Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-3 (2016) 209–214

Holding Control of Bus Bunching without Explicit Service Headways

Lucas Zimmermann * **Werner Kraus, Jr.** *
Luiz Alberto Koehler ** **Eduardo Camponogara** *

* *Department of Automation and Systems Engineering, Federal University of Santa Catarina, Florianopolis SC, Brazil, (e-mail: zimmermann.lucas@posgrad.ufsc.br).*
** *Department of Electrical and Telecommunications Engineering, Regional University of Blumenau, Blumenau SC, Brazil, (e-mail: luiz@furb.br)*

Abstract: Holding-based control methods for bus operation are examined to point out that allowing greater variance in headways between consecutive buses leads to possible gains in total delay, as compared to strict adherence to a service headway. This result, obtained empirically, indicates that optimal operation is not necessarily attained with even headways. Such finding is related to the well-known fact that there should not be too many control points for headway corrections when operating under the traditional method of scheduled departures from bus stations. Current feedback and predictive methods, however, can be productively applied at all stations, hence the importance of studying the effects of frequent control actions. Several feedback schemes are tested, as well as a rolling horizon predictive control method that seeks to minimize onboard and at station delays. The latter has no headway reference and hence yields larger headway variations. The scenario is a BRT corridor modeled in a microsimulation environment. Simulation results indicate gains of 29% in total delay for predictive control in relation to open loop operation, and superior performance when compared to the tested proportional feedback control methods.

Keywords: Bus bunching, headway control, predictive control, feedback control, microsimulation analysis.

1. INTRODUCTION

It has long been known that bus operation is an inherently unstable process. Newell and Potts (1964) were the first to analyze the so-called bunching of buses when the operation is subject to common disturbances such as variable boarding/alighting times. Bunching causes deterioration of service in terms of irregular headways, variations of expected time of arrival, passenger load distributions, and passenger travel time (waiting plus riding time).

In order to regulate headways, one of the most used forms of control is holding a bus that is ahead of schedule (if operation is based on timetables) or closing in on the preceding bus (if service headways are specified). We concentrate on the latter case, which is common in the operation of high-frequency lines. As presented in reviews like Strathman et al. (2001), in this case passengers arrive at the stations independently; also, the least delay at stations is obtained with even headways. However, holding-only headway control will necessarily delay some buses, thus increasing aggregated onboard delay. As a result, holding should be used sparingly; in fact, the holding computation problem may involve both deciding on few control points where to act as well as by how much a bus should be held, as discussed by authors like Eberlein et al. (2001), Strathman et al. (2001), among others. Recently developed feedback control methods, however, do not deteriorate with holding actions at every station for all buses. For instance, Xuan et al. (2011) show that three different proportional feedback control laws are quite insensitive to the number of control points, while a more traditional fixed-schedule control suffers significantly from a large number of points where holding is applied.

In this paper, we revisit two classes of control methods to evaluate their performance regarding control objectives and improvements in quality of service indicators that are not directly controlled. The aim is to point out that allowing greater variance in headways between consecutive buses leads to possible gains in total delay as compared to adherence to a prescribed headway. The classes considered are variations of proportional feedback control with reference service headway and predictive control based on a rolling horizon, mathematical programming approach.

Feedback methods are as follows. A unity-gain forward headway control is used for establishing a benchmark for the effects of strict adherence to the service headway; a smaller gain is also tested because, in practice, gains in the range of [0.6, 0.8] would be used, see Cats et al. (2011). Another proportional feedback structure is the two-way headway method, similar to the “prefol” method described by Turnquist (1982), also used in Xuan et al. (2011) and Cats et al. (2011). The third structure is similar

10.1016/j.ifacol.2016.07.035

to the first, but with a threshold rule that turns off the control everytime the headway between a bus and its pre-predecessor is more than twice the reference headway, indicating large delays of the preceding bus.

Predictive control is examined to test methods that do not prescribe a reference headway, hence allowing the accommodation of disturbances without resorting to strict headway adherence. The control law is the same as presented in Koehler et al. (2011).

System performance is assessed by means of control objectives given by headway regularity (for feedback control) and passenger delay times (for predictive control). Although not being directly considered in the control objectives, comfort is analyzed by the number of standees on the bus and its related indicator, the perceived passenger delay.

