vehicle Location (AVL) system of the transit vehicles for monitoring transit service performance measures. Ledwitz [12] proposes the initial approaches in utilizing GPS for the public transit and Wipke [13] focuses on more specific application of using GPS in transit-on-demand applications for reducing vehicle miles traveled (VMTs). Munizaga et al. [14] use GPS data from transit vehicles to monitor the commercial speed of buses as a direct performance measure. One of the major issues with transit vehicles is that they tend to bunch up and increase the expected passenger waiting times. The correlation between the speed of buses and the bunching itself needs further investigations as the speed value of buses is only one aspect various performance measures. Liao and Liu [15] propose a framework for transit performance analyses that incorporates the GPS-based transit probing and suggest that GPS may be able to help transit operators identify causes of bunching due to wheel-chair lifting, traffic signaling, and mix of bus types. The authors only discuss the issue qualitatively and do not quantify the "bunching" with respect to the expected passenger waiting times. Daganzo and Pilachowski [16] propose an adaptive control scheme that adjusts transit vehicles' speeds in efforts to reduce bunching. El-Geneidy et al. [17] assume that the reliability of a transit system can be represented by the route-level variations of run times and headways. Chen et al. [18] recognize that properly evaluating or monitoring transit performances requires discretization of route-level performance measures into stop-level performance measures. The authors collect data using a labor-intensive survey method.

Based on existing researches, there is a need for a methodology that is economically feasible and can provide transit performance measures relating to the bunching phenomena with respect to the expected passenger waiting times, at both stop-level and route-level that utilizes automatically collected data with GPS devices throughout the route on a daily basis. Owais and Hassan [19] also recognize that incorporating stop-level simulation helps transit assignment models. In rail transit, Jamili [20] and Tamannaei et al. [21] solve rail transit rescheduling problems under perturbations and double-track scenarios, respectively. In rail transit, a number of trains in operations are usually much smaller than fleet sizes in bus transit. By definition, they also have dedicated tracks and multiple location determining sensors. By utilizing GPS as such sensors for bus transit as in this paper, it opens new opportunities for bus transit where monitoring strategies have been traditionally only available for rail transit.

Once the access to the massive amount of automatically collected data is granted, there is no need to simulate such data [22]. After performance measures throughout a route at all stops are available and causes of delays are identified, the newly available information can help determining optimal schedules, for example, through the determination method introduced by Furth and Muller [23].

2 Transit System in Santiago, Chile (Transantiago)

Transantiago is the public transport system of Santiago, the capital of Chile, with a population of over 6 million people. It was launched in February 2007 and is considered as "the most ambitious transport reform ever tried by a developing country". (*The Economist*, London, U.K., February 9th, 2008). It has over 200 km of dedicated bus lanes and reorganized the old bus route to integrate with the city's metro. Since 2010, GPS devices have been installed on over 6000 buses operating on over 700 different routes, generating GPS data that are available every 30 s from all buses during the entire operation period. This is equivalent to 40,000,000 positions and speed values of buses during 1 week period. The data collection mechanism is already in place and continuously the data are fed into the transportation research lab at the Universidad de Chile (Santiago, Chile), currently, for post-processing. Figure 1 shows the typical Transantiago bus in operation and a scene of a congested street in downtown during the afternoon rush hour.

Fig. 1 Congested street with transit buses during the afternoon rush hour in Santiago, Chile (*bottom left*), a typical Transantiago bus operating on Las Condes line (401) in Santiago, Chile (*top right*)

3 Bunching vs Adherence to Scheduled Headway

Munizaga et al. [14] focus on commercial speeds of buses assuming that a faster speed would indicate a better level of service (LOS), which is partially true. However, after all, the primary objective of any transit system is to satisfy their passengers. One of the most important passenger LOS performance measures is a time consumed by individuals during a trip, which is a sum of waiting time and travel time in buses. Therefore, a transit service, in general, should satisfy its passengers by minimizing a combination of waiting times at the bus stops and travel times of the buses on the road. Often, because of an accumulation of randomness in conjunction with traffic congestions, further down the route, buses tend to experience a phenomenon known as "bunching". Bunching occurs when headways between consecutive buses become smaller than scheduled headways. Since a fleet size is generally limited, the bunched group of buses consequently leaves large headways before and after them. Some studies [24, 25] in the past indicate that the expected passenger waiting time, W, is a function of average headway, μ, and headway variance, S^2 [26]:

$$W = \mu \times \frac{\left(1 + \frac{s^2}{\mu^2}\right)}{2} \tag{1}$$

where μ = mean headways between buses, and S^2 = variances of headways between buses.

If a bunching occurs, inevitably, some of the buses deviate from scheduled headways resulting in longer expected waiting times for passengers. The situation gets worse when multiple buses arrive at the same bus stop simultaneously with a nearly zero headway. Passengers arriving at the bus stop after the bunched buses have left will most likely have to wait a much longer time period for the arrival of the next bus. In addition, the next bus will have a higher chance of getting fulfilled with a maximum number of passengers to its capacity forcing remaining passengers at the bus stop to wait for additional buses until the maximum-capacity loading issue dissipates over time. In conjunction with Automatic Passenger Counting (APC) system and Automatic Fare Collection (AFC) system that are planned to be installed in near future, buses' load information with some socio-economic variables can be included for computing more detailed performance measures when the data become available.

Adherences to scheduled headway directly indicate the deviations of observed headways from scheduled headways set by the transit authority, and bunching is the main "cause" of the deviation. It seems that they are very similar performance measures; however, one is monitoring the occurrence of the problem, while the other is referring to the cause of the problem. This paper develops scheduled headway adherence indices that, in turn, also indirectly capture the bunching phenomena.

It is important to note that a severity of bunching is not only a function of a number of vehicles in the bunched group, but also a function of the scheduled headway. In other words, if the scheduled headway is relatively short, and then even if buses are close to each other, the event is not considered as a bunching. In short,

Severity of Bunching = f (number of vehicles in a bunched group AND scheduled headway). (2)

It is a challenge to derive indices that can indicate an adherence to scheduled headways while also providing insights on severity of bunching.

