Modeling spatial-temporal patterns of bus delays at and between stops using AVL and APC data and semi-Markov techniques

YING SONG[1*]

[1] (Corresponding Author) Ph.D., Assistant Professor;
Geography, Environment and Society, University of Minnesota – Twin Cities;
478 Social Science Building, 267 19th Ave S, Minneapolis, MN, USA, 55455;
PH (612) 625-2056; FAX (612) 625-6080; E-Mail: yingsong@umn.edu

ABSTRACT

Most transit operators produce automatic vehicle location (AVL) and automatic passenger counter (APC) data to assess and improve transit performance. Conventional analyses usually extract data at stops, and therefore underutilize vehicle location data collected between stops. This paper develops a model to study bus stopping patterns using AVL and APC data collected at and between stops. The model contains three major steps: (1) linear-referencing AVL and APC data along transit routes, (2) visually exploring the spatio-temporal patterns of delays, and (3) modeling vehicle movements as continuous-time Semi-Markov processes and calibrating them using the revealed patterns. The model can be used to identify locations and times that are more likely to get congested and lead to delays, and provide more accurate arrival times to transit users. To demonstrate the model, the paper uses AVL and APC data collected along A-Line rapid route within 8-day period in Minneapolis, Minnesota, USA.

INTRODUCTION

Recent years have witnessed a paradigm shift toward sustainable transportation planning and practice (Banister 2008). Among various strategies, studies have proved that reliable transit services can reduce automobile dependency while providing access to various resources and opportunities (e.g. Handy 2002, Litman, 2017, Shatu and Kamruzzaman 2014). To assess transit performance and improve transit services, most transit authorities have produced and maintained automatic vehicle location (AVL) and automatic passenger counter (APC) data. However, the sampling frequency of AVL used to be low (e.g. 5 minutes), and analyses usually extract data only at the stop level (e.g. Cathey and Dailey 2003, Crout 2007, Farber et al. 2014, Shen et al. 2016).

>recent increase in temporal resolution

>more precise/ accurate location resolution

Advances in location technologies have significantly improved the temporal resolution and the spatial precision of vehicle location data. The sampling frequency of these data has been increasing, especially over the past few years (from 1 minute to 5 seconds for AVL). And the progress in Global Positioning System (GPS) enables us to collect more accurate and precise vehicle location data (from 25 meters in early 1970 to 5 meters now for AVL) (Okunieff 1997). Hence, a few studies have started to utilize high resolution bus location data (AVL and APC) to analyze congestion patterns and travel time reliability along urban arterials (e.g. An et al. 2016, Feng, et al. 2015, Glick and Figliozzi, 2017, Yang, et al. 2017). These studies also contribute to a broader theme on gathering, cleaning, visualizing, and analyzing spatial-temporal data to understand manage and predict traffic flows in the urban environment (e.g. Ng, Reaz, and Ali 2013, Yue, Rilett, and Revesz. 2016, Zhao et al. 2017).

This paper continues this research trajectory on utilizing AVL and APC data collected at and between bus stops to investigate travel time, congestion hotspots, and service reliability. Compared to existing studies, this paper contributes to this research topic in three aspects. First, instead of aligning discrete vehicle GPS coordinates stored in AVL to network links (road segments), the paper applies the dynamic segmentation method in Geographic Information Science (GIS) to linear-reference each sequence of GPS coordinates to its corresponding route and eliminate possible artifact movement patterns resulted from inaccurate/imprecise location data. Second, the paper visualizes the linear-referenced movement trajectories and explores spatio-temporal patterns of delays along the bus route. Third, the paper models bus movements among stops along a route as a continuous-time Semi-Markov process, and uses the revealed delay patterns as references to calibrate the model. The calibrated model can estimate arrival times at bus stops that are sensitive to the congestion levels along the route. To demonstrate the model, the paper uses AVL and APC data collected within 8-day period along A-Line rapid bus route in Minneapolis, Minnesota, USA. The paper discusses selected results, and concludes with a discussion of potential applications and future research.

METHOD

Model Schema

This paper develops a model to study bus stopping patterns using AVL and APC data collected at and between stops. Figure 1 shows the model schema. The inputs are the original AVL and APC dataset managed by transit authorities, and route and stop locations. The processes highlight the integration of space and time in the model, and its ability to capture spatial-temporal patterns of delays while estimating arrival times.

