Measuring Subway Service Performance at New York City Transit: A Case Study Using Automated Train Supervision (ATS) Track-Occupancy Data

Brian Levine
Operations Planning,
MTA New York City Transit,
2 Broadway, Cubicle A17.100,
New York, N.Y. 10004-2207
<Brian.Levine@nyct.com>

Alex Lu
Transit Analyst,
Haverhill, Massachusetts 01835-7234
http://www.lexciestuff.net/
<lexcie@gmail.com>

Alla Reddy (corresponding author)
Sr. Director, System Data & Research,
Operations Planning,
MTA New York City Transit,
2 Broadway, Office A17.92,
New York, N.Y. 10004-2207
<Alla.Reddy@nyct.com>

Word Count: 202 (Abstract) + 6,690 Words + (6 Figures * 250 Words) = 8,392 Words

ABSTRACT

A recurring challenge facing transit managers today is the persistent question of how to do more with less—to maintain and improve service despite deficits of historic proportions. New York City Transit (NYCT) responded by re-tooling performance measurement frameworks and procedures to better capture customers' perspective, respond to management initiatives, and incentivize proper operating decisions. NYCT's primary performance measure, Wait Assessment (WA), measures customers' maximum wait times while waiting to board at stations. Defined as percent of headways between trains not exceeding 125% of scheduled headways, a "Reach and Match" algorithm was developed to account for NYCT's irregularly scheduled service and ensure customer experienced headways are matched to the specific published scheduled headway in effect at that moment, regardless of which scheduled trip was supposed to arrive. Upgrading sample-based methods that gathered limited data manually, track-occupancy data was downloaded from the Automated Train Supervision (ATS) system for the No.1 through No.6 routes, providing 100% coverage, much lower public reporting time-lag, and the ability to take near-term corrective action. The increase in data availability also allows NYCT to easily consider corridor-level and track-level WA standards for internal diagnostic purposes, analyzing train performance in shared-track territory regardless of route designations, to provide better service.

INTRODUCTION

New York City Transit (NYCT) operates the third largest subway system in the world (by annual ridership), carrying about 5.0 million riders on an average weekday. The subway system extends 830 track miles through four boroughs, covering a service area of 321 square miles and serving a population of 8.0 million people 24-hours, seven days a week. NYCT's predominant role is to ensure that trains and buses operate safely, reliably, on-time, and provide convenient services to the customer in a cost effective manner. One of the tools available to ensure the mission is being carried out properly is an independent performance audit infrastructure—outside of both the operations management and the customer advocacy groups—and continuous applied research and improvements in not only monitoring methodologies but also how the service can be improved.

This paper describes New York City Transit's case study in using track-occupancy data to measure subway service performance, specifically Wait Assessment, and to understand how these metrics are useful in analyzing and improving service. WA is the percentage of actual headways between successive trains that are less than or equal to a given standard. An algorithm was developed that objectively matches actual observed train headways to scheduled service headways depicted in the timetable. Prior performance research tended to provide the matching based on trip identifiers, or provides a subjective matching by a schedule or operations expert *(1)*. This algorithm builds upon previous research by considering what actual headway a customer should be experiencing at that moment in time, instead of using an average, making standardized performance results sensitive to both service delivery and proper schedulemaking. The algorithm also avoids subjectivity or ambiguity in the analytical process of matching actual observations and scheduled timeslots where trips were dropped, as discussed in the literature *(1)*. Results from applying the algorithm to a variety of measures and standards are presented, as well as how operations management is provided with daily "flash" and outlier reports used to diagnose performance problems and ultimately recommend strategies for improving service delivery. This research is one piece of the overall performance measurement and service improvement framework at NYCT, which includes the Passenger Environment Survey *(2)*, general performance indicator reporting *(3)*, key performance indicators *(16)*, environmental justice monitoring *(17)*, fare evasion monitoring *(18)*, and others.

Characteristics of the New York City Subway

NYCT uses a relay-based interlocking system to control train traffic throughout its "A" Division (formerly the Interborough Rapid Transit) network of seven numbered subway routes. ATS (Automated Train Supervision) is a non-vital centralized dispatching system overlaid on top of existing local relay logic and remote control panels in master towers. ATS provides the Operations Control Center (OCC) with real-time track occupancy information and track the identity of each train as it proceeds throughout the network, allowing automated route-setting based on pre-loaded schedule data *(19)*. The real-time track occupancy information provided by ATS is stored in central servers, providing constant service reliability and performance monitoring capabilities, and aiding service delivery improvement strategies.

NYCT is unique in North America in providing frequent service throughout the day over multiple interconnected routes. During rush hours, all routes have typical headways between 2 and 8 minutes with minor exceptions, and passengers do not generally arrive according to prescribed schedules; instead they arrive at stations at random times since they know they will not have a long wait. NYCT's publicly available timetables do not generally give exact arrival times; instead they may say, e.g., service is provided every 4~7 minutes on Northbound “5” route on weekdays. During most off-peak hours, standard service headway is 10 minutes or more frequent, whereas in overnight periods, policy headway is every 20 minutes.

Another feature of scheduled subway service is that headways on trunk lines are rarely uniform, due to interactions between routes of different service frequency at flat junctions or merge points:

- Summer 2006 schedules contained off-peak service that required southbound “4” and “5” trains to merge at 149 St. in the Bronx on a 2:1 ratio, with “4” trains operating every five minutes north of 149 St., but changing to a four-and-six minute pattern to the south, allowing a “5” train to operate in the six-minute gap between successive “4” trains.
- Winter 2009 schedules required “M” trains to cross over from a division where the prevailing headways are multiples of six, to one where prevailing headways are multiples of eight. A schedule conflict therefore occurs predictably every two hours, which must be resolved by introducing irregular intervals.
- Prior to 2009, southbound “R” and “N” trains splits on the Broadway Line at even intervals, yet “R” runs ‘via Tunnel’ and incurs a 10-minute trip time to Dekalb, whereas “N” runs ‘via Bridge’ and incurs an 8-minute trip time to Dekalb, making headways potentially uneven at all points south.

NYC subway, an amalgamation of three previously independent subway systems, required a good deal of passengers to transfer trains to get to their final destination; passengers are thus generally more concerned about waiting time, than on-time performance at the train’s terminus.

Performance Indicators

The Performance Indicator (PI) program was established in 1994 in response to the MTA Inspector General’s research *(4)* recommending the need for measures of service reliability other than the traditional Terminal On-Time Performance (TOTP). TOTP is a good operational measure for commuter railroads where the majority of customers are traveling to the final stop in the central business district (CBD). However, transit routes tend to drop off and pick-up many passengers at intermediate stations, which requires more sophisticated measures capable of blending waiting time and travel time experiences from a customer perspective.

Turnquist and Bowman *(5)* describe the effects of network structure on service reliability. They find that controlling link travel time variability and scheduling to ensure easy transfers are both important, and that that service reliability is very sensitive to frequency of service. This becomes a key factor in the NYCT subway system and requires a reliability measure that can distinguish between minor differences in service.

Extensive research had been conducted to understand transit service reliability from passengers’ and transit managers’ perspectives *(6)*, building on prior models of headway variance *(7,8)*. Furth and Muller *(9)* describe a method to determine potential passenger waiting time using automated vehicle location data. The authors assume that for short headways, passengers arrive independent of vehicle arrivals. While this may be true on single route lines, many of NYCT’s stations in the CBD are large complexes where transfers are possible between numerous routes, hence leading to heavy loading when trains arrive. Statistical measures of service reliability, such as root-mean-squared average passenger wait time *(10)*, were considered too complex for use as public measures, as a key criterion of NYCT’s performance measures is that people would not need a mathematical background to understand its significance.