4 Study Objective

The objective of this paper is modeling indices and utilizing a massive amount of GPS data collected by transit buses for monitoring transit performances throughout routes. The research is conducted in three phases and their respective objectives are:

1. To analyze the running time and headway distributions of a route using GPS data;
2. to develop an index of bunching with respect to the expected passenger waiting time and to provide tempo-spatial graphical representations of the computed indices;

3. to develop an index of scheduled headway adherence associated with bus operations and tempo-spatial graphical representations of computed indices.

5 Study Region and Data Collection

Route 506I in Fig. 2, that runs, from the west-end to the east-end of the central Santiago, is chosen for data collection, because it is one of the busiest major bus trunk lines of Transantiago with typical running time of 1.5 h. The route is considered to be a typical major route in terms of ridership, and it includes 1.5 km of exclusive bus lanes locally acting as BRT (Bus Rapid Transit) operations. In other words, the route experiences the most typical traffic conditions throughout the city with a short section with exclusive bus lanes. Even though this route is chosen due to the unique feature of having a short section of dedicated bus lane, it is not expected to significantly affect the analyses when compared to the other sections of the route due to its relatively short length with respect to the entire route. The route generally experiences the bunching phenomena in the latter half of the route on the east side. The focus of this paper is to propose a methodology of how raw GPS data from buses can be processed and utilized for transit performance monitoring. For the period of 1 week, including 5 weekdays and 2 weekend days, the GPS data from all buses on the Route 506I were collected and post-processed. It is noteworthy that the developed methods are applicable to all other buses operating on other routes.
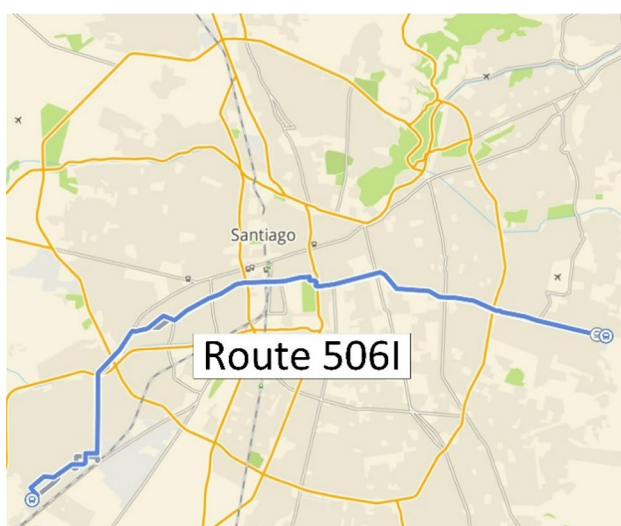

Fig. 2 Transit bus route 506I of Transantiago in Santiago, Chile (Accessed https://www.openstreetmap.org in December 10th 2016)

6 Required Sample Sizes

Li et al. [27] investigate general sample size requirements for establishing a consistent method for GPS-based network monitoring. Their work focuses on determining how many probes on a given link are needed for reliably estimating traffic conditions on that link. This work contributes towards resolving the "significant sampling challenges" that Smith et al. [28] refer to. Authors find that a reliable traffic monitoring usually requires 5–10 readings to estimate link travel time, delay, and work zone conditions. Byon et al. [4] recommend that, in an AGPS cell phone traffic monitoring system, for each desired road link, at least five AGPS cell phones need to be queried at once.

From a perspective of public transit monitoring, based on other relevant research works, it can be interpreted as there should at least be five sample readings between any particular pair of successive bus stops. Even though it is theoretically acceptable to have five sample readings from one bus for estimating conditions on a link over a longer time horizon, it would make sample observations more reliable if there are multiple buses operating between two stops. If achieving such high-frequency services of operating more than 1 bus between two successive stops is not economically feasible, at least the sections with multiple bus routes simultaneously operating on them can benefit by merging the data from multiple routes by post-processing the GPS data together.

7 Phase 1: Run-Time and Headway Distribution

In the first phase of this study, the current state of transit operating conditions is analyzed by observing some basic performance measures that are running times and headways. First, the distribution of running times is analyzed for weekdays. Then, the headway distributions of the entire route for different days and periods are analyzed. The main purpose of this phase is to illustrate how the basic performance measures can be found directly and solely from the raw GPS data from transit buses.

7.1 Run-Time Distribution

Figure 3 shows the distribution of running times of Route 506I on weekdays of all buses aggregating all operation time periods. Durations of most trips are found to be between 80 and 120 min. It is noted that run times can easily be computed at the final bus stop, since the dispatching times at the first bus stop is known. GPS-based data have great advantages when the information throughout the route is of interest.
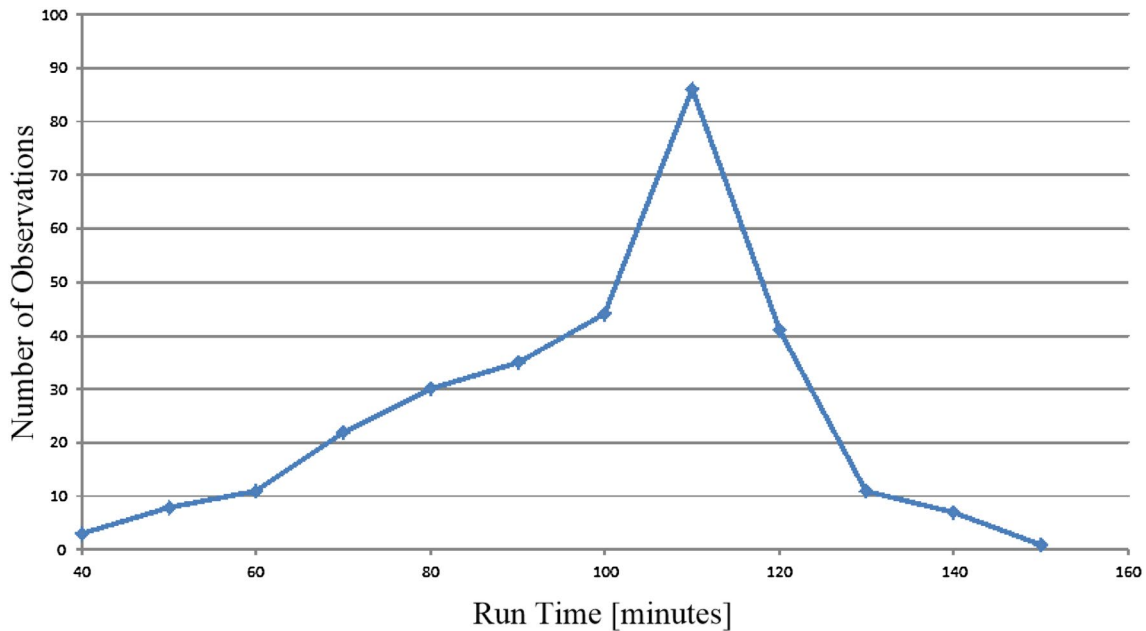
Fig. 3 Distribution of run times of 506I on weekdays