2. CONTROL STRUCTURES

The following indices, parameters and variables are used to model the bus system:

λ_k	passenger arrival rate at station k (pax/s);
$a_{i,k}$	arrival time of bus i at station k (s);
C_0	time required to start boarding and alighting operations after bus arrival (s);
C_1	time for passenger boarding (s/pax);
C_2	time for passenger alighting (s/pax);
$d_{i,k}$	departure time of bus i from station k (s);
H	service headway (s);
I	set of buses in the system, $I = \{1, \ldots, n_I\}$;
i	bus index;
k	station index;
K_c	proportional control gain for feedback structures;
$l_{i,k}$	number of onboard passengers in bus i upon departure from station k (pax);
n	number of doors for alighting and boarding;
N_i	set of stations belonging to the prediction horizon of bus i;
n_I	number of buses of the system;
q_k	fraction of onboard passengers alighting at station k;
r_max	maximum holding time at stations (s);
$r_{i,k}$	holding time of bus i at station k (s);
$s_{i,k}$	duration of boarding and alighting process for bus i at station k (s);
t_k	nominal travel time between stations $k-1$ and k (s).

All control methods presented below act to regulate headways, if necessary, by holding buses at any station after alighting and boarding processes.

2.1 Forward control (FH)

The forward headway controller applies holding whenever after a bus finishes the alighting and boarding processes with a headway lower than the service headway. The holding will last for the time needed to restore the service headway, being calculated by:

$$r_{i,k} = K_\mathrm{c}[H - (a_{i,k} + s_{i,k} - d_{i-1,k})]^+ \qquad (1)$$

in which $[u]^+ = \max\{0, u\}$. The headway is calculated as the difference between bus i's expected departure time $(a_{i,k} + s_{i,k})$ from station k and the departure time of its preceding bus $(d_{i-1,k})$.

Despite being simpler than other more elaborate control methods, such as predictive control, forward headway control is suited for benchmark as a headway control policy that seeks to correct any headway shorter than the service headway. Letting $K_c = 1$ implies strict adherence to the prescribed headway; in practice, $K_c < 1$ is used to avoid large holding actions.

2.2 Two-way headway control (TWH)

This method holds a bus i to balance the headway with its preceding and following buses. More precisely the holding time is given by:

$$r_{i,k} = K_\mathrm{c}[(d_{i+1,k'} - d_{i,k'}) - (a_{i,k} + s_{i,k} - d_{i-1,k})]^+ \qquad (2)$$

in which k' is the station last visited by bus $i+1$, $(a_{i,k} + s_{i,k} - d_{i-1,k})$ is the expected headway between buses i and $i-1$ without holding, and $(d_{i+1,k'} - d_{i,k'})$ is the last observed headway between $i+1$ and i.

2.3 Forward headway with threshold control (FTH)

This method is derived from FH in which the holding time $r_{i,k}$ given by (1) is applied, unless the headway between i and the pre-preceding bus $i-2$ is more than twice the service headway. In mathematical notation, this method is given by:

$$r_{i,k} = \begin{cases} 0, & \text{if } (a_{i,k} + s_{i,k} - d_{i-2,k}) \geq 2H \\ \text{Eq. (1)}, & \text{otherwise} \end{cases}$$

If the headway between bus i and bus $i-2$ exceeds twice the service headway, it is considered that bus $i-1$ is over delayed. In such a situation, holding is not applied for bus i in order to prevent a ripple holding effect on all the succeeding buses, which would invariably degrade overall system performance.

2.4 Predictive control (opt.H)

The predictive control method is based on the mathematical programming model presented by Koehler et al. (2011). The control method assumes the availability of the following historical data:

- passenger arrival rates at stations;
- passenger alighting rates at stations;
- dwell time function parameters;
- bus travel times between stations;
- departure time at last visited station;
- number of onboard passengers.

The model is based on the following assumptions:

- passenger load capacity is not considered (no residue of queues at stations);
- bus travel time between stations is approximated by the expected value;
- boarding and alighting times are approximated by a deterministic linear function;
- no overtaking of buses is allowed.