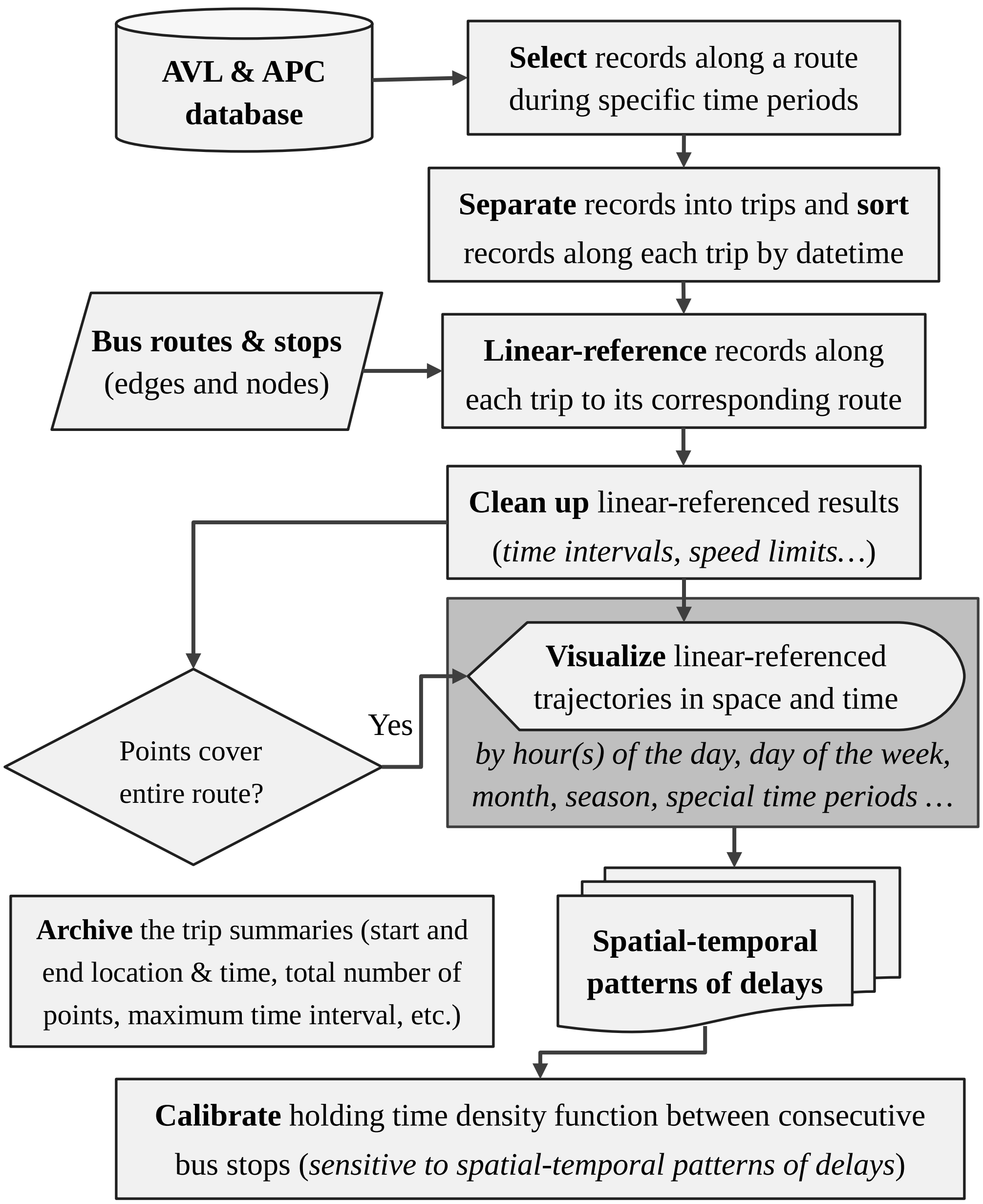


Figure 1. Processes to explore and model bus delays at and between stops.

The two core model components are (1) *linear referencing* the recorded vehicle locations to the corresponding routes, and (2) modeling vehicle movements between bus stops along the route as *continuous-time semi-Markov process*. The following two sections introduces the two core concepts and their application to AVL and APC dataset.

Linear-Referencing and Dynamic Segmentation

In transportation studies, linear referencing systems have been widely applied to record incidences and conditions along highways (Dueker and Vrana 1992). The incidents and lengths are represented as point and linear features along a highway route. Their locations can later be retrieved by identifying the route and the replacement along that route. For instance, the St. Anthony Falls Bridge over the Mississippi River can be represented as Interstate 35W, MP 18.374–18.736. For network beyond highways, the Topologically Integrated Geographic Encoding and Referencing (TIGER) products of the U.S. Census Bureau include digital representations of transportation networks using the edge/node schema (Coppock and Rhind, 1991). This enables us to expand the linear referencing systems for highway routes to arterial roads and local streets. In GIS, these are also known as dynamic segmentations (Goodchild 1998). Any geographic location, therefore, corresponds to a relative position along an encoded road segment. The paper will adopt the dynamic segmentation method and linear-reference an ordered sequence of vehicle locations recorded in the AVL and APC data along their corresponding route.

The bus routes are the basis of the linear referencing system. The displacement of a location along a bus route is measured as its distance from the first stop along the route. Stops along each bus route are also linear-referenced using this method, so each stop is represented as (*route ID, stop ID, stop sequence, distance from the first stop*). Note that the distances here use linear units such as meters and miles, instead of travel time along road segments such as minutes and seconds.

After initially linear-referencing vehicle locations to the corresponding route, the paper further processes the results considering: (1) the maximum achievable speeds or speed limits (e.g. no more than 100 mph), (2) the moving direction (e.g. no frequently moving back and forth at bus stops), and (3) possible device failure (e.g. time interval between two consecutive recorded GPS locations not larger than 10 minutes). The final linear-referencing result for any trip is an ordered sequence of points along that route, each represented as (*route ID, point sequence, datetime, distance from the first stop*).