7.2 Headway Distribution

Passengers perceive a level of service of a transit system directly from the observed headways. Equation 1 emphasizes that observed headways are major variables for determining an expected passenger waiting time. However, to manually collect such information, one employee per each bus stop is required to log the observations. GPS devices on transit buses can replace such labor-intensive approaches. In this section, current headway distributions at different time periods on weekdays and weekends are analyzed.

Figure 4 shows observed headway distributions of the Route 506I on weekdays at four different time periods; 7–9 AM, 2–4 PM, 6–8 PM, and 9–11 PM. The dotted lines indicate the respective scheduled headways that

	7:00-9:00	14:00-16:00	18:00-20:00	21:00-23:00
Desired Headway	232	208	230	225
Mean	213	200	202	208
Standard Deviation	110	58	72	70

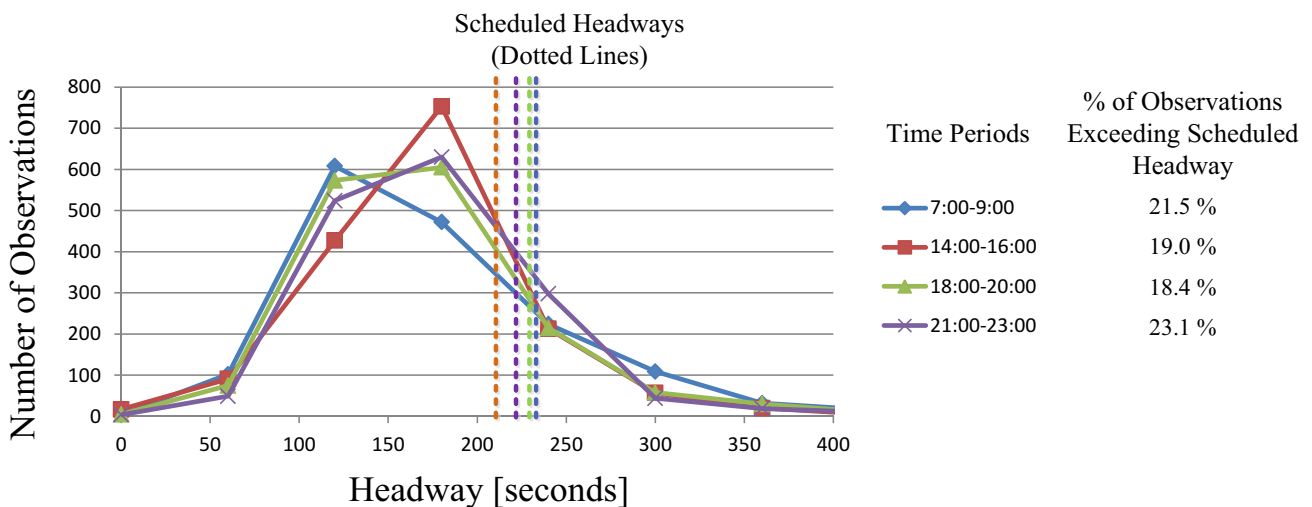| Time Periods	% of Observations Exceeding Scheduled Headway
7:00-9:00	21.5 %
14:00-16:00	19.0 %
18:00-20:00	18.4 %
21:00-23:00	23.1 %

Fig. 4 Weekday headway distributions

Transantiago desires. On weekdays, for all four time slots, mean observed headways are lower than scheduled headways. However, it is noted that still a considerable amount of headway observations are higher than the desired headways that may make the level of service of the transit system worse.

Figure 5 shows the headway distributions of weekend days. For the first two time periods of 7–9 AM and 2–4 PM, mean observed headways are longer than the scheduled headways, while the other time periods resulted in lower observed headways than the scheduled headways. Comparing to the weekday results in Fig. 4, the headway performance measures are generally found to be worse on weekends.

8 Phase 2: Bunching Analysis

In this phase, an index that indicates a severity of bunching with respect to expected passenger waiting times is derived and analyzed. It is noted that the bunching is not directly measured but rather indirectly captured.

8.1 Derivation of a Bunching Index

A bunching phenomenon can occur anywhere along a route. From a passenger's standpoint, a passenger who is waiting at a bus stop will recognize the bunching phenomenon regardless of where the bunching had originated from. By keeping that in mind, typically bus stops in Santiago are

separated by only 200–500 m. Therefore, the bus stops that are finely distributed along routes can act as good sensors for detecting the bunching phenomena. At a bus stop level, if a passenger waits longer due to bunching (or variations from scheduled headways due to bunching), the bunching index should capture the incident and increase the index value. As a starting point, the following relative waiting time ratio (namely RWTR) is considered as follows:

$$RWTR = \frac{OWT}{IWT} \tag{3}$$

where OWT is an observed waiting time, and IWT indicates an ideal and perfect condition waiting time with no bunching (i.e., perfect scheduled headway adherence).