Holding times are obtained by minimizing the total passenger delay in the bus system, according with the following cost function:

$$f = \sum_{i\in I}\sum_{k\in N_i}\left[\frac{\lambda_k}{2}(\hat{d}_{i,k} - \hat{d}_{i-1,k})^2 + (1-q_k)\tilde{l}_{i,k-1}(\hat{r}_{i,k} + \hat{s}_{i,k})\right] \tag{3}$$

in which $\hat{u}$ denotes estimated values and $\tilde{l}_{i,k-1}$ is computed iteratively in the horizon as in Koehler et al. (2011).

The cost function (3) considers all buses in the system and the stations that are within the prediction horizon N_i of each bus i. The first term of (3) represents the delay of passengers at stations, while the second term accounts for the delay of onboard passengers.

The constraints that represents the bus arrival times at stations, bus dwell times, bus departure times from stations, onboard bus passengers, and maximum holding times are given by (for all $i \in I$ and $k \in N_i$):

$$\hat{a}_{i,k} = \hat{d}_{i,k-1} + t_k \tag{4}$$

$$\hat{s}_{i,k} = \frac{nC_0 + C_1\lambda_k\left[(\hat{a}_{i,k} - \hat{d}_{i-1,k}) + C_2 q_k \hat{l}_{i,k-1}\right]}{(n - C_1\lambda_k)} \tag{5}$$

$$\hat{d}_{i,k} = \hat{a}_{i,k} + \hat{s}_{i,k} + \hat{r}_{i,k} \tag{6}$$

$$\hat{l}_{i,k} = \lambda_k(\hat{d}_{i,k} - \hat{d}_{i-1,k}) + (1-q_k)\hat{l}_{i,k-1} \tag{7}$$

$$\hat{r}_{i,k} \leq r_{\max} \tag{8}$$

$$\hat{a}_{i,k}, \hat{s}_{i,k}, \hat{d}_{i,k}, \hat{l}_{i,k}, \hat{r}_{i,k}, \geq 0 \tag{9}$$

Given that the model is represented by the convex cost function (3) and linear constraints (4) to (8), the resulting problem is of the quadratic programming class. This type of problem can be solved by off-the-shelf software packages in real time. The solution to the headway control problem determines, for each station within the prediction horizon, how long each bus in the system should be held to minimize passengers delay.

The optimization problem is solved every time a bus is ready to leave a station. Although the prediction horizon usually covers multiple stations for all buses in the system, holding is applied (if necessary) only to the bus that is about to depart. The other holding decisions belonging to the prediction horizon are not implemented. When holding time computation is triggered, the current system state is updated with data relative to the latest stations visited by the buses, the prediction horizon is rolled forward, and a new control action is calculated by solving problem (3)-(9). Such a recurring revision of control actions compensates for stochastic phenomena not taken into account in the model.

3. SIMULATION OF CONTROL METHODS

To investigate the performance of the control methods, a scenario was built on the Aimsun microsimulation software, see Transport Simulation Systems (2012). Its application program interface (API) enables the setting of dwell times according with passenger activity (boarding/alighting) and holding decisions.

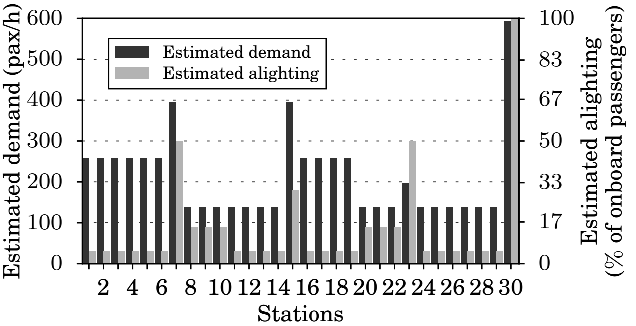

Fig. 1. Estimated demand and fraction of passengers alighting at each station

The particular bus line studied in this work is part of a BRT system proposed for the city of Florianópolis, Brazil. The line operates in a route with 31 km and 30 stations. Buses ride on exclusive corridors and receive full priority in all intersections. The itineray is circular, so a bus arrives at the terminal station and immediately starts a new round trip.