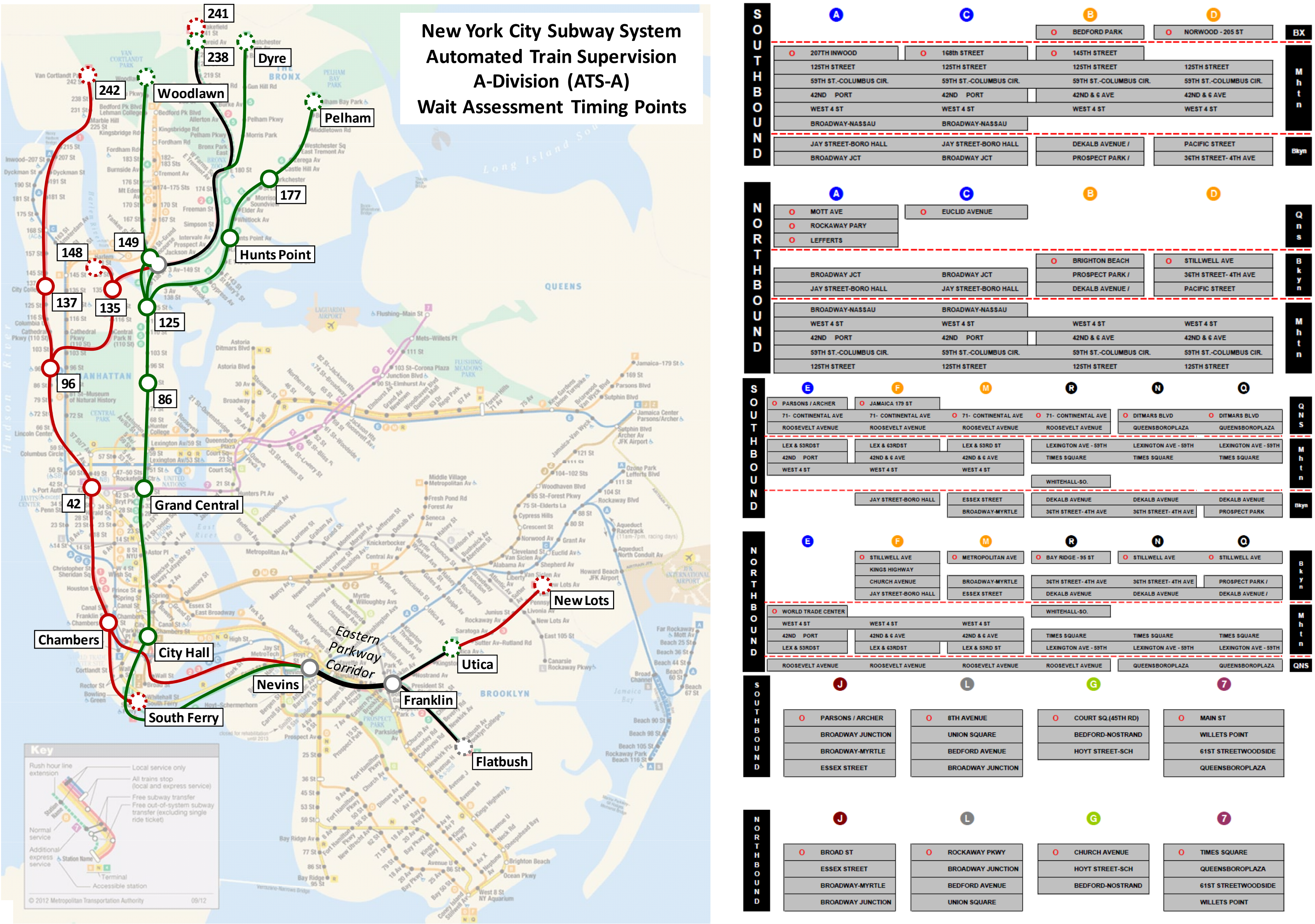
NYCT developed a simplified version of algorithms and ideas described above, more easily understood by passengers and operational management. The result was Wait Assessment (WA), which effectively measures how long customers may wait for a given train at a given station, calculated by comparing successive headways between trains to a prescribed standard. Though this measure does not focus on specific passengers, it ensures that the longest time passengers have to wait is within the defined standard, even if many passengers were not waiting as long. In effect, it describes a worst-case scenario. This strict standard promotes consistent service throughout the whole system; in fact, many passengers have wait times shorter than this amount, especially at large transfer complexes. Service during overnight periods is provided less frequently; WA is not calculated during these periods since many people arrive according to schedule, making on-time performance a more important measure.

The Performance Indicators' main purpose is to monitor how well NYCT is providing service to the public. Wait Assessment is publicly reported at the systemwide and route levels. These results are used by rider advocacy groups for their annual rating of subway routes—the *State of the Subways (11)*. Results are also often reported in the media. Maintaining a transparent and accountable performance reporting process is critical to achieving public trust. Indeed, stakeholder and watchdog groups have adopted the MTA's measures as the basis of their performance reporting *(12)*. Data is also used to specifically identify operations issues.

WAIT ASSESSMENT "REACH AND MATCH" ALGORITHM

Wait Assessment applies an analytical algorithm on raw data collected either by surveyors or the ATS system, specifically using departure times of all vehicles passing a location. Data is collected at all en-route timepoints (Figure 1(a)), because majority of riders enter and depart the system at intermediate stops along the route, not just at terminals. It is important to provide service reliability measures at these locations. Typical routes have anywhere from 5-15 timepoints (25-50% of total stops), a subset of which are used for WA reporting purposes.

Locations used for WA reporting are agreed upon by management in Operations Planning and Rapid Transit Operations, and generally include major transfer stations, hubs, and departing terminals, since many customers board at these locations. Figure 1(a) shows current Wait Assessment timing points for ATS-enabled A-Division routes. Timepoints with a dashed circle are only used when the train originates at that location. For example, the "2" route, which travels from the 241 St in the Bronx to Flatbush Ave in Brooklyn via Manhattan, contains 13 timepoints out of 49 total stops. Currently, 8 of these timepoints are used for WA reporting: 1-2 in the Bronx, 4 in Manhattan, and 2-3 in Brooklyn. Timepoints for the non-ATS territory are shown in Figure 1(b).



IGURE 1. New York City Subway System Wait Assessment Timing Points: (a) ATS-A enabled A-Division (numbered lines); (b) Non-ATS erritory.

The algorithm (described below) matches actual headways to scheduled headways, and then applies a uniform standard across all trains and all routes to ensure consistent reporting. The different standards applied are discussed later. This methodology accounts for varying headways on a single route due to operational characteristics, instead of applying the standard against an average headway. Consider a hypothetical service that runs every 3-4 minutes during rush hour. An average headway would imply a 3.5 minute standard, when in actuality at a given moment in time the scheduled headway may either be three or four minutes. The goal of this algorithm is to match the actual service at a specific point in time to the scheduled service that is supposed to be provided at that moment.

For all actual observations of consecutive trains passing a given timepoint, actual departure headways must be matched to all scheduled headways based on the daily schedule with supplemental schedules applied, to calculate which trains meet the WA standard. Supplements are alterations to the schedule due to construction and/or maintenance on the train track which necessitates re-routes of trains or added running time. The effects of these schedule changes are published online, and it is assumed that the public is aware of the altered schedule in terms of different headways. The matching process is governed by the "Reach and Match" algorithm, described in brief below and in more detail at the end of the paper, and shown in Figure 6(a). The matches are made for <u>timepoint</u> locations only, since those are the only locations where scheduled departure time and/or scheduled headways are published. Departure times are used because a feature of NYCT's subway schedule is scheduled holding times at key transfer stations. A train may be scheduled to arrive at 10:52am but scheduled to depart at 10:54am. Even though passengers may sit on the train once it arrives, the train does not actually provide service until its departure time.

This algorithm's primary element is the train matching step, whereby actual service headways are matched to scheduled service headways based on departure time. Actual trains can only be matched if they are within acceptable matching boundaries, based on the scheduled headway between itself and adjacent trains in both temporal directions (i.e. prior and following trains). Trains may be matched to more than one scheduled train only when there are gaps in service, allowing for multiple opportunities to pass Wait Assessment (though the train may only pass in one of these intervals). Trains may not be matched when there is too much service within a given scheduled headway, because extra trains should not help the WA metric if they do not provide service in distinctly separate intervals. In other words, a single train that has successfully picked up passengers cannot be evaluated again, however a train with a long actual headway may fail to pick up passengers in multiple scheduled headways.

This algorithm was developed to account for the "drift" that naturally occurs as actual service trains move out of scheduled slots in normal daily operations. The intent of the algorithm is to describe the service headways as experienced by customers who are expecting a train to arrive every *h* minute(s), where *h* is the headway specified in the timetable, i.e. regardless of what <u>specific</u> train was supposed to arrive at that time. The intent of this "Reach" criterion is to prevent actual and scheduled trains from drifting too far away from each other. While up to one headway of give is allowed, to account for operational schedule adjustments (called the "flex"), as soon as the actual train headways drift out-of-sync with the timetable, it is considered "Out-of-Reach" and is not used to make a "Match".