In Santiago, transit buses operate based on scheduled headways instead of scheduled arrivals at bus stops. Therefore, passengers do not know exactly when a bus should arrive; they just know the schedule frequency and, consequently, the scheduled headway. Therefore, if passenger arrivals are uniformly distributed at a bus stop i, the total waiting time of all passengers at a bus stop i (TW_i) as a function of an arbitrary headway of h will be computed as follows:

$$TW_i = \sum_{j=1}^{n} \left(\frac{h_{i,j}}{2} * h_{i,j} * k \right) \tag{4}$$

where $h_{i,j}$ presents an observed headway for vehicle j at bus stop i, and k is a constant arrival rate of passenger. In Eq. (3), $\frac{h_{i,j}}{2}$ represents an expected average waiting time during the headway ($h_{i,j}$), and ($h_{i,j} * k$) indicates a total number of

	7 to 9	14 to 16	18 to 20	21 to 23
Scheduled Headway	257	202	240	290
Mean	285	233	212	246
Standard Deviation	81	104	59	80

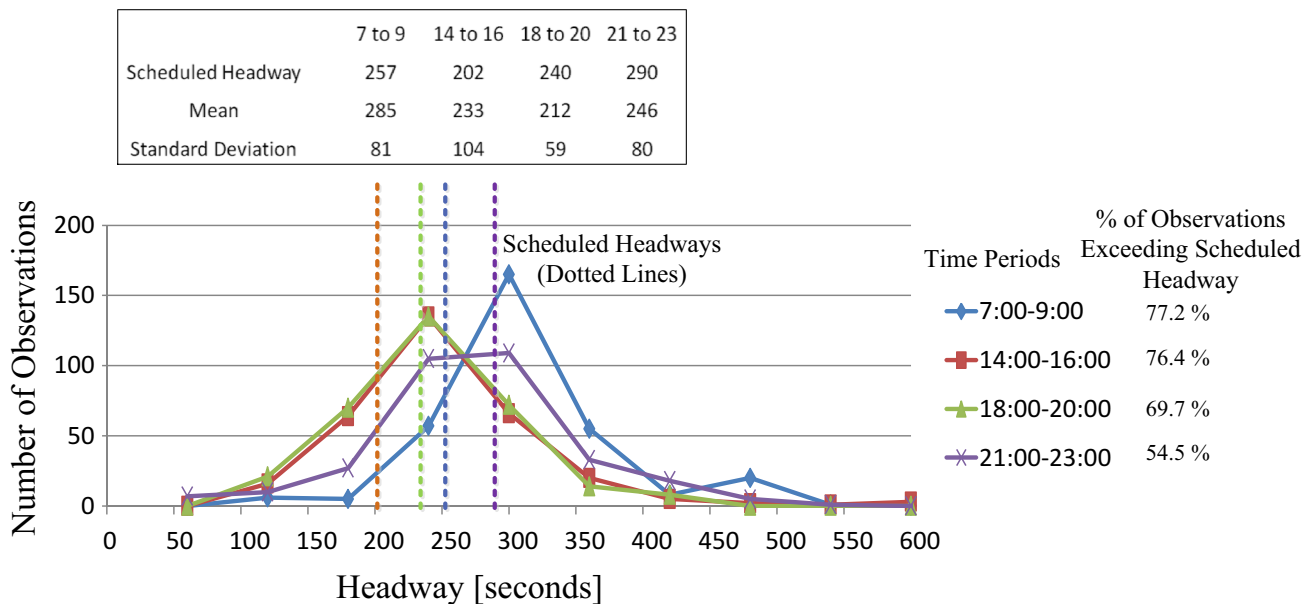| Time Periods	% of Observations Exceeding Scheduled Headway
7:00-9:00	77.2 %
14:00-16:00	76.4 %
18:00-20:00	69.7 %
21:00-23:00	54.5 %

Fig. 5 Weekend headway distributions

passengers during the headway ($h_{i,j}$). Thus, the total observed waiting times of all passengers at a bus stop i during the observed headway $h_{i,j}$ reduce down to $\sum_{j=1}^{n} \frac{h_{i,j}^2 * k}{2}$.

Similarly, if no bunching is observed (which means no deviations of headways from the scheduled one due to bunching), the total waiting times of all passengers at bus stop i should equal to $\sum_{j=1}^{n} \frac{h_{i,j}^{*2} * k}{2}$, where $h_{i,j}^*$ is a scheduled headway for vehicle j at bus stop i.

Therefore, the summation of RWTR reduces down to an index that captures the severity of bunching, which increases the expected passenger waiting times with respect to the ideal case waiting times:

$$\text{Index}_i = \sum_{j=1}^{n} \left(\frac{\frac{h_{i,j}^2}{2} * k}{\frac{h_{i,j}^{*2}}{2} * k} \right) = \sum_{j=1}^{n} \left(\frac{h_{i,j}}{h_{i,j}^*} \right)^2. \tag{5}$$

Once the summation of the ratio for each tempo-spatial cell is computed, it can either be divided by the monitoring time duration or a certain arbitrary time unit. We then differentiate the Index per Hour (IPH) from the Index per Observation (IPO). The former is divided by the monitoring time, while the latter is a normalized measure as it is divided by the number of observed buses, n. In this paper, a time window frame of 30 min (0.5 h) is arbitrarily chosen and consistently used in all the computations. Analytically,

$$\text{IPH}_i = \frac{\sum_{j=1}^{n} \left(\frac{h_{i,j}}{h_{i,j}^*} \right)^2}{0.5} \tag{6}$$

$$\text{IPO}_i = \frac{\sum_{j=1}^{n} \left(\frac{h_{i,j}}{h_{i,j}^*} \right)^2}{n}. \tag{7}$$

It is worth noting that the proposed indices, IPH and IPO, are similar to the Percentage Regularity Deviation Mean (PRDM), which is a popular index for describing the regularity of a bus transport service. The PRDM [29] is expressed as follows:

$$\text{PRDM}_i = \frac{\sum_{j=1}^{n} \left| h_{i,j}^* - h_{i,j} \right|}{n}. \tag{8}$$

A lower value of PRDM means better regularity of a bus service. However, the PRDM does not recognize the polarity of the differences. In practice, for passengers, $h_{i,j} > h_{i,j}^*$ (waiting longer than scheduled) is more significant than $h_{i,j} < h_{i,j}^*$ (waiting less than scheduled) with regard to the bus service regularity. Thus, it is needed to develop a regularity index to distinguish between the two

cases. Unlike the *PRDM*, our proposed indices can capture the transit operation states as in the following:

$$\text{IPH}_i(\text{or IPO}_i) \begin{cases} > 1, & \text{when } h_{i,j} > h_{i,j}^* \\ = 1, & \text{when } h_{i,j} = h_{i,j}^* \\ < 1, & \text{when } h_{i,j} < h_{i,j}^* \end{cases} \tag{9}$$

Figure 6 illustrates how normalized index values can be assigned to a confined tempo-spatial cell resulting in a grid-like graph that provides a visual representation of one day's performance measuring indices.