To prepare for calibrating the holding time density function, the paper also adds travel time between two bus stops and two recorded GPS locations. For bus stops, the travel time indicates the minimum time to move from one stop to the next stop along the route given the length and speed limit of edges between these two stops. For recorded GPS locations, the travel time indicates the travel time between two locations.

Semi-Markov Processes

A Markov process is a "memoryless" stochastic process whose future behavior is conditioned on its present status, and independent of its past history (Paul et al. 1972). A *semi-Markov process* is a stochastic process with time delays (holding times) before reaching to the consecutive status. For semi-Markov processes evolving continuously in time, two components to model the processes are (1) an imbedded Markov process with transition probabilities $p_{ij}(t)$, and (2) a holding time density function $h_{ij}(\tau)$ that describes the probability that one transition from i to j takes τ amount of time (Howard 2012). In theory, $h_{ij}(\tau)$ can be any probability density function (PDF) that is right-continuous such as the exponential distribution and log-normal distribution.

Song et al. (2016) adopt and modify the semi-Markov techniques to model individuals' movements within transportation networks. They model the network as a directed graph $G = (V, E)$ with nodes $V = \{v_1, v_2, v_3, \ldots\}$ and links $E = \{e_{ij}: (v_i, v_j)\}$. To account for the minimum travel time t_{ij} calculated based on the length and speed limit of link e_{ij}, they modify the holding time density function $h_{ij}(\tau)$ as $h_{ij}(\tau - t_{ij})$. They calibrate $h_{ij}(\tau - t_{ij})$ using empirical vehicle trajectories collected in New York City, USA, and find that it fits best with exponential distributions. Song et al. (2017) prove that $h_{ij}(\tau - t_{ij})$ in different study areas may fit best with different distributions: the vehicle trajectories collected in Phoenix, Arizona, USA show that the holding times in Phoenix follow lognormal distribution instead of exponential distribution.

The paper models bus movements between stops along a route as a continuous-time semi-Markov process. It represents the transit system as a directed graph *G* with bus stops as a set of vertices *V*, headways with direction as a set of directed edge *E*, and the distance from one stop to the next by minimum travel time $t_{ij} = t_{SN}(v_i, v_{i+1})$. A state for the semi-Markov process corresponds to a movement from stop v_i to next stop v_{i+1} along the headway e_{ij}. The state space, therefore, is movements between any two consecutive stops along a route, which is finite in nature.

For the holding time density function, the paper adopts the method by Song et al. (2016) and use extra time spent (delays) between stops as the variable $h_{ij}(\tau - t_{ij})$. Instead of assuming the holding times follow some right-continuous distribution, the paper uses the empirical data to find the best fit distribution(s) using methods such as Kolmogorov-Smirnov-Test (KS-Test) (Young 1977) and tests for skewness, kurtosis, and normality (Bai and Ng 2005). This enables us to capture the spatial-temporal variations of delays, including the different patterns along the same headway across various time periods and the different patterns along different headways during the same time periods. To get the empirical arrival time at each bus stop (or the time passing it), the paper first extracts vehicle locations within a certain distance from each stop (e.g. 30 meters), and then checks the average speed from its previous recorded location.