The buses have a capacity of 150 passengers (58 seated) with three doors for both alighting and boarding. To comply with the peak demand of approximately 6300 pax/h a service headway of 3 min is set. The peak-hour cycle time is approximately 51 min which requires an operational fleet of 17 buses. Overtaking is not allowed and all stations have a bay for two vehicles.

Fig. 1 shows the passengers arrival rates and the fraction of alighting passengers at each station. Arrival rates are the average of Poisson distributions whereas the alighting fractions are normally distributed with standard deviation σ equal to half the average. All onboard passengers alight at the terminal (station 30).

Boarding/alighting parameters are: $C_0 = 3\,\mathrm{s}$, $C_1 = 1.5\,\mathrm{s/pax}$, and $C_2 = 1\,\mathrm{s/pax}$, according to Wright and Hook (2007). Passengers that arrive while the bus is at station also board. Holding is limited to 40 s.

Bus characteristics are defined in terms of acceleration, deceleration and speed acceptance, which are set to $1\,\mathrm{m/s^2}$ ($\sigma = 0.3\,\mathrm{m/s^2}$), $2\,\mathrm{m/s^2}$ ($\sigma = 0.3\,\mathrm{m/s^2}$) and 1 ($\sigma = 0.05$), respectively. Maximum speed of bus corridors are 50 km/h, however, real cruise speed of each bus will be defined by the random speed acceptance value.

The warm-up period lasts until the first bus finishes a complete trip (reaching the terminal). During warm-up, no headway control is active and no data is collected for performance analysis.

4. PERFORMANCE INDICATORS

Several indicators are computed in order to compare performance of implemented control methods. The first is average headway deviation, which is obtained by the arithmetic mean of the difference between all observed headways and the service headway. The feedback control methods examined in the simulations seek to directly control headways, so we expect headway deviations to be smallest for this class of holding control.

Passenger delay at stations and onboard passenger delay are indicators used explicitly by predictive control and will also be examined. For performance measurement purposes, passenger delay at station is defined as the time interval between the arrival of the passenger at the station and the arrival of the next bus at the same station. Passengers board in order of arrival. Delay at station is computed for every passenger that is allowed to board the bus, and those who are prevented from boarding due to capacity limits are the first to board the following bus. Passengers waiting at the terminal station have their delays at station computed by the time interval between the arrival of the passenger at the terminal and the departure of the next bus.

Onboard passenger delay is defined as the time interval between the arrival and departure of a bus from a station for those onboard passengers that do not alight at the station and for those passengers who board at this station. At the terminal station all passengers alight, then no onboard passenger delay is computed.

Perceived passenger delay is introduced as the sum of onboard and at station delays. Onboard delay, however, is weighted by a time multiplier that reflects the crowding conditions on the bus. As a result, perceived delay combines the three user-related factors affected by headway control methods: delay at station, onboard delay, and bus crowding.

The concept of time multiplier used in this work to attribute more weight to onboard passenger delay when faced with uncomfortable conditions is based on the proposal by Whelan and Crockett (2009). They proposed a linear relationship between standee density (pax/m^2) and time multiplier based on a large scale program of market research at UK rail system. The linear equations adapted from their work are:

$$Tm^{\text{sit}}_{i,k} = 1 + \frac{0.63 l^{\text{sup}}_{i,k}}{L^{\text{sup}}_{i}} \tag{10}$$

if the passenger is seated, and

$$Tm^{\text{sup}}_{i,k} = 1.53 + \frac{0.51 l^{\text{sup}}_{i,k}}{L^{\text{sup}}_{i}} \tag{11}$$

if the passenger is standing up. In the above, Tm is the time multiplier, L^{sup}_{i} is the capacity of standees for bus i (6 pax/m^2 equivalent), and $l^{\text{sup}}_{i,k}$ the number of standee passengers for bus i at station k. From (10) and (11) it can be noted that while standee passengers experience a higher discomfort level when bus is not crowded, discomfort of seated passengers increases in a higher rate. Time multiplier is applied for all passengers to weight their onboard delay in every station.

To complement the performance analysis, we also present the operational speed, calculated as the spatial speed of all buses; and total holding obtained by adding up all applied holding actions over the one hour simulations.