After all trains are matched, the Wait Assessment results are calculated by comparing actual service headways to scheduled service headways. If the actual service headway is greater than scheduled service headway by an allowable margin, that interval is denoted as failing WA. A discussion of the different WA measures follows.

MEASURES OF WAIT ASSESSMENT

T represents the threshold headway which delineates passing and failing Wait Assessment headways. The exact values used depend on how strict the performance measure is intended to be. Examples of these WA standards and measurements are discussed below.

Absolute vs. Relative WA

As first conceived, Wait Assessment was an absolute measure of relative performance. It's an absolute measure because the thresholds of what constitutes an acceptable excess wait time *(1)* is a fixed quantity for a given time period (+2 minutes peak/+4 minutes off-peak). However, it measures relative performance because it is based on headways between trains—obtained by comparing a train's departure time with its predecessor, and not by comparing it with the fixed schedule. The rationale for this was to provide customers with a fixed standard of excess wait time above which the service headway is considered unacceptable. This type of measurement metric has one interesting property: routes scheduled with shorter headways tend to score higher in Wait Assessment, because there is simply a higher probability of a train—any train—achieving that two-/four-minute window above headway.

In discussions with operations management, it became apparent that this property does not give dispatchers correct incentives. High frequency service routes are often very congested, where the smallest perturbation in headways or ridership volume can quickly snowball into bunched service and big gaps *(14)*. On lower frequency routes, dispatchers have more latitude to adjust schedules; proper headway isn't as critical to maintaining service performance as holding for connections at major transfer points. To prevent such imbalance, the standard to which each route is held must be a function of service frequency, with busier routes held to more exacting standards. In turn, the tolerable excess wait time must be specified relative to the headway.

New York City's "7" route runs from Times Square in Manhattan to Flushing Main St. in Queens. Service on this route runs every 4-5 minutes even during off-peak periods. With a standard of up to four minutes of acceptable excess wait time, dispatchers can still achieve an above 90% WA score even if delayed trains were allowed to depart in bunches, which does not incentivize dispatchers to properly perform train regulation functions. Earlier research *(20)* had shown that dispatching decisions are often driven by crew requirements; absent strong incentives to maintain proper headway, dispatchers sometimes allowed trains following a heavily delay to operate without being checked, in an effort to "catch up" to schedule and therefore minimize required real-time crew manipulations at terminals.

After much consultation with stakeholders and management, the Wait Assessment threshold was modified to be within +25% of scheduled headway, thereby making WA a strictly relative performance measure. This corresponds to $T = {}^{5}/_{4}(t_{i+1} - t_i)$, where $(t_{i+1} - t_i)$ is the scheduled headway between consecutive trains. The strictest standard at +25% was selected in an effort to assure the public that NYCT is looking for continuous improvements in service delivery. The formal definition of Wait Assessment for subway routes is thus defined as percentage of headways between trains that does not exceed 25% of scheduled headways. For example, for scheduled headways of four minutes, an actual headway of less than five minutes would be permissible.

The revised Wait Assessment shows results of tighter Headway +25% standard, however, it does not give any information on distribution of service headways not meeting standard. Line Managers (responsible for a single route) felt that simply knowing 20% or more of their trains were more than 25% outside the scheduled headway did not help them pinpoint the source of bunching problems, which often started with an overcrowded train that sometimes became more than double the headway later than its predecessor. To assist the management, the "failing" Wait Assessment headways were further broken into three subcategories: 25%-50% more than headway, considered a "minor gap", 50%-100% above headway, a "medium gap", and more than double headway, a "major gap". This type of reporting clearly shows the different levels of service provided, and provides a multi-dimensional picture of actual service quality delivered and system performance without modifying existing standards.

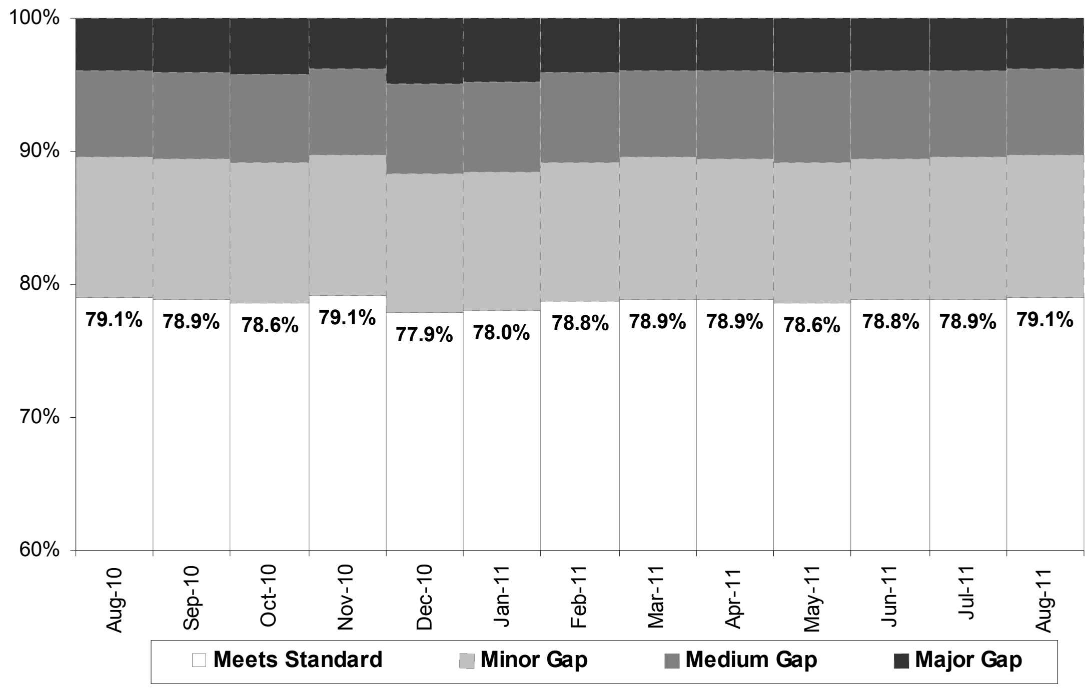


FIGURE 2. Wait Assessment Results with Distribution

An example of this is shown in Figure 2. Though overall Wait Assessment at the 25% standard for the entire system hovers around 79%, approximately 95% of actual train headways are within 100% difference of the scheduled headways, showing that despite incidents and crowding affecting the quality of service, service is in fact provided consistently on these routes albeit at increased headways.

In the future, Figure 2 will become interactive, allowing line management to drill down to see location(s) and time(s) of day that WA did not meet the 25% standard and to determine where major gaps in service tend to occur on a regular basis—which could be pro-actively filled in with an appropriately named "gap train". The reports shown in Figure 5 provide an example of location- and time-based drill-down, which will be expanded to provide distribution data.

Using ATS data, the Wait Assessment performance measure can be calculated for all routes and all time periods. This allows peak time periods to be reported separately, allowing explicit monitoring of performance during maximum ridership.

Line-, Corridor-, and Track-Level Wait Assessment

New York has many subway routes, some of which are co-routed on the same physical line infrastructure. The Queens Boulevard Line, a major four-track subway corridor in Queens, actually hosts two local routes, the "R" and "M" Trains, and two express routes, the "E" and "F" Trains. The local routes share local tracks, which makes all stations stops; express routes share express tracks that has station platforms only at major transfer points.

This creates a dilemma when measuring Wait Assessment at stations served by more than one route. At those stations, customers have a choice as to which route they would like to use. A passenger may take the first train that arrives, using it to get as close to their destination as the route permits, and make transfers to different trains later on in their journey. Yet other customers prefer one-seat rides and will wait on the platform for the exact route they require. Wait Assessment is a route-based measure; it measures the headways between trains of the same route—and does not consider headways between trains sharing the same track if they are assigned different route letters or numbers.