IMPLEMENTATION AND RESULTS

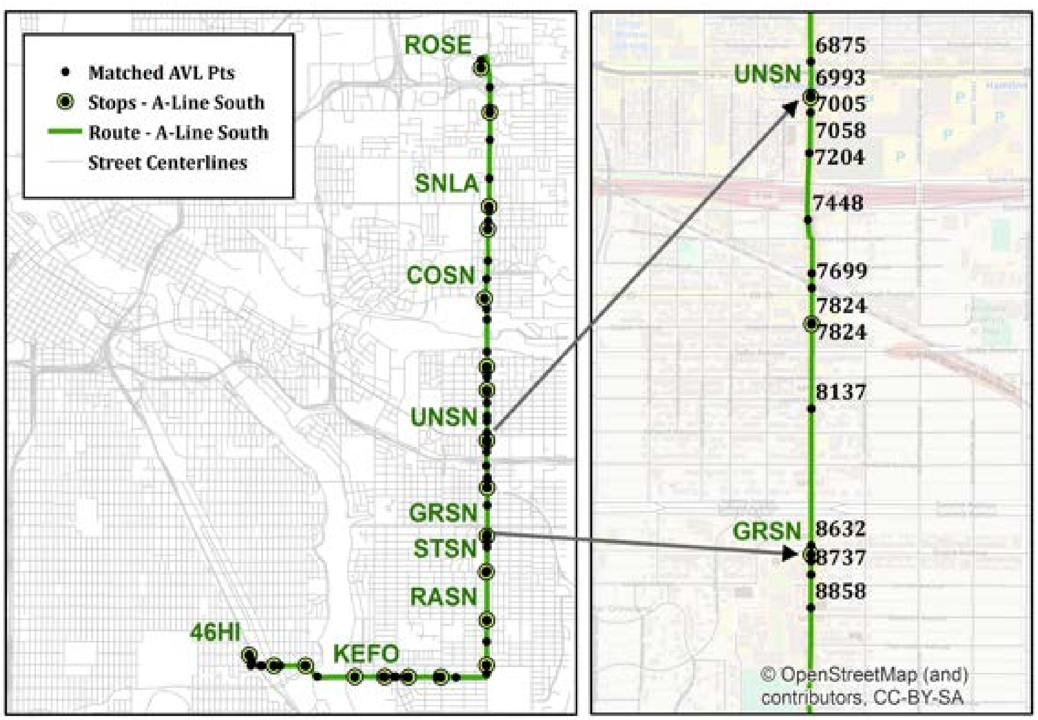
Software Tools and Study Area

The paper implements the methodology using two programming languages: Python and R. For Python, the paper uses two modules: (1) Numpy for creating and managing large-size arrays, and (2) ArcPy in ESRI ArcGIS software for conducting network analysis and visualizing results. For R, the paper uses the package "fitdistrplus" to get the best fit univariate parametric distribution(s) to the empirical holding times.

To demonstrate the methods and processes, the paper uses messages for the A line rapid service route (https://www.metrotransit.org/Route/921) from 2016-10-01 to 2016-10-08. This dataset is provided by Metro Transit Authority, the primary public transportation operator in the Minneapolis – St. Paul area (Twin Cities), Minnesota, USA. The data represents a sampling of normal service within a single version of the schedule. There are twelve unique vehicles assigned to this route, each equipped with a radio transmitter. The total number of messages is 126,883 in the original dataset, which are divided to 1,625 trips.

Results and Discussion

Figure 2(a) shows with the linear-referend stops along the route, and Figure 2(b) shows part of a linear-referenced trip, labeled with distances (meter) from the first stop.



(a) all Points along the trip (b) selected points with distance

Figure 2. Linear Referenced AVL data for one trip along A-Line Southbound

Figure 3 shows visualization results for all trips along A-Line northbound. Since these trips have different departure times from the first stop, the paper uses the time leaving the first stop as the start time of each trip (marked as “0 minute” along x-axis). The results indicate that: (1) Monday, Tuesday and Sunday have relatively more reliable services with less delays and delay variations, (2) the most significant delays locate near 8,000-10,000 meters along the route between stops around Macalester College, especially for Wednesday, and (3) Friday and Saturday tend to have delays in the later part of the trip toward the north, while Thursday has delays more evenly distributed along the trips.