8.2 Results of Bunching Indices

8.2.1 Index Per Hour (IPH)

The IPH indices are computed for a typical weekday and a weekend day in a same week. The IPH indices are found to be generally higher on the weekday than the weekend day. Due to high congestions throughout the route and throughout the day, as expected, the IPH values turned out to be higher during a weekday than on a weekend day.

Each cell in Fig. 7 is colored with a corresponding IPH value according to the scale shown at the top of the diagram. It is noted that long horizontal blue bars indicate missing data that results from technical errors on the field. On the weekday, more cells show yellow-to-red colors indicating poor adherences to scheduled headways, which, in turn, give insights on the bunching phenomena occurring more often and throughout the route when compared with the weekend day which shows more blue-to-green cells representing less bunching occurring. In practice, transit operators can easily detect the problematic sections, so that they can consider re-positioning bus stops to avoid consistently congested sections.
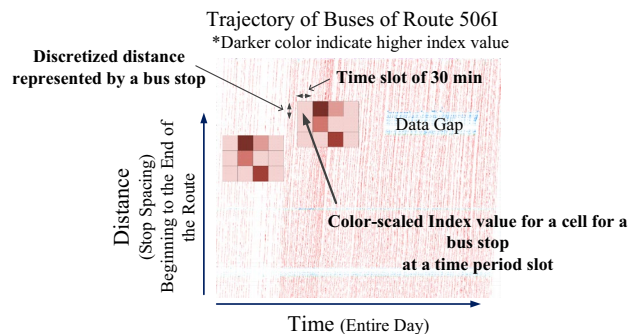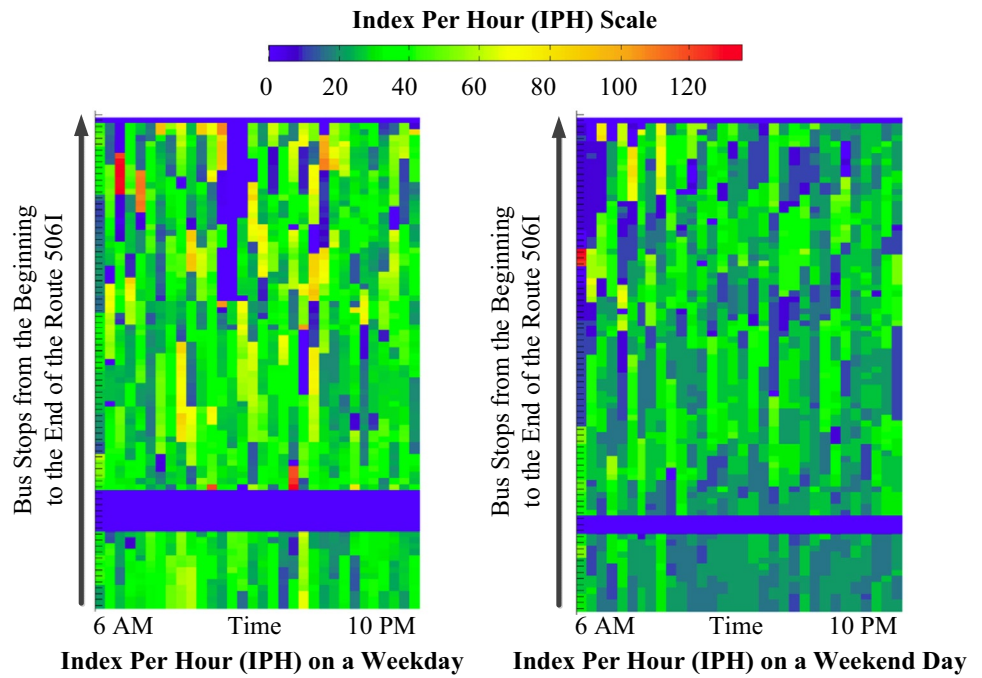

Fig. 6 Division of cells for assigning indices by a road-section represented by a bus stop and a time period of 30 min

Index Per Hour (IPH) Scale

0   20   40   60   80   100   120

**Index Per Hour (IPH) on a Weekday**   **Index Per Hour (IPH) on a Weekend Day**

8.2.2 Index Per Observation (IPO)

The IPO values for a typical weekday and a weekend day are computed and presented in Fig. 8. It is noted that IPO values are higher in the morning and afternoon peak hours on the weekday as expected. It seems that IPO and IPH follow similar patterns validating that both indices are behaving as expected, that is IPH and IPO should intuitively be higher on weekdays than weekend days due to generally heavier congestions.