There is an internal NYCT debate about how such shared-track corridors should be managed. Customers destined for branching section of trunk lines (e.g. Concourse, Rockaway, Culver) often require a specific route to reach their final destination, and therefore are interested in knowing if their route is having bunching and spacing problems. Customers who use trunk sections exclusively (e.g. 8 Avenue, 6 Avenue), or make inter-divisional transfers at major transfer points, are usually more interested in corridor-level measures because route designations are only of passing interest to them—for they simply require any train headed in the direction they're traveling. When routes share track, it is sometimes operationally important for crowding reasons to keep even spacing between trains even if they have different route designations, therefore line management usually are more interested in corridor-level measures.

Furthermore, when dealing with incidents affecting service such as sick passengers or an inoperative switch, transit supervisors have the ability to reroute trains from express tracks to local tracks, or vice versa, to provide service where there otherwise would have been none. The fact that this happens quite often in the NYCT subway network makes a strong case for a track-level wait assessment measure whereby the route designation of the train itself does not matter, rather the fact that service is being provided on a particular section of track, better reflecting what an actual customer is experiencing.

Having both route-level and track-level WA results provide a complete picture of service as experienced by different types of passengers. Passengers waiting for a particular service to arrive are more concerned with route-level results, whereas passengers who are indifferent to service designation are more concerned with track-level wait assessment results. Overall these results together are more reflective of customer experience.

Figure 3(a) shows comparative WA results for route level, corridor level, and track level for Friday, September 23, 2011 on Eastern Parkway corridor in Brooklyn only for the period from 17:00 to 21:00.

The Eastern Parkway corridor is shown in Figure 3(b) and consists of the "2" and "3" routes which are serve all stops, and the "4" and "5" routes which serve only express stops. These routes travel from the Bronx / Northern Manhattan to Brooklyn. The "2" and "5" routes have roughly the same termini in both the Bronx and Manhattan, and the "3" and "4" routes have roughly the same terminal in Brooklyn. Accordingly, service can be adjusted across routes to provide better service when an incident occurs.

(a)

	Route/Line	Scheduled Headways	Passing Headways	Percentage Passed
WA 25% 9/23/2011 at Atlantic Ave 1700-2100 Southbound	4	39	27	69.20%
	5	26	43	64.20%
	4/5 Corridor	67	43	64.20%
	4/5 Express Track	67	40	59.70%
	2	27	12	44.40%
	3	27	9	33.30%
	2/3 Corridor	56	23	41.10%
	2/3 Local Track	56	25	44.60%
WA 100% 9/23/2011 at Atlantic Ave 1700-2100 Southbound	4	39	34	87.20%
	5	26	24	92.30%
	4/5 Corridor	67	54	80.60%
	4/5 Express Track	67	52	77.60%
	2	27	13	48.10%
	3	27	13	48.10%
	2/3 Corridor	56	30	53.60%
	2/3 Local Track	56	36	64.30%

(b)

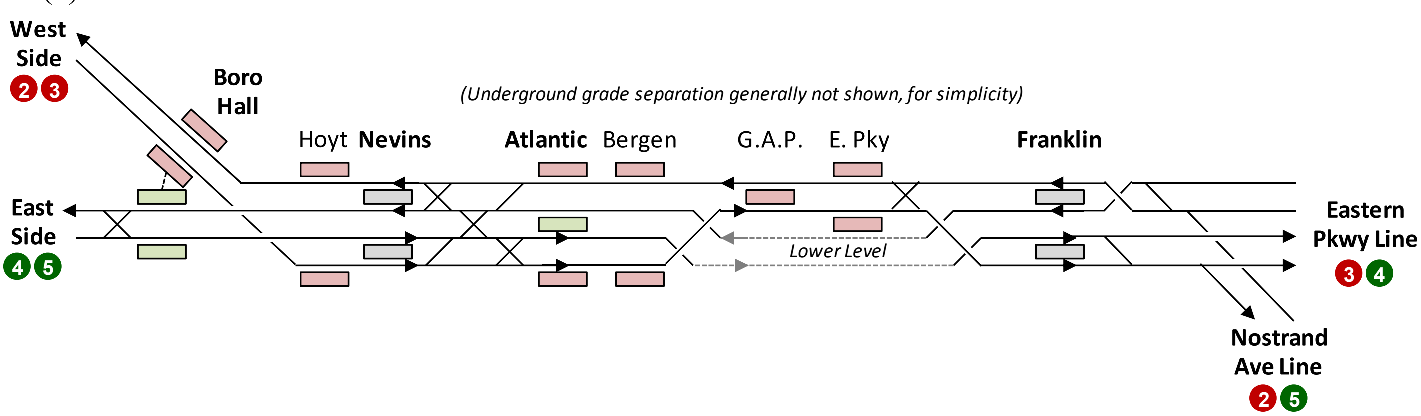


FIGURE 3 Wait Assessment Experiment on the IRT Eastern Parkway Corridor in Brooklyn: (a) Single-route, Corridor-, and Track-level Wait Assessment Results for Afternoon Peak Period on September 23, 2011; (b) Functional Track Layout of the Segment Discussed.

On this day there were numerous incidents on the 7 Ave IRT in Manhattan, affecting evening Southbound service in Manhattan and Brooklyn. The Department of Subways attempted to balance service by rerouting trains. In particular, since local "2" and "3" Trains experienced a partial blockage in Manhattan, certain express "4" and "5" trains were rerouted to local track once they arrived in Brooklyn, at Nevins Street Interlocking, to serve Bergen Street, Grand Army Plaza, and Eastern Parkway stations. Figure 3(a) shows local track WA was slightly higher than corridor WA due to re-routes of express trains to local track. Similarly, express track WA along Eastern Parkway was lower than corridor WA due to removal of “4” and ”5” trains from express track. Across the whole day, differences in Wait Assessment results were minimal (63.2% vs 62.2% 4/5 corridor versus track; 57.0% versus 57.2% 2/3 corridor vs track), but track-level performance is a better reflection of what customers experienced that day.

Local track WA 25% was not much higher than corridor level WA 25% due to fewer overall trains in service. Since the Wait Assessment measure is indirectly a function of train throughput at a given stop, regardless of operational changes put into effect to alleviate inconsistencies caused by incidents, WA at the 25% standard cannot improve drastically when fewer trains are in service. However, track-level WA 100% is higher than corridor-level WA by over 10%, indicating that although fewer trains were in service, they were spaced at approximately double scheduled headways, quantifying attempts made to provide service where there otherwise would have been none. Although trains providing service on the local track were officially designated “4” trains, customers used these trains to reach local stops. Express track WA decreased slightly as compared to corridor WA, as expected, but it shows that express service was not severely degraded when diverting some trains to “run local”.

ALGORITHM RESULTS

Figure 4(a) shows results from the headway matching algorithm for the “2” route in Brooklyn headed southbound at Atlantic Avenue for 9/23/2011, along with the calculation of Wait Assessment at the 25% and 100% standard. Each train is assigned a train ID, which indicates the route number of the train (05), the departure time from the origin (1301+, 1309+, etc), followed by a plus sign if the train is scheduled to depart on the half minute (e.g. 1301+ is 13:01:30), and also codes for origin terminal (241) and destination terminal (FLA). The primary terminals for the “2” train are 241, Wakefield–241 Street in the Bronx, and FLA, Flatbush Avenue–Brooklyn College in Brooklyn.

Based on algorithm previously described, the schedule is first used to determine the scheduled headway at that particular station at the time the train is scheduled to arrive. The “reach and match” part of the algorithm matches train headways based on the relationship between their scheduled time, scheduled headway, and actual departure time, to determine whether an actual observed departure is within reach of a scheduled slot. Actual headways that are too wide or too narrow can easily cause actual departure times to fall out of reach of the scheduled slots. The results in Figure 4(a) show a few interesting and important properties of the algorithm, which are discussed in Figure 4(b).