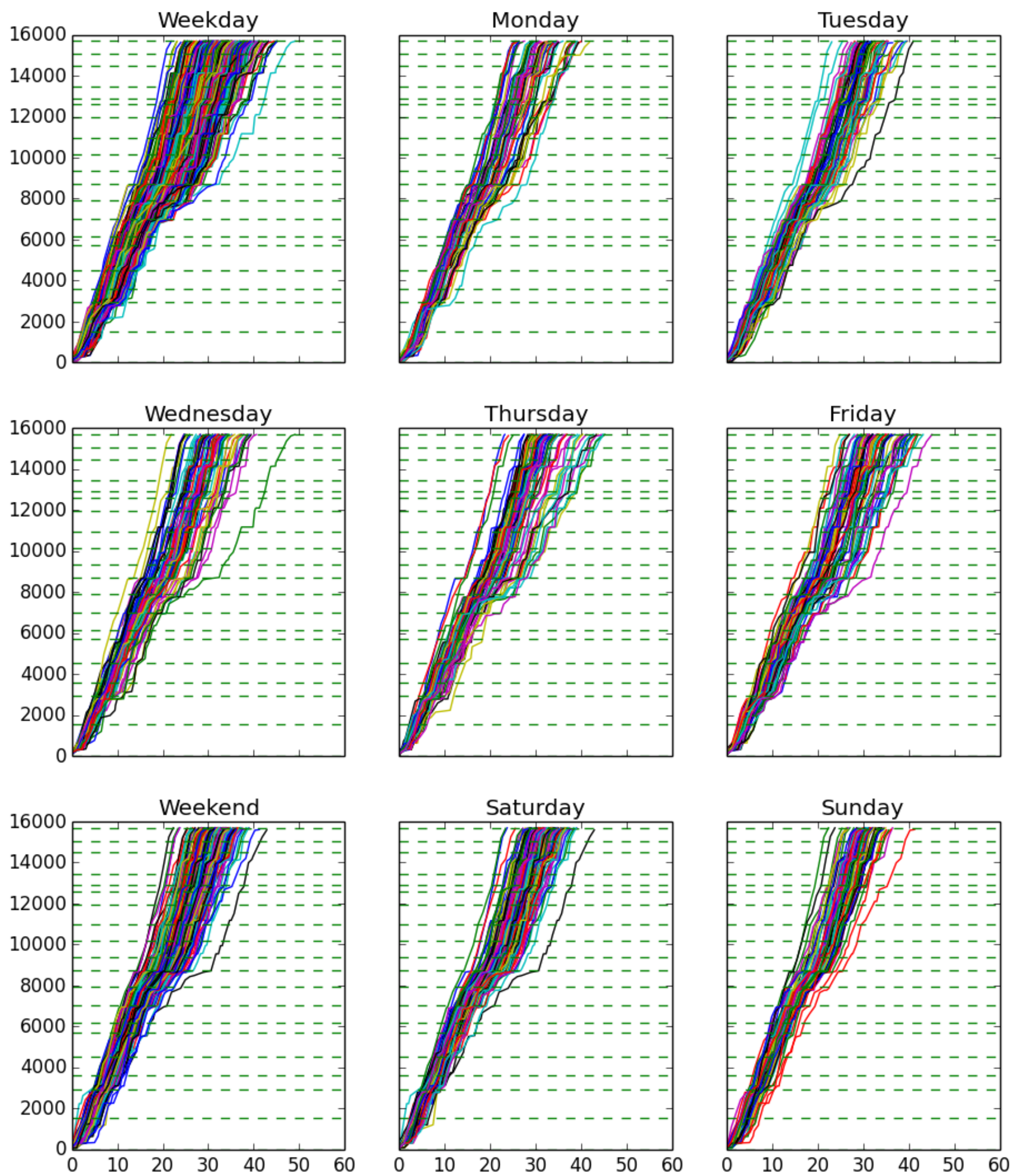


Figure 3. Linear-reference trips by days of the week.

Figure 4 shows visualization results for trips along A-Line northbound during different time periods within a day. The delays from 3pm to 6pm are most significant, and are least significant from 9pm to 6am the following day. These are consistent with the afternoon peak hours, and low traffic volumes during late night and early morning. The delays from 9am to 12pm have a significantly large variation near 8,200 meters along the route, which is likely to caused by high transit ridership and/or general traffic near Macalester College during morning class hours and/or during lunch hours.

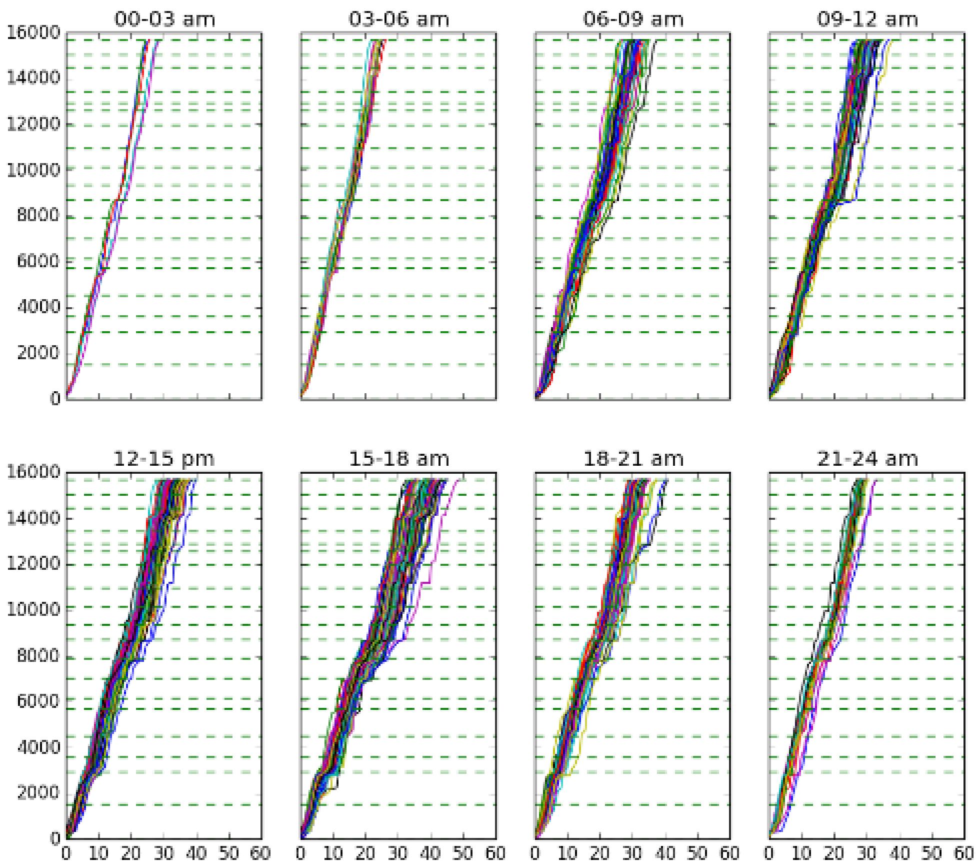


Figure 4. Linear-reference trips by time of the day.

Based on the delay patterns implied by the visualizations, the paper calibrates holding time density function $h_{ij}(\tau - t_{ij})$ from each stop to its next stop during certain time periods. First, the paper plots the kurtosis and squared skewness of the sampled delays at each stop, which is the extra travel time beyond minimum travel time from its previous stop along the route; and compare it to typical distributions including normal, uniform, exponential, logistic, beta, lognormal, and gamma distribution. Figure 5 shows results for all stops during weekdays. Although lognormal and gamma distribution appear to the best fit distributions for delays at all stops, the degree of fitness varies significant.