Index Per Observation (IPO) Scale

1   3   6

Index Per Observation (IPO) on a Weekday   **Index Per Observation (IPO) on a Weekend Day**
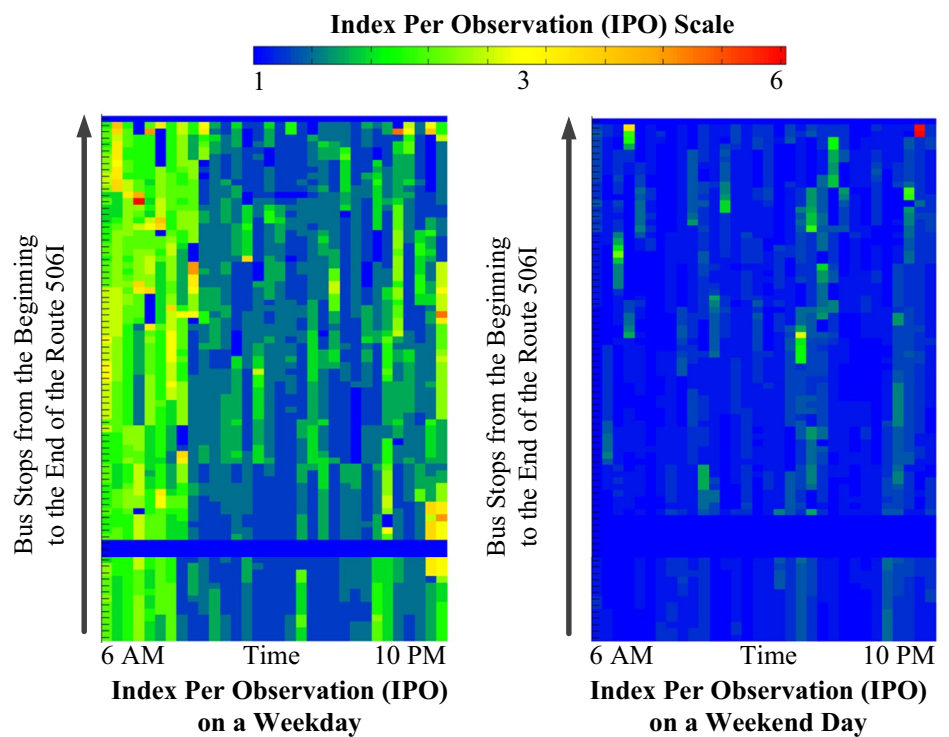
9 Phase 3: Scheduled Headway Adherence Analysis

In Phase 2, bunching indices have been developed and are functions of observed headways and scheduled headways. In Phase 3, slightly different performance measures that monitor a scheduled headway adherence are developed. Deviations from scheduled headways occur due to bunching, which, in turn, are affected by traffic conditions, such as congestions and traffic-signal optimization issues. Therefore, in addition to the proposed bunching indices in Phase 2, transit operators may also be simply interested in a tempo-spatial status of a route with respect to the scheduled headway adherence performance measures. Scheduled headway adherence monitoring can be approached from two different perspectives:

1. Adherence with respect to the scheduled headway;
2. adherence with respect to a variable headway demanded by passengers.

The formulations are shown below. The first perspective is to compare the observed headway against the scheduled headway to capture the deviation of observed headways with respect to what has been scheduled as promised by the transit authority to the public. The second perspective is to compare the observed headway against a variable headway that may be arbitrarily demanded by the public, which evaluates the current status of the transit service against the desired headway of the public. The second perspective is not a monitoring tool but rather a stress-testing evaluation tool to compare the current status of operation to presumably changing demands from the public, through time and space. Such tool would reveal critical sections of a route with low performances that need to be checked for enhancements and it may also reveal surprisingly well-performing sections even with newly demanded shorter headways.

Formulation of adherence indices with respect to scheduled headway:

- $SA_{schedule}$ for a single observation $(SASO_{schedule})$ = Observed Headway – Scheduled Headway.
- Absolute $SA_{schedule}$ for a single observation (ABS $SASO_{schedule}$) = |Observed Headway − Scheduled Headway|.
- $SA_{schedule}$ for a bus stop = Average of all $SASO_{schedule}$ values at a stop.
- Absolute $SA_{schedule}$ for a bus stop = Average of all Absolute $SASO_{schedule}$ values at a stop.
- $SA_{schedule}$ of a day = Average of all $SASO_{schedule}$ values at all stops of all day.

- Absolute $SA_{schedule}$ of a day = Average of all Absolute $SASO_{schedule}$ values at all stops of all day.

Formulation of adherence indices with respect to desired headway set

- SA_x for a single observation $(SASO_x)$ = Observed Headway − Passenger Desired Headway, X.
- Absolute SA_x for a single observation (ABS $SASO_x$) = |Observed Headway − X|.
- SA_x for a bus stop = Average of all $SASO_x$ values at a stop.
- Absolute SA_x for a bus stop = Average of all Absolute $SISO_x$ values at a stop.
- SA_x of a day = Average of all $SASO_x$ values at all stops of all day.
- Absolute SA_x of a day = Average of all Absolute $SASO_x$ values at all stops of all day.

9.1 Adherence Results

Figure 9 shows schedule adherence-related indices for a same weekday. The top two graphs show the schedule adherence indices computed against the scheduled headways. The first graph is based on the summation of $SA_{schedule}$ which adds the positives and negatives, sometimes cancelling out each other. The second graph is based on the $|SA_{schedule}|$ which penalizes both the early and late arrival of buses with respect to the scheduled headways. The $|SA_{schedule}|$ graph shows higher values with a high contrast, so that transit operators can easily identify the problematic locations and times.

The bottom 5 graphs in Fig. 9 show the adherence graphs computed against a set of varying headways arbitrarily defined that passengers may demand. Some studies in the past [15, 17, 18, 22] indicate that passenger arrival patterns switch from random arrivals to an organized pattern when the headways are larger than 10~12 min. The bottom 5 graphs are intentionally designed to include the 10~12 min scenario. The intention of these measures is to test the current transit operations, quantifying how it would perform when passengers demand a certain headway threshold as a minimum acceptable level of service. As intuitively expected, the adherence indices increase (more red and green cells) as passengers demand for shorter headways. In the "With Set Headway" section of Fig. 9, "8 min" scenario represents a case where passengers demand 8 min or shorter headway to be acceptable. However, because there are more red/green cells than other scenarios, the current operating condition is not satisfactory on certain sections at certain time slots. In this case, the transit operator should assign more buses in this route to increase the level of service. In the other extreme case, the "16 min" scenario