5. SIMULATION RESULTS

The first comparison is made with the BRT system operating in open loop, with forward headway control and predictive control. Fig. 2 shows the resulting headway profiles. In 2(a), no control (open loop) operation is depicted showing that the service headway of 3 min cannot be sustained without control, leading to bunching behavior (represented by headways approaching 0 min). Note that when bunching occurs the operation is further deteriorated due to prohibited overtaking. Fig. 2(b) and 2(c) depict the headways under forward headway control (FH) and predictive control (opt.H), respectively. It is clear that forward control is able to prevent bunching. However, the occurrence of headways higher than the service headway is observed, since this method takes no action for delayed buses.

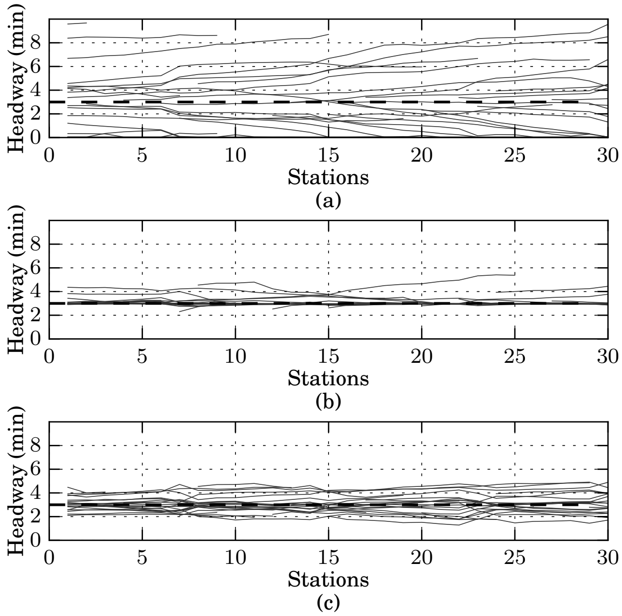

Fig. 2. Bus headway profiles during one hour of simulated time; (a) no control case; (b) headways induced by forward headway control (FH) with $K_{\text{c}} = 0.7$; (c) headways induced by predictive control with a prediction horizon extending to 10 stations (opt.H(10)). The dashed horizontal line represents the service headway of 3 min.

Predictive control is also able to prevent bunching, but there is no strict adherence to the service headway. Yet, such slack in control does not necessarily mean degraded service, as discussed next.

Table 1 reports the performance indicators yielded by the no control and headway control methods. To serve as a benchmark, the most rigid control, FH with $K_{\text{c}} = 1.0$, is presented. All other implementations of feedback control methods are done with $K_{\text{c}} = 0.7$. Three implementations of predictive control are presented, with prediction horizons of 7, 10, and 15 stations.

Total holding data from the table show that FH control typically requires more holding time that the opt.H in the effort to keep buses on the service headway. Headway standard deviations are smaller with feedback control, even though at the cost of a decrease in operational speed. Still, passenger delay at stations is the minimum with feedback control at the expense of onboard delay due to holding actions.

Table 1. Performance comparison of headway control methods; FH(1.0) refers to the forward headway control with unity gain; other feedback methods have gains of 0.7; opt.H(.) refer to predictive control with different horizon lengths (in terms of station lookahead).

Performance indicators	No control	FH(1.0)	FH	TWH	FTH	opt.H(7)	opt.H(10)	opt.H(15)
Average headway (s)	180	188	187	191	184	184	184	186
Headway standard deviation (s)	150	22	24	14	41	62	46	35
Total holding (s)	-	3024	2820	3794	2005	1941	2082	2574
Operational speed (km/h)	36.5	35.8	35.9	35.1	36.4	36.5	36.4	36.0
Perceived passenger delay (s)	278	216	212	230	208	200	197	202
Passenger delay at station (s)	139	81	81	80	83	88	83	82
Passenger delay onboard (s)	90	102	100	113	92	85	86	92
Average onboard standees (pax)	43	28	27	29	30	29	29	28