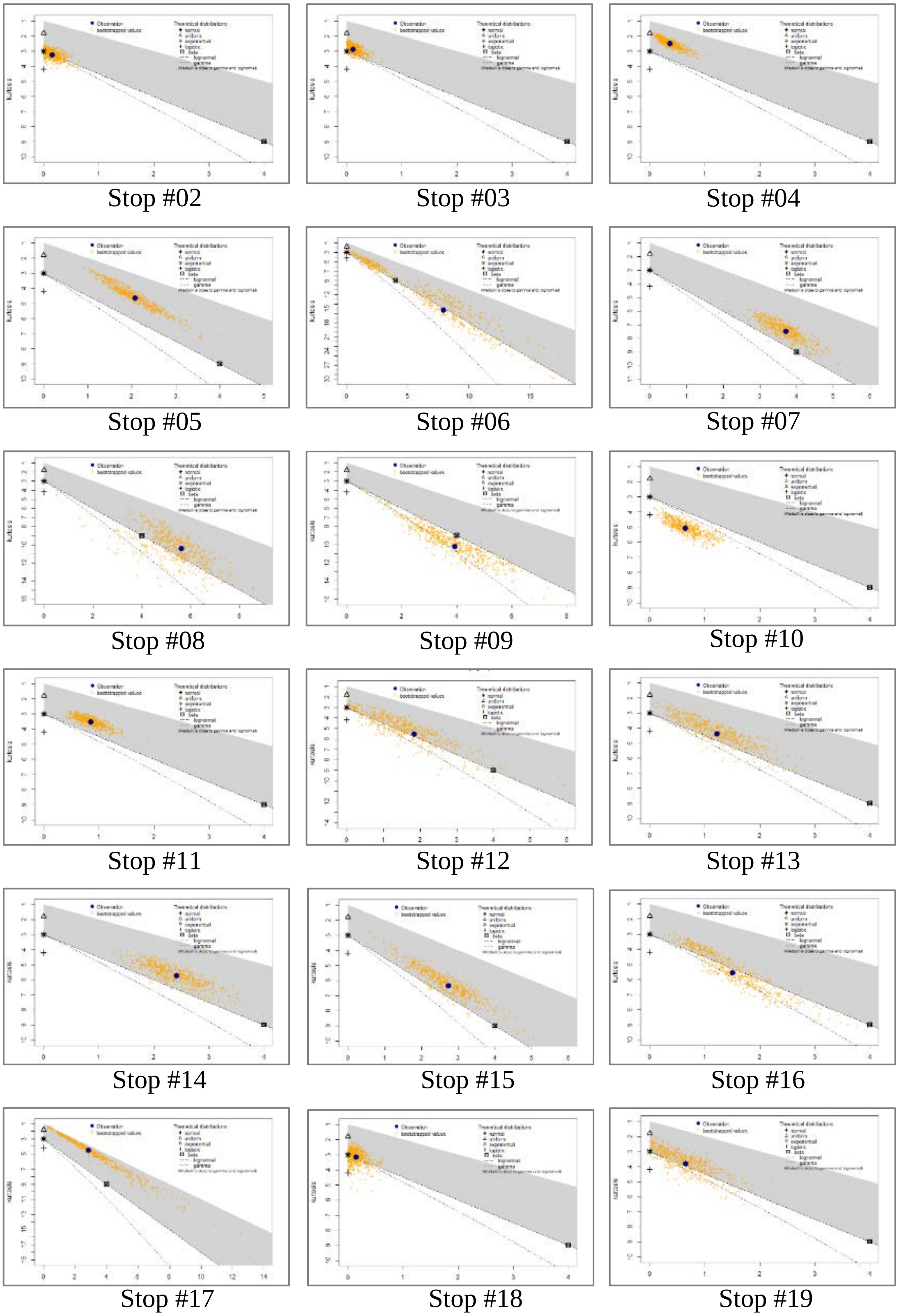


Figure 5. Fitting distribution based on descriptive characters using R.

Second, the paper calibrates the $h_{ij}(\tau - t_{ij})$ at selected stops and quantify the spatial-temporal variations of delays indicated by visual exploration results. The paper uses probability density (PDF), cumulative density (CDF), Q-Q plot and pp plot (in R) to compare the empirical distribution with the best fitting lognormal and gamma distributions (by maximum likelihood in R). Figure 6 shows results for stop 5, 11, and 16 along A-Line Northbound on Wednesday. These results support the previous visual explorations of delay patterns (see Figure 3). Delays from stop #4 to stop #5 are quite minor and seldom exceed 100 seconds. Delays from stop #9 to stop #10 always exist and quite a few of them are over 2 minutes. Delays from stop #15 to stop #16 are relative more evenly distributed. This prove our intuition that the delays are not evenly distribution throughout the trip, and it is necessary to use different holding time density functions to estimate arrival time at stops.