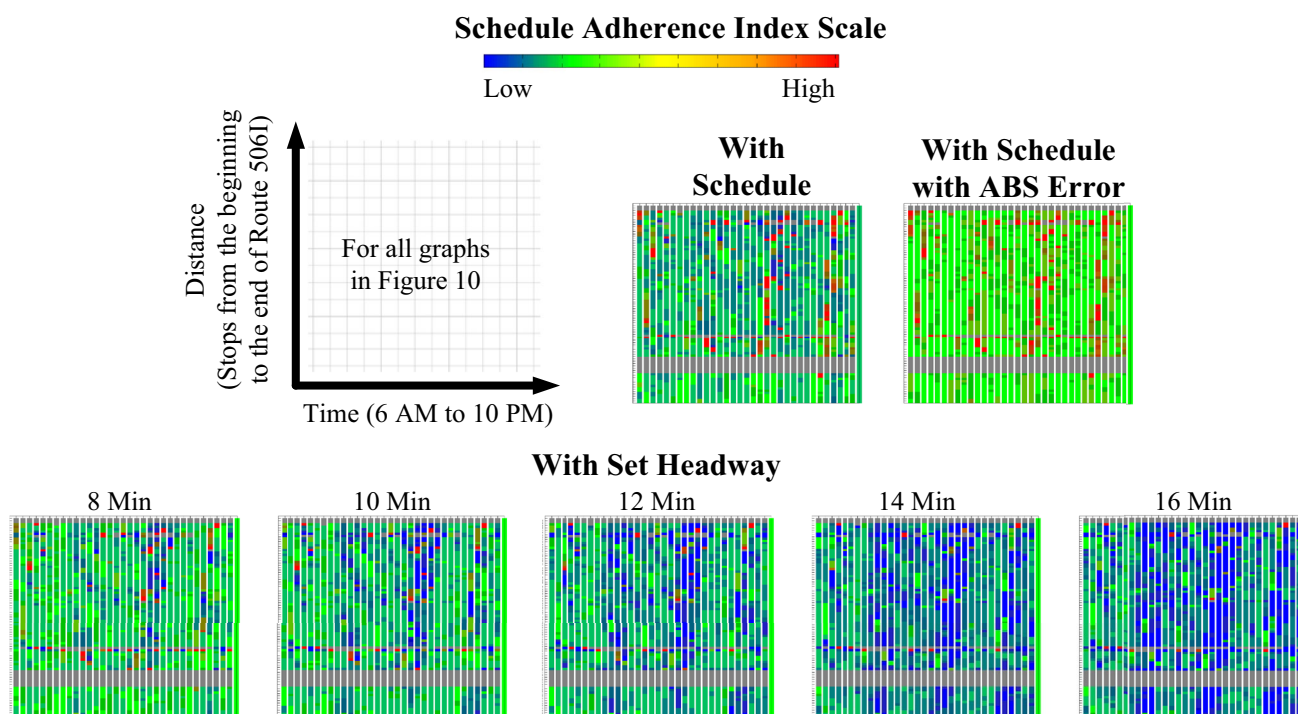## Schedule Adherence Index Scale

Fig. 9 Schedule adherence indices for a typical weekday

has more cells with blue which means that the current system is generally satisfying the passengers' demand. In this case, the transit operator can pull out some fleets from the route to reduce costs. Figure 9 is a collection of zoomed out graphs for a quick visual comparison purpose only. The developed and implemented code displays each graph in high resolutions that are sufficient for quick visual inspections of the current state of bus operations.

10 Dedicated Bus Lane

The chosen route 506I does have a 1.5 km of relatively very short section of dedicated bus lane. After visually analyzing the results at the bus stops in the vicinity of the dedicated bus lane, it is found that the indices did not significantly behave differently.

Provided that we have access to the significantly long section of dedicated bus lanes, our initial hypothesis would be that dedicated bus lanes would result in a better adherence to scheduled headways. However, it is not guaranteed that such trend will actually be observed, since the transit performance degrades over the route in a cumulative fashion. For example, if a bus is delayed heavily by congestions in an upstream of a section with dedicated bus lanes, the bus would still arrive off schedule, yet possibly operate with constant headways that are shifted from the scheduled arrivals (if Transantiago adopts scheduled arrival policy in

the future) at bus stops. Until a longer section of dedicated bus lane becomes available, it is not yet possible to conduct such analyses.

11 Conclusions

With newly available massive GPS data from transit buses operating in Santiago, Chile, this paper develops various transit performance measures useful for transit operators and planners.

- In the first phase, the running times of buses are extracted from the GPS data. Then, the distribution of headways on a weekday and a weekend day for different time slots of the day is found.
- In the second phase, a new index that captures the increase in expected passenger waiting times due to the bunching phenomena is developed and their respective tempo-spatial graphs are produced for a weekday as well as for a weekend day.
- In the final phase of this paper, it is recognized that the adherence to scheduled headway is slightly different from bunching in a sense that bunching "causes" deviations of observed headways from the scheduled headways. Therefore, a schedule adherence monitoring method is presented along with a stress-testing case study for a varying headway demand from the public.

- It is found that equipping transit vehicles with GPS devices can provide enormous amount of data that are sufficient for continuously monitoring various transit performance measures throughout the route daily.

12 Future Works

In this paper, public transit buses installed with GPS devices are used to collect data on a typical week and weekend days for monitoring bunching and headway adherences. This paper is unique and contributes to the field of study as a case study in a public transit system with GPS bus probes in a capital city in Latin America (Chile). This initial study can be extended to more in-depth analyses as the database size expands with continued GPS data collection efforts with Transantiago.

In the near future, when multiple years' worth of data will be available, it is likely that we will eventually have significant number of unusual week and weekend days that include the days with natural distractions, such as rainfall, snowfall, and minor earthquakes (as earthquakes are quite common in Chile). It will be interesting to see how public transit performances react to such phenomena.

This paper develops mainly from observing collected GPS data and analyzing them in the dimensions of bus stops (distance) and time. However, each vehicle (bus) has capacity limitations and each bus stop is associated with typical passenger arrival patterns. If a particular bus stop attracts significantly more passengers than others, there is a higher chance that the next arriving bus will not be able to board all passengers at the stop forcing left-over passengers to wait longer for following successive buses. Procedures developed in this paper would only capture abnormal incidents in public transit operations and alert transit operators. However, it is also important to develop mechanisms to resolve such issues. If observed indices indicate certain bus stops consistently performing poorly, the bus route itself, location, and the number of bus stops before and after the problematic stops can be modified. Integrating automatic passenger counters (APC) would also help providing transit load information to further aid with monitoring bus capacity limitations and related level of service measures.

Dedicated bus lanes would help achieving higher schedule adherences to scheduled headways or arrivals, in general. Constructing new or newly assigning existing lanes as bus dedicated lanes would result in enormous financial and social costs, and there is a research need for a careful consideration of practical potential benefits per dollar invested in such mega construction projects. It is possible to conduct a simulation-based study in the near future that may show reductions in travel times of passengers with respect to capital costs of such projects. It is also a research challenge to consider and model dynamic assignments of certain roads as timed dedicated bus lanes. Such efforts may maximize the utility of the existing road networks for both automobile and transit users.