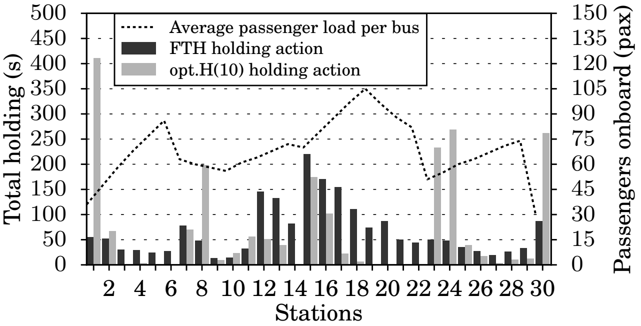

Fig. 3. Total holding time per station caused by forward headway with threshold with $K_c = 0.7$ and predictive control with prediction horizon of 10 stations, and average passenger load.

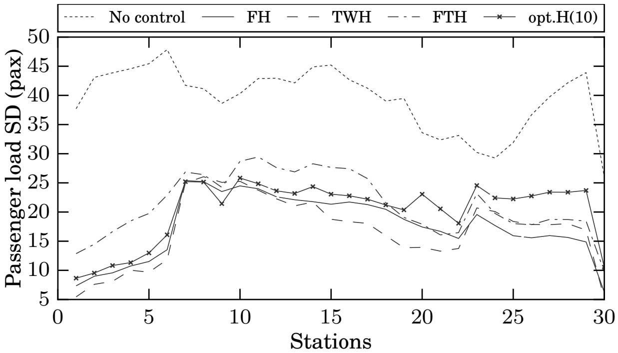

Fig. 4. Standard deviation of passenger load at each station for all headway control methods. Feedback methods are presented with $K_c = 0.7$, and predictive control is presented with prediction horizon of 10 stations.

Predictive control applied to the tested scenario produced a more balanced picture. Total holding in the three cases (opt.H(7) to opt.H(15)) is less than in the feedback cases with the exception of forward headway with threshold (FTH), which held buses less than opt.H(10) and opt.H(15). As a result, onboard delay is lower than with feedback, even in comparison with FTH: in this case, although less holding is exerted, it is used anywhere on the itinerary. Predictive control, however, applies more holding when buses are less loaded as will be discussed below. Hence, the delay onboard averaged over the total number of riders is less (equal, in opt.H(15)) with predictive control. The price payed in terms of station delay is not high and hence the perceived passenger delay reaches the smallest values with predictive control.

More insights into control behavior are provided by the graphs in Fig. 3, which presents holding time per station for FTH and opt.H(10). The dashed line indicates the average passenger load of all buses. Notice that feedback control issues holding times for buses at stations regardless of the passenger load, thereby incurring longer onboard delays per passenger. Unlike FTH, predictive control favored longer holding times for buses at stations when passenger load is relatively low. Conversely, note that opt.H avoids holding buses when loads are high (e.g., at stations 5, 6, and 17 to 22), which is in line with the approach of choosing few control points typically employed by schedule-based headway control methods. The advantage of predictive control is the automatic selection of control points based on dynamic conditions along the operation.

Granted, feedback methods are deployed to keep operation within the desired service headway. Accordingly, and for the case studied here, these methods performed better than predictive control in regulating service with the reference, as shown by the smallest values for headway standard deviations in Table 1. FTH presents higher deviations than other feedback methods, which is explained by its less intense holding actions. Still, the plot of average passenger load per bus in Fig. 3 shows that FTH control did not actuate at points with less passengers onboard as done by predictive control.

In line with the greater regularity of headways, feedback methods presented the lowest variations in passenger loads along the itineraries, as illustrated in Fig. 4. Predictive control, in turn, allowed greater load variability although being able to recover at some particular stations, as shown by the dips observed at stations 9, 21, 22, and 30. Overall, the flexibility on deviation from service headway displayed by predictive control did not lead to high variation on passenger load and delay at stations. This observation is confirmed by the passenger delay at stations under predictive control reported in Table 1, in the similar range of those achieved with feedback methods.

In terms of standing passengers, data in Table 1 show that there is little difference in average numbers for all control methods, being in the range of [27,30], while open loop operation yielded an average of 43 passengers standing up. In other words, all control methods achieved a similar and significant improvement on passengers comfort. Fig. 5 confirms that such similarity of the averages is kept along the itinerary for all control methods.