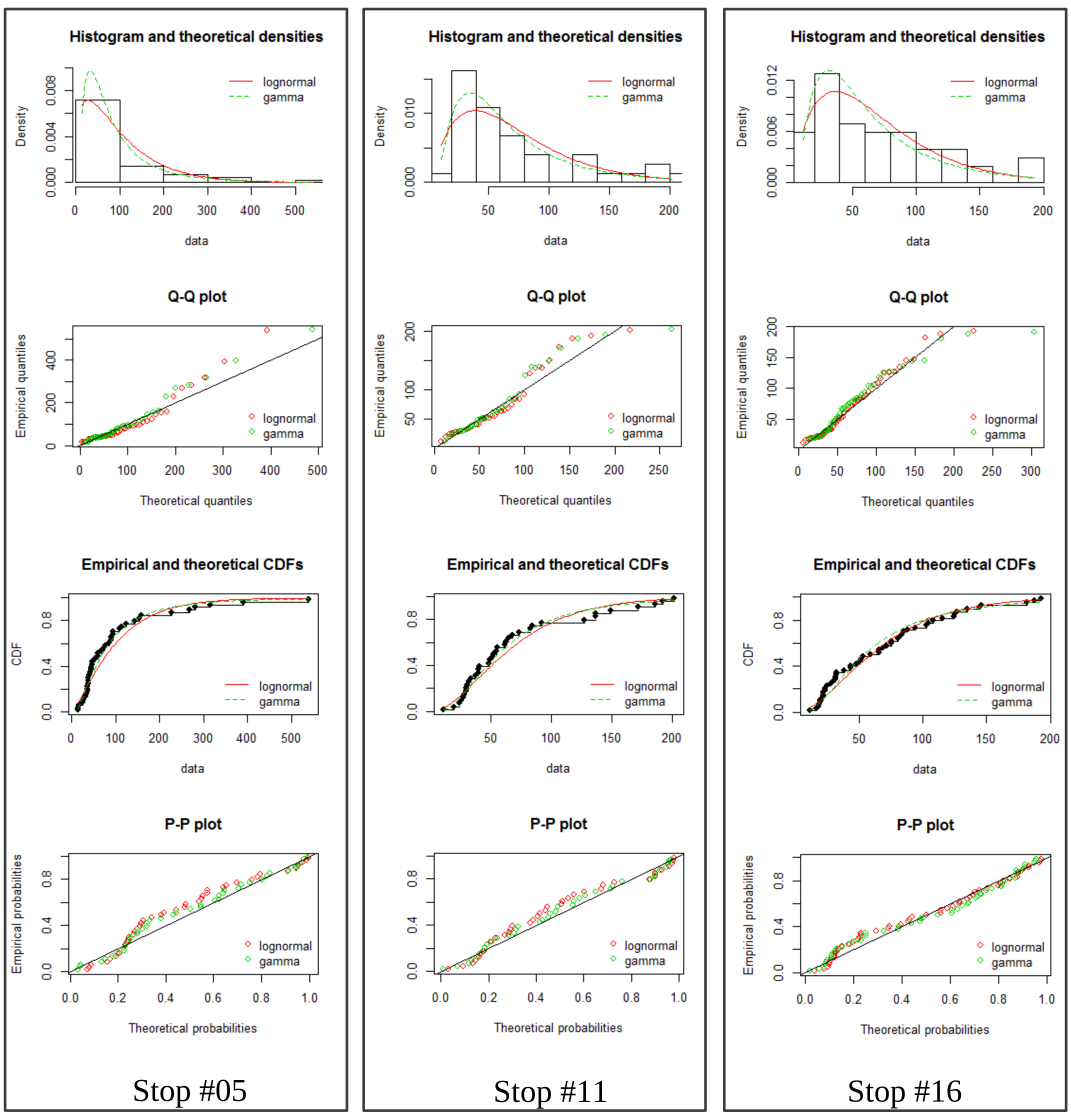


Figure 6. Comparing empirical and theoretical distributions using R.

CONCLUSION

This paper develops a model to process, visualize and analyze AVL and APC data and investigate spatial-temporal patterns of delays along bus routes. Compared to existing methods, the model considers vehicle locations as ordered placements along a bus route, and applies dynamic segmentation method in GIS to linear-reference vehicle locations to the route. Visualizations of the linear-referenced trips reveal that delays are not evenly distributed along the route, and the delay patterns vary with respect to time (e.g. hour of the day, day of the week). Then, the paper models bus movements along a trip as a continuous-time semi-Markov process among stops. The model calibrates the holding time density function (time delays) between stops using linear-referencing results, which can provide arrival time estimations that are sensitive to spatial-temporal patterns of delays along the route.

The paper provides new insights in bus delay patterns using AVL and APC data collected at and between stops. The findings could help transit operators and planners to identify locations and/or times that are more likely to get congested and lead to service delays. The transit operators can then use these findings to determine strategies that can effectively mitigate the delays and congestion costs. Using these delay patterns revealed by historical data, the transit operators can apply the semi-Markov techniques to provide more accurate estimations of arrival times to transit users.

The current work could be extended by obtaining additional data. Since A-Line route is a local rapid route with frequent services and short distances between stops, the revealed patterns may not apply to other types of service route with low frequent services and longer commute distances (e.g. commuter express service between urban and suburban areas through highways). Moreover, data collected in one-week may not be sufficient to represent the global patterns, and cannot detect patterns such as monthly and seasonal patterns (e.g. delay patterns during summer near Macalester College may be quite different from patterns revealed by this paper). The additional dataset can also allow validation of the methods. Given sufficient data, it is possible to divide the data into two sets: one for training that can calibrate the holding time density function, and the other as “observed pattern” to be compared to the calibrated functions for validation.

Another extension could be integrating the AVL and APC data with GTFS data, and aligning linear-referenced trips to the GTFS scheduled trips. The AVL data used in this paper has an average sampling frequency of 60 seconds, and buses along express bus routes like A-Line are loosely restricted by scheduled times at time stops. Therefore, the matching results are not optimistic and have been excluded from this paper. When more frequently sampled data become available, the future research can refine the empirical delays by considering the scheduled arrival times and compare the analysis and modeling results.