ID	SCHD_TRAIN_ID	SCHD TIME	SCHD HDWY	MATCHED_TRAIN_ID	ACT TIME	ACT HDWY	WA 25%	WA 100%
1	02 1301+ 241/FLA	141700	480	02 1301+ 241/FLA	141836	430	PASS	PASS
2	02 1309+ 241/FLA	142500	480	02 1309+ 241/FLA	142546	363	PASS	PASS
3	02 1317+ 241/FLA	143300	480	02 1317+ 241/FLA	143149	330	PASS	PASS
4	02 1325+ 241/FLA	144100	480	02 1325+ 241/FLA	143719	610	FAIL	PASS
5	02 1333+ 241/FLA	144900	480	02 1333+ 241/FLA	144729	724	FAIL	PASS
6	02 1341+ 241/FLA	145700	480	02 1341+ 241/FLA	145933	483	PASS	PASS
7	02 1349+ 241/FLA	150500	480	02 1349+ 241/FLA	150736	216	PASS	PASS
8	02 1357+ 241/FLA	151300	480	02 1357+ 241/FLA	151112	715	FAIL	PASS
9	02 1405+ 241/FLA	152100	480	02 1405+ 241/FLA	152307	352	PASS	PASS
10	02 1413+ 241/FLA	152900	480	02 1413+ 241/FLA	152859	211	PASS	PASS
11	02 1421+ 241/FLA	153700	480	02 1421+ 241/FLA	153230	1411	FAIL	FAIL
12	**02 1431+ 241/FLA**	**154500**	**420**				**AUTO**	**AUTO**
13	02 1440+ 241/FLA	155200	600	02 1431+ 241/FLA	155601	82	PASS	PASS
14	02 1450+ 241/FLA	160200	420	02 1440+ 241/FLA	155723	860	FAIL	FAIL
15	02 1457+ 241/FLA	160900	390	02 1450+ 241/FLA	161143	300	PASS	PASS
16	02 1504 241/FLA	161530	510	02 1457+ 241/FLA	161643	406	PASS	PASS
17	**02 1512+ 241/FLA**	**162400**	**540**	**02 1504 241/FLA**	**162329**	**1224**	**FAIL**	**FAIL**
18	**02 1520+ 241/FLA**	**163300**	**630**	**02 1504 241/FLA**	**162329**	**1224**	**FAIL**	**PASS**
19	02 1528 241/FLA	164330	360	/2 1512+ 241/FLA	164353	224	PASS	PASS
20	02 1533 241/FLA	164930	300	02 1520+ 241/FLA	164737	82	PASS	PASS
				02 1528 241/FLA	**164859**	**232**		
21	02 1543+ 238/FLA	165430	330	02 1533 241/FLA	165251	365	PASS	PASS
22	02 1545 241/FLA	170000	390	02 1543+ 238/FLA	165856	99	PASS	PASS
23	02 1551+ 241/FLA	170630	420	02 1545 241/FLA	170035	408	PASS	PASS
24	02 1600+ 241/FLA	171330	390	02 1551+ 241/FLA	170723	1174	FAIL	FAIL
25	**02 1606+ 241/FLA**	**172000**	**510**	**02 1606+ 241/FLA**	**172657**	**1012**	**FAIL**	**PASS**
26	**02 1615+ 241/FLA**	**172830**	**420**	**02 1606+ 241/FLA**	**172657**	**1012**	**FAIL**	**FAIL**

FIGURE 4. Wait Assessment Detail Results from Signal System Data: (a) Raw Results; (b) Description of Results as it Relates to the Wait Assessment Algorithm.

- Slots 1 through 11 indicate trains and corresponding scheduled headways matched to trains scheduled to provide service during those headways, as indicated by schd_train_id = matched_train_id. Even though matching is successful, certain slots fail wait assessment at the 25% standard. The `1325+ 241/FLA` train arrives approximately 4 minutes early, therefore over 10 minutes ahead of its follower. Compared to the scheduled 8 minute headway between trains, this headway fails under the 25% standard.
- Slots 11 through 13 indicate a large timeframe without any actual trains passing by (15:32:30 to 15:56:01). This causes slot 11 to fail wait assessment due to the higher than scheduled headway, and slot 12 fails because no train departed within the acceptable range of the algorithm.
- Beginning at slot 13, actual trains have drifted out of sync with the scheduled trains; however this does not necessarily cause the slots to fail Wait Assessment. As long as *any* actual train arrives within reach of scheduled train headway, the algorithm matches these trains because service is provided to the customer, even though it is not the specific train scheduled to arrive.
- Slots 17 and 18 indicate a repeat match, whereby only one train arrived within reach of two separate headway intervals. The repeat match allows a specific actual train to be matched to more than one scheduled train. Since scheduled headways change from train to train, WA 100% fails for the first slot, but WA 100% passes for the second slot. This feature of the algorithm allows more opportunities for passing credit to be given.
- Notice the `1528  241/FLA` train is not matched to any scheduled train. The train arrived slightly over a minute later than the previous train, and out of reach of the next scheduled arrival. This feature of the algorithm prevents giving credit to actual trains not providing service in a distinct service slot.
- Slots 25 and 26 indicate another repeat match. In this case, slot 25 passes wait assessment at WA 100%, therefore the following slot 26 must automatically fail wait assessment, to prevent a single train from being credited towards two distinct service slots. Extra trains within a given scheduled headway do not help the wait assessment metric.
- During the beginning of the rush hour, beginning around 1600 hours, the scheduled headway ranges from 5 to 10 minutes and varies greatly from interval to interval. Actual headway intervals are matched to scheduled headway intervals that are in effect at the time the train actually arrives.

From the results of this algorithm, line managers can identify strategies that have potential to improve service. First, it is clear that trains in slots 3, 4, and 5 were ahead of schedule, which negatively impacts performance even when later trains arrive on time. In this case, holding trains to scheduled departure times at key timing points along the trip can help alleviate the problem. Sometimes, during rush hours there are gaps in service that arise solely because of congestion and merging within the system, or due to an incident such as a sick passenger. In these cases, "gap" trains can be strategically placed within the system to be activated in service when such an incident occurred. The gap between slots 11 and 13 lasting nearly 24 minutes could have been partially alleviated by filling in service with a gap train. Additionally, during these scenarios, service could be rerouted or diverted from other routes/lines to provide service on the "2" route.

Daily Reporting & Operational Impacts

The availability of extensive data downloaded from the Automated Train Supervision (ATS) system provides 100% coverage for the IRT division (except "7" route) of NYCT's subway system and yields a much lower time-lag for compiling performance measures. This allows near-term corrective action by operations supervisors. Daily "outlier" reports are issued to assist managers to identify troublespots.

Figure 5(a) shows a daily summary report, with all information about a single train route's performance summarized on one page. For each hour and for each time-point location, the Wait Assessment result during that hour is given, as is the throughput (count of trains passing that location during that hour). This report allows a line manager to see at a glance how the route performed the previous day—and more importantly, if there were an incident on the route, what the performance impact was for that incident. Currently, only route-level results are calculated and reported to management on a daily basis. Track-level and corridor-level results have not been reported yet but NYCT is in the process of exploring this option.

Figure 5(b) is a slightly different daily report designed to improve troubleshooting. Without detailed knowledge of the day's incidents, it can be difficult to determine whether a lower Wait Assessment score during a certain hour was typical, and whether problems were ongoing and recurring or if they're due to a specific non-repeating incident. Operation managers generally prefer to focus on recurring problems rather than unusual incidents. This report compares today's Wait Assessment statistics (at the hourly and location level) with the rolling average over the past 30 days (where data is available). The "low outlier" hour-and-location combinations are printed out, allowing managers to take corrective action.

These reports are given to management by 10:00am the day thereafter and along with these reports the managers are provided with a list of the day's incidents so they can focus on both recurring problems (i.e. slow train operator) as well as a review of how dispatchers responded to an incident and what level of service was provided. This can be used to better inform supervisors how to respond to certain incidents in order to provide better service to passengers.

The 09:00 hour on the Southbound "1" route shows Wait Assessment scores in the 50-70% range for many of the stations in the middle to end of the route. These entries do not show up in the outlier report, and therefore are about similar to results for the previous 30 days. This implies a recurring problem that could perhaps be solved by schedule adjustments. On the other hand, the 07:00 hour on the Southbound "1" Train does show up on the outlier report, implying that there was an incident causing lower than expected performance results. Managers review incident reports and logs to determine whether the appropriate action was taken and how they might change the response in the future.