In terms of tuning the control methods, there was no extensive search for gain values (in the case of feedback

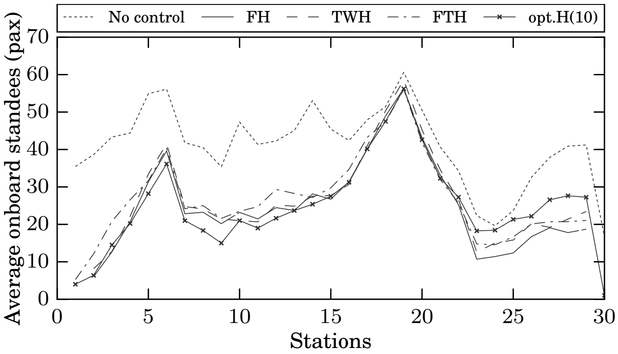

Fig. 5. Average onboard standees at each station for headway control methods; feedback methods presented have $K_c = 0.7$, and predictive control has a prediction horizon of 10 stations.

methods) nor weighting of cost terms (in the cost function of predictive control). Yet, the length of the prediction horizon was varied to provide some insight into the tradeoff between shortsightedness in adopting small horizons versus large estimation errors in the case of long predictions into the future. We remark that the model uses average estimates only, being deterministic. For the case studied, results presented in Table 1 endorse that better overall performance can be achieved with prediction horizon of 10 stations in the 30 station route. As expected, given the stochastic nature of the scenario it is advisable to restrict the horizon in order to avoid control actions computed with large estimation errors.

6. CONCLUSION

We presented comparative simulation results of feedback and predictive control strategies for bus operation. Results confirmed that trying to keep headways strictly at the reference causes relatively large onboard delays, although passengers delays at stations are small. Progressively relaxing adherence to the reference headway improves overall delays, albeit at the expense of added delay at stations and increasing variations in passenger loads.

The best overall performance in terms of passenger delay (onboard and at the stations) was achieved with predictive control. It appears that its advantages more than compensate for the extra real-time data needed for control calculations.

Future work will aim at extending the scenarios being tested with larger operational disruptions. Also, analysis of simplified scenarios for closed-form solutions will be undertaken.

REFERENCES

Cats, O., Larijani, A., Koutsopoulos, H., and Burghout, W. (2011). Impacts of holding control strategies on transit performance. *Transportation Research Record: Journal of the Transportation Research Board*, 2216, 51–58.

Eberlein, X.J., Wilson, N.H.M., and Bernstein, D. (2001). The holding problem with realtime information available. *Transportation Science*, 35(1), 1–18.

Koehler, L., Kraus, W., and Camponogara, E. (2011). Iterative quadratic optimization for the bus holding control problem. *IEEE Transactions on Intelligent Transportation Systems*, 12(4), 1568–1575.

Newell, G.F. and Potts, R.B. (1964). Maintaining a bus schedule. In *2nd Australian Road Research Board Conference*, volume 2, 388–93. Australian Road Research Board (ARRB), Melbourne, Australia.

Strathman, J.G., Kimpel, T.J., Dueker, K.J., Gerhart, R.L., Turner, K., Griffin, D., and Callas, S. (2001). Bus transit operations control: review and an experiment involving tri-mets automated bus dispatching system. *Journal of Public Transportation*, 1–26.

Transport Simulation Systems (2012). AIMSUN dynamic simulator users manual v. 7.

Turnquist, M. (1982). Strategies for improving bus transit service reliability. Technical Report NTIS/DOT/RSPA/DPB-50-81-27, Northwestern University, Evanston, IL.

Whelan, G. and Crockett, J. (2009). An investigation of the willingness to pay to reduce rail overcrowding. In *Proceedings of the First International Conference on Choice Modelling*. Harrogate, England.

Wright, L. and Hook, W. (2007). *Bus Rapid Transit Planning Guide*. Institute for Transportation & Development Policy.

Xuan, Y., Argote, J., and Daganzo, C.F. (2011). Dynamic bus holding strategies for schedule reliability: Optimal linear control and performance analysis. *Transportation Research Part B: Methodological*, 45(10), 1831 – 1845.