REFERENCE

An, S., Yang, H., Wang, J., Cui, N., and Cui, J. (2016). Mining urban recurrent congestion evolution patterns from GPS-equipped vehicle mobility data. *Information Sciences*, 373, 515-526.

Bai, J., and Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business and Economic Statistics*, 23(1), 49-60.

Banister, D. (2008). The sustainable mobility paradigm. *Transport policy*, 15(2), 73–80.

Cathey, F. W. and Dailey, D. J. (2003). A prescription for transit arrival/departure prediction using automatic vehicle location data. *Transportation Research Part C: Emerging Technologies*, 11(3), 241-264.

Coppock, J.T., and Rhind D.W. (1991) The history of GIS. In Maguire, D.W., Goodchild, M.F., and Rhind, D.W. edited *Geographical Information Systems: Principles and Applications*. London: Longman Scientific and Technical, 1, 21–43.

Crout, D. (2007). Accuracy and precision of the transit tracker system. *Transportation Research Record: Journal of the Transportation Research Board*, 1992, 93-100.

Dueker, K.J., and Vrana, R. (1992) Dynamic segmentation revisited: a milepoint linear referencing system. Journal of the Urban and Regional Information Systems Association 4(2): 94–105.

Farber, S., Morang, M. Z., and Widener, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. *Applied Geography*, 53, 149-159.

Feng, W., Figliozzi, M., and Bertini, R. L. (2015). Quantifying the joint impacts of stop locations, signalized intersections, and traffic conditions on bus travel time. *Public Transport*, 7(3), 391-408.

Glick, T. B., and Figliozzi, M. A. (2017). Novel methodology to estimate traffic and transit travel time reliability indices and confidence intervals at a corridor and segment level, *Proceedings in the 96th Annual Meeting of the Transportation Research Board, 2017*, No. 17-06136.

Goodchild, M.F. (1998) Geographic information systems and disaggregate transportation modeling. *Geographical Systems*. 5(1–2): 19–44

Handy, S. L. (2002). Accessibility-vs. mobility-enhancing strategies for addressing automobile dependence in the US. Paper presented at the ECMT round table on Transport and spatial Policies: The Role of Regulatory and Fiscal Incentives, RT124, Paris, November.

Howard, R. A. (2012). *Dynamic Probabilistic Systems, Volume II: Semi-Markov and Decision Processes*. John Wiley & Sons, Inc.

Litman, T. (2017). *Evaluating public transit benefits and costs: best practices guidebook*. Unpublished Report, Victoria, BC, Canada: Victoria Transport

Policy Institute. July 18, 2017

Metro Transit, Twin Cities, Minnesota, USA (accessed on Jan. 2018) *A Line Roseville-St Paul-Minneapolis*, https://www.metrotransit.org/Route/921

Ng, K. M., Reaz, M. B. I., and Ali, M. A. M. (2013). A review on the applications of Petri nets in modeling, analysis, and control of urban traffic. *IEEE Transactions on Intelligent Transportation Systems*, 14(2), 858-870.

Okunieff, P. E. (1997). AVL systems for bus transit, *TCRP Synthesis 24*, *Transportation Research Board*, National Academy Press, Washington, D.C.

Paul G. H., Sidney, C. P., and Charles, J. S. (1972). *Introduction to stochastic processes*. Boston: Houghton Mifflin.

Shatu, F. and Kamruzzaman, M. (2014). Investigating the link between transit oriented development and sustainable travel behaviour in Brisbane: a case control study. *Journal of Sustainable Development*, 7 (4), 61–70.

Shen, Y., Xu, J., and Zeng, Z. (2016). Public transit planning and scheduling based on AVL data in China. *International Transactions in Operational Research*, 23(6), 1089-1111.

Song, Y., Miller, H. J., Zhou, X., and Proffitt, D. (2016) Modeling visit probabilities within network‐time prisms using markov techniques. *Geographical Analysis*, 48(1), 18-42.

Song, Y., Miller, H. J., Stempihar, J., and Zhou, X. (2017). Green accessibility: Estimating the environmental costs of network-time prisms for sustainable transportation planning. Journal of Transport Geography, 64, 109-119.

Yang, Y., Gerstle, D., Widhalm, P., Bauer, D., and Gonzalez, M. (2013). Potential of low-frequency automated vehicle location data for monitoring and control of bus performance. *Transportation Research Record: Journal of the Transportation Research Board*, 2351, 54-64.

Yang, Y., et al. (2017). Efficient traffic congestion estimation using multiple spatio-temporal properties. *Neurocomputing*. 267(6), 344-353.

Young, I. T. (1977). Proof without prejudice: use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources. *Journal of Histochemistry and Cytochemistry*, 25(7), 935-941.

Yue, H., Rilett, L. R., and Revesz, P. Z. (2016). Spatio-temporal traffic video data archiving and retrieval system. *GeoInformatica*, 20(1), 59-94.

Zhao, H., Sun, D., Yue, H., Zhao, M., and Cheng, S. (2017). Using CSTPNs to model traffic control CPS. IET Software. 11(3), 116-125.