(a)

Daily Wait Assessment by Hour and Location DRAFT

Service Date **11/03/2011** *Line* ❶

Direction S **Hourly Results at Gap Locations (Wait Assessment 25%/Thruput)**

	AM Peak (0600-0859)			Off-Peak (0900-1559)							PM Peak (1600-1859)			Off-Peak (1900-2359)				
Station	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
51020 238 ST	86%/7	60%/10	89%/9	91%/11	92%/12	30%/10	90%/10	60%/10	100%/10	92%/12	67%/12	83%/12	67%/12	91%/11	80%/10	89%/9	100%/6	100%/5
51100 168 ST	67%/9	47%/15	77%/13	90%/10	75%/12	30%/10	90%/10	70%/10	100%/10	83%/12	87%/15	75%/12	75%/12	92%/12	70%/10	80%/10	100%/6	100%/5
51130 137 ST	67%/9	50%/14	64%/14	70%/10	75%/12	20%/10	100%/10	90%/10	90%/10	83%/12	79%/14	79%/14	58%/12	92%/12	70%/10	78%/9	100%/6	100%/6
51170 103 ST	86%/7	50%/14	61%/18	54%/13	67%/12	0%/10	100%/10	90%/10	70%/10	92%/12	64%/14	69%/13	58%/12	92%/12	82%/11	67%/9	100%/7	100%/5
51210 96 ST	71%/7	36%/14	72%/18	50%/12	83%/12	10%/10	100%/10	100%/10	80%/10	75%/12	73%/15	77%/13	83%/12	83%/12	80%/10	90%/10	83%/6	80%/6
51250 66 ST	86%/7	46%/13	72%/18	54%/13	83%/12	0%/10	100%/10	80%/10	80%/10	58%/12	79%/14	64%/14	67%/12	67%/12	80%/10	80%/10	86%/7	75%/5
51280 42 ST	88%/8	45%/11	84%/19	54%/13	92%/12	0%/11	100%/10	100%/10	90%/10	91%/11	86%/14	50%/14	67%/12	67%/12	82%/11	80%/10	83%/6	80%/6
51380 CHAM ST	71%/7	40%/10	71%/17	73%/15	67%/12	18%/11	50%/10	90%/10	90%/10	91%/11	77%/13	67%/15	67%/12	67%/12	73%/11	80%/10	75%/8	75%/6

Direction N **Hourly Results at Gap Locations (Wait Assessment 25%/Thruput)**

	AM Peak (0600-0859)			Off-Peak (0900-1559)							PM Peak (1600-1859)			Off-Peak (1900-2359)				
Station	**06**	**07**	**08**	**09**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
51380 CHAM ST	100%/6	50%/10	53%/15	82%/17	83%/12	55%/11	30%/10	100%/10	100%/10	100%/11	100%/12	73%/15	85%/13	83%/12	82%/11	100%/10	100%/9	100%/6
51280 42 ST	50%/6	88%/8	57%/14	56%/18	100%/12	50%/12	20%/10	80%/10	100%/10	90%/10	75%/12	60%/15	92%/13	83%/12	83%/12	90%/10	80%/10	100%/6
51260 59 ST	40%/5	67%/9	54%/13	67%/18	92%/13	45%/11	10%/10	80%/10	90%/10	82%/11	73%/11	67%/15	77%/13	83%/12	83%/12	82%/11	67%/9	100%/7
51210 96 ST	33%/6	86%/7	54%/13	61%/18	77%/13	67%/12	10%/10	70%/10	80%/10	70%/10	67%/12	86%/14	64%/14	83%/12	100%/12	80%/10	70%/10	83%/7
51130 137 ST	40%/5	71%/7	56%/9	53%/17	64%/14	75%/12	20%/10	60%/10	80%/10	70%/10	73%/11	86%/14	64%/14	75%/12	100%/12	82%/11	90%/10	50%/7
51100 168 ST	50%/4	63%/8	38%/8	59%/17	64%/14	42%/12	40%/10	60%/10	90%/10	70%/10	64%/11	93%/14	57%/14	75%/12	100%/12	82%/11	80%/10	86%/8
51070 DYCK ST	50%/4	86%/7	33%/9	60%/15	67%/15	67%/12	30%/10	50%/10	80%/10	80%/10	55%/11	100%/13	47%/15	75%/12	92%/12	83%/12	80%/10	83%/7
51030 231 ST	100%/4	71%/7	50%/8	58%/12	54%/13	58%/12	60%/10	30%/10	90%/10	60%/10	64%/11	92%/13	47%/15	92%/12	83%/12	75%/12	80%/10	86%/8
51020 238 ST	100%/4	71%/7	38%/8	58%/12	58%/12	55%/11	73%/11	30%/10	90%/10	82%/11	45%/11	83%/12	79%/14	82%/11	90%/10	58%/12	70%/10	86%/8

(b)

Daily Flash Report: Outliers* by Location by Hour DRAFT

Date 11/03/2011 *Sample Period Begins 10/05/2011*

Line and Service Type ❶ **Direction** S

Location	Time (Hour)	Today's WA25	Month Rolling Avg at This Location	# of Days
51100 168 ST	06 00	67%	91%	28
51130 137 ST	06 00	67%	90%	30
51210 96 ST	06 00	71%	89%	30
51020 238 ST	07 00	60%	88%	28
51100 168 ST	07 00	47%	78%	28
51130 137 ST	07 00	50%	89%	30
51170 103 ST	07 00	50%	85%	30
51210 96 ST	07 00	36%	86%	30
51250 66 ST	07 00	46%	83%	30
51280 42 ST	07 00	45%	81%	30
51380 CHAM ST	07 00	40%	87%	30
51210 96 ST	09 00	50%	77%	30
51020 238 ST	11 00	30%	89%	25

FIGURE 5. Wait Assessment Daily Flash Reports: (a) Report by Hour and Location; (b) Outlier Report Indicating Worst Performing Locations.

WAIT ASSESSMENT "REACH AND MATCH" ALGORITHM DETAIL

Before describing the wait assessment "reach and match" algorithm in more detail, the following notation is presented. Let:

I = Set of trains in the schedule having the same route identifier, direction, and timepoint location
J = Set of trains in the actual data, sorted by actual departure time
i = Current scheduled train from set **I** being processed
j = Current actual train from set **J** being processed
n = Pointer to actual train from set **J** to be used for matching
t_i = Scheduled departure time for train *i*
t_j = Actual departure time for train *j*
t_{Min} = Minimum (earliest) matching limit for scheduled departure time of train *i*
t_{Max} = Maximum (latest) matching limit for scheduled departure time of train *i*
Tag_j = Tag assigned to actual train *j* if it has been used in matching process
$Match_{i,j}$ = Array of tags assigned to the match between scheduled train *i* and actual train *j*
WA_i = Wait Assessment result for scheduled train *i*
T = Headway threshold for Wait Assessment to pass

(a)

(b)