References

1. Byon YJ, Shalaby A, Abdulhai B (2006) GISTT: GPS-GIS integrated system for travel time surveys. TRB 85th Annual Meeting Compendium of Papers CD-ROM. Transportation Research Board of the National Academies, Washington, D.C, p 16
2. Herrera JC, Work DB, Herring R, Ban X, Jacobson Q, Bayen AM (2010) Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. Trans Res Part C 18:568–583. doi:10.1016/j.trc.2009.10.006
3. Work DB, Tossavainen O, Blandin S, Bayen AM, Iwuchukwu T, Tracton K (2008) An ensemble Kalman filtering approach to highway traffic estimation using GPS enabled mobile devices. 47th IEEE Conference on Decision and Control. 5062–5068. doi:10.1109/CDC.2008.4739016
4. Byon YJ, Abdulhai B, Shalaby A (2009) Real-time transportation mode detection via tracking global positioning system mobile devices. J Intel Trans Sys 13(4):161–170. doi:10.1080/15472450903287781
5. Byon YJ, Liang S (2014) Real-time transportation mode detection using smartphones and artificial neural networks: performance comparisons between smartphones and conventional global positioning system sensors. J Intel Trans Sys 18(3):264–272. doi:10.1080/15472450.2013.824762
6. Kwon J, Petty K, Varaiya P (2012) Probe vehicle runs or loop detectors? Effect of detector spacing and sample size on the accuracy of freeway congestion monitoring. J Trans Res Rec 2012:57–63. doi:10.3141/2012-07
7. Byon YJ, Shalaby A, Abdulhai B (2010) Traffic data fusion using SCAAT Kalman Filters. TRB 89th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C, p 16
8. Cathey FW, Dailey DJ (2002) Transit vehicles as traffic probe sensors. J Trans Res Rec 1804:23–30. doi:10.3141/1804-04
9. Kumar SV, Vanajakshi L, Subramanian SC (2011) A model based approach to predict stream travel time using public transit as probes. 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, Germany, pp 101–106. doi:10.1109/IVS.2011.5940413
10. Song MK, Li M, Nakamura H (2010) Real-time Arterial performance measurement using signal timing data and brt as probes. TRB 89th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C. p 17
11. Agarwal A, Goel N (2009) GPS-RDS enabled location based smart transit. ICIN 2009. 13th International Conference on Intelligence in Next Generation Networks, Bordeaux, France, pp 1–5. doi:10.1109/ICIN.2009.5357105

12. Ledwitz P (1992) GPS technology applied to public transit systems. IEEE 42nd Vehicular Technology Conference, Denver, USA. 1:83–84. doi:10.1109/VETEC.1992.245391

13. Wipke KB (1996) Reducing VMTs through transit-on-demand with GPS and satellite communications. IEEE Technical Applications Conference Northcon/96, Seattle, USA. pp 404–408. doi:10.1109/NORTHC.1996.564971

14. Munizaga MC, Cortés E, Gibson J, Gschwender A, Zúñiga M (2010) Bus commercial speed diagnosis based on GPS Monitored data. TRB 89th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C., p 16

15. Liao CF, Liu HX (2010) Development of a Data Processing Framework for Transit Performance Analysis. TRB 89th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C., p 17

16. Daganzo CF, Pilachowski J (2011) Reducing bunching with bus-to-bus cooperation. Trans Res Part B Metho 45(1):267–277. doi:10.1016/j.trb.2010.06.005

17. El-Geneidy AM, Horning J, Krizek KJ (2008) Analyzing transit service reliability using detailed data from automatic vehicular locator systems. TRB 87th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C., pp 18

18. Chen X, Yu L, Zhang Y, Guo J (2009) Analyzing urban bus service reliability at the stop, route and network levels. Trans Res Part A Pol Prac 43(8):722–734. doi:10.1016/j.tra.2009.07.006

19. Owais M, Hassan T (2016) Incorporating dynamic bus stop simulation into static transit assignment models. Int J Civ Eng. doi:10.1007/s40999-016-0064-8

20. Jamili A (2016) A mathematical model and a hybrid algorithm for robust periodic single-track train-scheduling problem. Int J Civ Eng. doi:10.1007/s40999-016-0089-z

21. Tamannaei M, Saffarzadeh M, Jamili A et al (2016) A novel train rescheduling approach in double-track railways: optimization model and solution method based on simulated annealing algorithm. Int J Civ Eng 14(3):139–150. doi:10.1007/s40999-016-0002-9

22. Milkovits MN, Wilson NHM (2010) Simulating bus service reliability: another use for automatically collected data. TRB 89th Annual Meeting Compendium of Papers DVD. Transportation Research Board of the National Academies, Washington, D.C., p 17

23. Furth PG, Muller THJ (2007) Service reliability and optimal running time schedules. J Trans Res Rec 2034:55–61. doi:10.3141/2034-07

24. Holroyd EM, Scraggs DA (1966) Waiting times for buses in Central London. Traf Eng Cont 8(3):158–160

25. Welding PI (1957) The instability of a close-interval service. Oper Res Q 8(3):133–148. doi:10.2307/3007157

26. Fan W, Machemehl RB (2009) Do Transit users just wait for buses or wait with strategies? Some numerical results that transit planners should see. Trans Res Rec 2111:169–176. doi:10.3141/2111-19

27. Li S, Zhu K, van Gelder BHW, Nagle J, Tuttle C (2002) Reconsideration of sample size requirements for field traffic data collection with global positioning system devices. Trans Res Rec 1804:17–22. doi:10.3141/1804-03

28. Smith BL, Zhang H, Fontaine M, Green M (2003) Cellphone Probes as an ATMS Tool. Research Project Report for the National ITS Implementation Research Center and US. University Transportation Center, Center for Transportation Studies at the University of Virginia

29. Hakkesteegt P, Muller T (1981) Research increasing regularity. Verkeerskundige werkdagen, pp 415–436