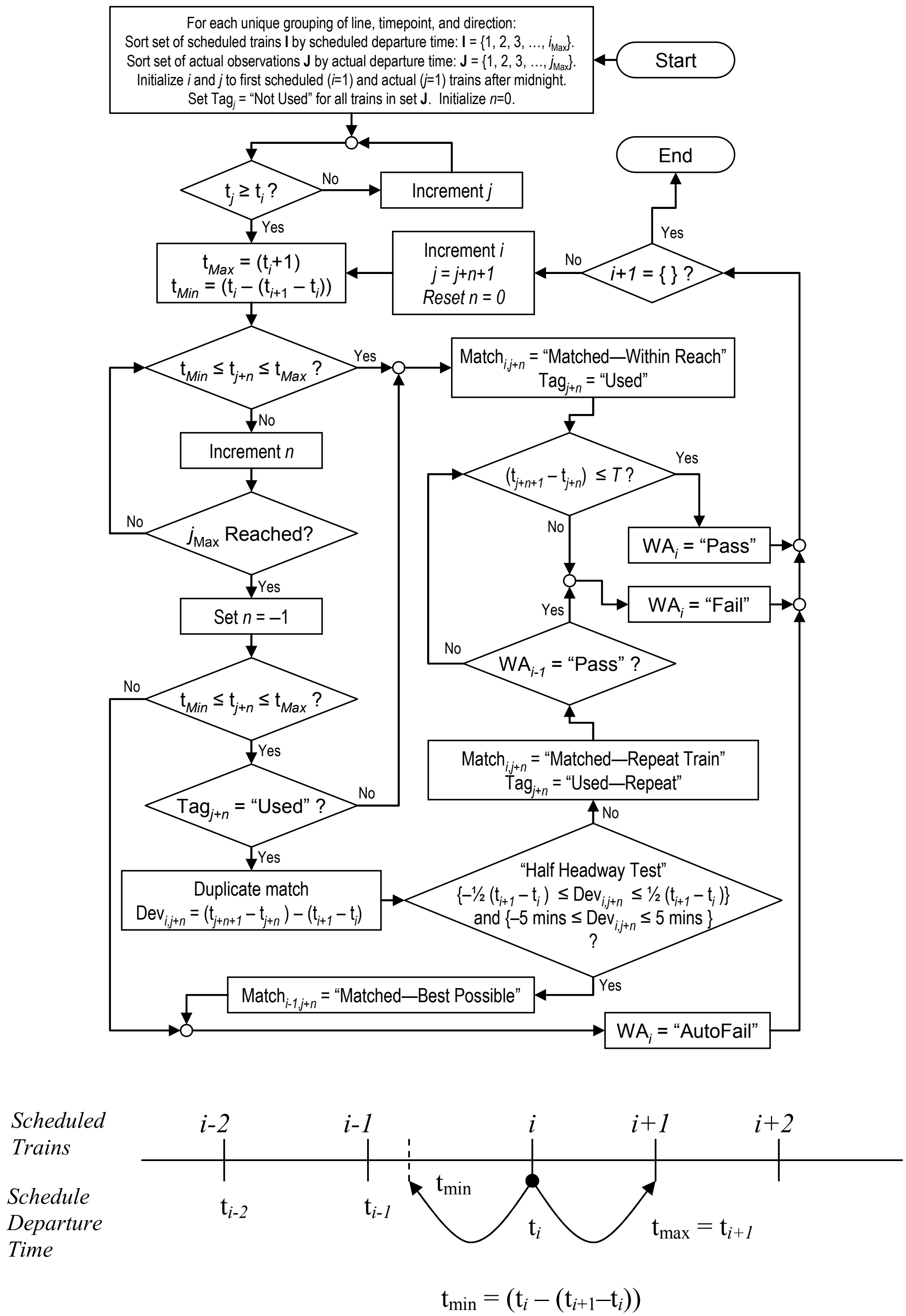


FIGURE 6. Wait Assessment Processing: (a) Flowchart of the Analytical Process; (b) Acceptable Matching Boundaries in the Wait Assessment Algorithm

Step 0: Input

For a given date, set **I** with corresponding t_i. For a given date, set **J** with corresponding t_j. Repeat the steps below for each unique grouping of route, timepoint, and direction.

Step 1: Sorting

Sort the set of scheduled trains **I** by scheduled departure time, such that **I** = {1, 2, 3, …, i_{Max}}. Sort the set of actual train observations **J** by actual departure time, such that **J** = {1, 2, 3, …, j_{Max}}.

Step 2: Initialization

Initialize *i* and *j* to the first scheduled and actual trains of the day, respectively. *i* is initialized to the first scheduled train after midnight (*i*=1). Set *j* = 1. Then, increment *j* until the actual departure time of the train is later than or equal to the first scheduled train after midnight, i.e. until $t_j \geq t_i$, where *i* = 1. Set *n* = 0. *n* is a pointer representing how far ahead or backward the algorithm looks to find a matching actual train. Set Tag_j = "Not Used" for all trains *j*. Perform steps 3 to 5 for each scheduled departure *i* of the route, direction, and timepoint group.

Step 3: Train Matching

Determine the acceptable matching boundaries. The maximum acceptable t_j for a match to *i* is (t_{i+1}), the scheduled departure time of the next train *i*+1, also called t_{Max}. The minimum acceptable t_j is $t_{Min} = (t_i - (t_{i+1} - t_i))$, the scheduled train departure time (t_i) minus the scheduled headway ($t_{i+1} - t_i$). The reason t_{Min} is not equal to t_{i-1} is because during transition periods between peak hour service and off-peak service, some of the transit authority's routes have headways that are somewhat irregular for a variety of operational reasons. To facilitate proper matching of these irregular headways, each train's acceptable matching boundary is based on the headway between itself and the following train, and not the prior train. This is illustrated in Figure 8(a). If a train falls within acceptable matching boundaries, it is denoted as "Within Reach."

Determine if the actual departure time of train *j* (t_j) is within acceptable matching boundaries and process the train accordingly:
a) If $t_{Min} \leq t_{j+n} \leq t_{Max}$ (i.e. train *j+n* is within acceptable matching boundaries for scheduled train *i*), then $Match_{i,j+n}$ = "Matched—Within Reach", and Tag_{j+n} = "Used".
b) If $t_{j+n} \geq t_{Max}$ or $t_{j+n} \leq t_{Min}$ (i.e. train *j+n* is not within acceptable matching boundaries), check all future trains for a potential match:

1. Increment *n=n+1* until train *j+n* satisfies $t_{Min} \leq t_{j+n} \leq t_{Max}$ or $j+n = j_{\text{Max}}$. If match is found, then $Match_{i,j+n}$ = "Matched—Within Reach"; Tag_{j+n} = "Used". Go to Step 4.

2. If no match is found, set *n* = –1 and check to see if train *j+n* satisfies $t_{Min} \leq t_{j+n} \leq t_{Max.}$ If yes:
   a. If Tag_{j+n} = "Used" this indicates actual train *j+n* <u>may be</u> matched to multiple scheduled trains *i*, i.e. Tag_{j+n} will be "Used—Repeat" and $Match_{i,j+n}$ will be "Matched—Repeat Train". To determine the best possible match, a "Half Headway" test is applied.
      i. Determine the scheduled headway ($t_{i+1} - t_i$) and the actual headway ($t_{j+n+1} - t_{j+n}$). Let the headway deviation ($Dev_{i,j+n}$) be the difference between scheduled and actual headways, i.e. ($t_{j+n+1} - t_{j+n}$) – ($t_{i+1} - t_i$).

ii. Determine if the headway deviation ($Dev_{i,j+n}$) is within ±50% of the scheduled headway ($t_{i+1} - t_i$), or 5 minutes, whichever is less. If $-\frac{1}{2}(t_{i+1} - t_i) \le Dev_{i,j+n} \le \frac{1}{2}(t_{i+1} - t_i)$ and $-5 \text{ min} \le Dev_{i,j+n} \le 5$ min, then the "Half Headway" test passes, and $Match_{i-1,j+n}$ = "Matched—Best Possible". Since only actual train $j+n$ fits within the acceptable train matching boundary, this actual train is a best match to previous scheduled train *i-1*, and since an actual train can only be credited once, this implies scheduled train *i* "Autofails". Set WA_i = "Autofail". Go to Step 4.

iii. If not, then Tag_{j+n} = "Used—Repeat" and $Match_{i,j+n}$ = "Matched—Repeat Train". This signals that although the previous scheduled headway *i-1* is a technical match, the current scheduled train *i* is a better match to the actual train $j+n$ being considered. Go to Step 4.

b. If Tag_{j+n} = "Not Used", then set $Match_{i,j+n}$ = "Matched—Within Reach"; Tag_{j+n} = "Used". This should never occur, since the algorithm works in increasing order of actual train observations **J**. Having to set $n = -1$ (go backwards in time to find a match) implies there is a shortage of trains, i.e. actual train throughput is lower than scheduled throughput, and the previous schedule departure should have matched this actual train as the algorithm looks predominantly ahead in time to find possible matches. However, pedantic implementation of this algorithm usually provides a check to ensure that every train is correctly matched. Go to Step 4.

3. If $t_{Min} > t_{j+n}$ or $t_{j+n} > t_{Max}$, then for all $t_{j*} \in \{t_{j-1}, t_j, t_{j+1}, \ldots, t_{j\text{Max}}\}$, $t_{j*} < t_{Min}$ and $t_{j*} > t_{Max}$ i.e. the departure times of all actual trains $j*$ are not within the acceptable train matching boundary for scheduled train *i*. This implies scheduled train *i* auto-fails: WA_i = "Autofail". There is no possible match to an actual train *j*, thus scheduled train *i* is determined to have automatically failed, by default. Go to Step 4.

Step 4: Result Calculation

After each "Match" result is computed (Matched—Within Reach, Matched—Best Possible, Matched—Repeat Train), a Wait Assessment (WA) result is calculated using the current scheduled train *i*.

1. If $Match_{i,j+n}$ = "Matched—Repeat Train" and WA_{i-1} = "Pass", then WA_i = "Fail". This is to prevent the same actual headway *j* for being credited against two scheduled headways i_1 and i_2. If WA_{i-1} = "Fail", then scheduled train *i* has the opportunity to pass Wait Assessment.
2. Otherwise, calculate Wait Assessment (WA). Recall *T* is the headway threshold by which Wait Assessment passes:
   a. If $(t_{j+n+1} - t_{j+n}) \le T$, then WA_i = "Pass". If the actual headway is less than or equal to the permissible threshold, then Wait Assessment is "Pass".
   b. If $(t_{j+n+1} - t_{j+n}) > T$, then WA_i = "Fail". If the actual headway is greater than the permissible threshold, signifying a gap in service, then Wait Assessment is "Fail".

Step 5: Increment Counter

If $i+1 \neq \{\ \}$ (i.e. there is another scheduled departure), increment $i = i+1$; $j = j+n+1$. Note that *n* may be negative or positive. After *i* and *j* are incremented, reset $n = 0$. Return to Step 3.

CONCLUSIONS

MTA New York City Transit (NYCT) has responded to the challenge of "doing more with less" by re-tooling its performance measurement frameworks to better capture performance from a customer's perspective, respond to management system improvement initiatives, and better incentivize operating decisions that deliver excellent customer service.

The "Reach and Match" algorithm is a crucial piece in the Wait Assessment calculation process. By applying a uniform standard across all trains and all routes, consistent reporting to the public is ensured. The algorithm takes the schedule into account but allows for flexibility due to "on-the-fly" changes that are made daily by dispatchers to improve service. Recent improvements to the WA standard have made it a more meaningful relative performance measure that is stricter for routes with more frequent service. The former pass/fail standard has now been replaced with by a distribution for failing headways that provides customers a more detailed view of system performance. NYCT is continuing to improve its performance standards by understanding how WA could fairly and best applied to shared-track territories where different routes can be treated as the same service corridor and train performance analyzed without reference to route designation. Upgrading a previous sample-based method that gathered limited data manually, extensive data was downloaded from the Automated Train Supervision (ATS) to provide 100% coverage and much lower time-lag for compiling performance measures. This allowed near-term corrective action by operations supervisors.

These improvements to NYCT's customer-centric service performance indicators were developed with extensive consultation with operations management, have been ratified by the MTA Board, and endorsed by stakeholders and public advocacy groups. In the tradition of improved reporting, NYCT will continue to explore new ways of assessing its own performance and reporting it for both internal diagnostic purposes and for public accountability. It is important for the Wait Assessment measure to be consistent with customer experience in the system. If a customer experiences worse than scheduled service, the Wait Assessment measure should drop accordingly.

Future Research

Future research involves applying the "Reach and Match" algorithm to bus operations using data from on-board GPS devices that track a buses' location. The difficulty in this endeavor is that buses do not necessarily arrive at a given stop in the same order as they leave the terminal, whereas this is the case for trains that travel on a fixed track. In New York City, many bus routes operate at frequencies higher than trains (*20, 21*), and wait assessment is a crucial performance measure (*3, 15*). Additional research focuses on determining the travel path of individual passengers, allowing us to compute a weighted waiting time measure that is more reflective of actual passenger experience.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions of Barry Greenblatt, Paul McPhee, Dan Mazzella, Herbie Lambert, Tom Calandrella, Bill Shrage, Theresa Cheung, Karl Steel, Leon Coakley, Nancy Yeoh, Bill Fitzgerald, the Subway On-Time Performance Task Force, Hercules Mack, Jean-Raymond Theobal, Chi Chan, Tanya Lipsmann, Aaron Berkovich, Larry Gould, Glenn Lunden, Patrick Diskin, John Cucarese, Anthony Cramer, H. Robert Menhard, and Steven Aievoli during the development phase of this algorithm.

REFERENCES

(1) Trompet, M., Liu, X., Graham, D, 2011. Development of Key Performance Indicator to Compare Regularity of Service Between Urban Bus Operators. *Transportation Research Board 90th Annual Meeting*, Washington, D.C., No. 11-0540.

(2) Lu, A., Aievoli, S., Ackroyd, J., Carlin, C., Reddy, A., 2009. Passenger Environment Survey: Representing the Customer Perspective in Quality Control. *Transportation Research Record 2112*, 93-103.

(3) Cramer, A., Cucarese J., Tran, M., Lu A., Reddy, A.V., 2009. Performance Measurements on Mass Transit: New York City Transit Authority Case Study. *Transportation Research Record 2111*, 125-138.

(4) Office of the Inspector General, State of New York, 1990. Regularity Indices for Evaluating Transit Performance. MTA/IG Technical Report 90-32. New York, N.Y.,

(5) Turnquist, M.A., Bowman, L.A., 1980. The effects of network structure on reliability of transit service. Transportation Research B 14, 79-86.

(6) Henderson, G., Kwong, P., Adkins, H., 1991a. Subway Reliability and the Odds of Getting There On-Time. Transportation Research Record 1298, 10-13.

(7) Adebisi, O., 1986. A Mathematical Model for Headway Variance of Fixed Bus Routes. *Transportation Research B* 20B(1), 59-70.

(8) Henderson, G., Kwong, P., Adkins, H., 1991b. Regularity Indices for Evaluating Transit Performance. Transportation Research Record 1297, 3-9.

(9) Furth, P.G., Muller, T.H.J., 2006. Service Reliability and Hidden Waiting Time: Insights from Automated Vehicle Location Data. Transportation Research Record 1995, 79-87.

(10) Wilson, N. H. M., Nelson D., Palmere A., Grayson T., Cederquist C., 1992a. Service Quality Monitoring for High-Frequency Transit Lines. *Transportation Research Record 1349*, 3-11.

(11) New York Public Interest Research Group—Straphangers' Campaign. State of the Subways Report Card, New York, N.Y. Retrieved from http://www.straphangers.org/statesub11/ on November 12, 2011.

(12) Permanent Citizens Advisory Committee to the MTA, 2011. Research Report—Minutes Matter: A Review of Performance Metrics at the MTA. New York, N.Y.

(14) Wilson, N. H. M., Macchi, R., Fellows, R., Deckoff, A., 1992b. Improving Service on the MBTA Green Line through Better Operations Control. *Transportation Research Record 1361*, 296-304.

(15) Wang, T. 2005. A Stakeholders' Approach To Bringing Bus Rapid Transit To New York City. *Master's Thesis, Queens College, City University of New York*.

(16) Reddy, Alla V., A. Lu, and T. Wang. Subway Productivity, Profitability, and Performance: A Tale of Five Cities. TRB Paper No. 10-0487. *Transportation Research Records 2143*, National Academies, Washington D.C., 2011.

(17) Reddy, Alla V., T. Chennadu, and A. Lu. Safeguarding Minority Civil Rights and Environmental Justice in Service Delivery and Reductions – New York City Transit Authority Title VI Program Case Study. TRB Paper No. 10-1155. *Transportation Research Records 2163*, National Academies, Washington D.C., 2011.

(18) Reddy, Alla V., J.A. Kuhls, and A. Lu. Measuring and Controlling Subway Fare Evasion: Improving Safety and Security at New York City Transit. TRB Paper No. 11-2016. *Transportation Research Records 2216*, National Academies, Washington D.C., 2012.

(19) Federal Highway Administration, California Division. New York City Automated Train Supervision (ATS). In Systems Engineering Guidebook for ITS, Version 3.0. Retrieved from http://www.fhwa.dot.gov/cadiv/segb/views/document/sections/Section8/8_5_1.htm on November 14, 2012.

(20) Carrel, A., R.G. Mishalani, N.H.M. Wilson, J.P. Attanucci, and A.B. Rahbee. Decision Factors in Service Control on High-Frequency Metro Line: Importance in Service Delivery. In *Transportation Research Records 2146*, National Academies, Washington, D.C.

(21) Beaton, Eric, J. Barr, J. Chiarmonte, T. Orosz, D. Paukowits, A. Sugiura. Select Bus Service on M15 in New York City: BRT Partnership Between New York City Department of Transportation and Metropolitan Transportation Authority New York City Transit. Transportation Research Records, Transit Volume 4, National Academies, Washington, D.C., 2012.

(22) Barr, Joseph E., E.B. Beaton, J.V. Chiarmonte, T.V. Orosz. Select Bus Service on Bx12 in New York City: Bus Rapid Transit Partnership of New York City DOT and Metropolitan Transportation Authority New York City Transit. *Transportation Research Records 2145*, National Academies, Washington, D.C., 2010.