TRB Manuscript # 09-1141

Performance Measurements on Mass Transit – New York City Transit Authority Case Study

*Anthony Cramer, John Cucarese, Minh Tran, Alex Lu, and Alla Reddy**
* Corresponding Author

Alla Reddy
Senior Director, System Data & Research (SDR)
Operations Planning
MTA New York City Transit
2 Broadway, A.17.92, New York, NY 10004
Tel: (646) 252-5662
Email: *Alla.Reddy@nyct.com*

Abstract 2
Introduction 3
Program Design 3
Route Selection 4
Measurement Standards 6
Headway Regularity 6
Wait Assessment (WA) 6
En-route Schedule Adherence (ESA) 7
Sampling 10
Confidence Levels and Accuracy 10
Sample Design 11
Data Processing 12
Sampling, Scheduling and Inventory System (SSIS) 12
Data Collection 12
Data Validation and Analysis 14
Reporting 16
Public Reporting 16
Internal Supplemental Reporting 16
Operational Uses of Performance Indicators 20
Scheduled Running Time Recalibration 20
Dispatcher Programs 20
Pilot Program Monitoring 21
Dispatching Strategy Evaluation 21
Data Mining 21
Summary 23
Future Work 23
Acknowledgements 24
References 25

Word Count: 248 (Abstract) + 4,990 (Text) + 10 × 250 (Figures) = 7,738 Words

ABSTRACT

For all organizations, be they public or private, it is essential to establish measurements ensuring that the services provided are being done well and when they are not, that the organization can diagnose problems. New York City Transit's (NYCT) mission is essentially to provide timely and reliable mass transit to over seven million daily riders. NYCT has established three main performance indicators to ascertain how closely they are meeting this mission: En-route Schedule Adherence (ESA), Headway Regularity, and Wait Assessment (WA). While ESA (–1 to +5 minutes) and WA is easily explained, Regularity (± 50%) is a useful diagnostic for operations management.

NYCT selected 23 subway routes and 42 borough-representative principal bus routes for performance analysis. A stratified sample, designed to prevent undesirable sample bias, is generated using a fully automated system and achieves an accuracy of 95 ± 5% at the route level.

Computerized data processing and analysis was implemented in 1995. Recently, paperless data collection was initiated, further decreasing reporting lag and improving data quality. Indicators are reported semi-annually to the public, while detailed internal diagnostic reports are issued frequently to assist operations management in improving service performance.

PI statistics are now used by senior management for setting goals and by rider advocacy groups to assess agency performance. A partnership and spirit of cooperation has developed between operating areas and analytical staff in troubleshooting delay issues and continuous quality improvement. The PI infrastructure is tapped by pilot program to assess performance of operations improvement initiatives.

INTRODUCTION

MTA New York City Transit (NYCT) carries over seven million passengers daily on its 6,388 subway cars and 4,576 buses. The system is geographically large, covering 321 square miles, and is diverse in both demographics and traffic conditions. The system also operates 24 hours, 7 days a week, even during construction projects and necessary system maintenance work.

NYCT's mission is "to provide customer with safe, reliable and convenient public transportation in a cost effective manner." To fulfill this mission, NYCT monitors the reliability of its operations through an extensive Performance Indicator (PI) program.

The PI program was established in 1994 in response to the MTA Inspector General's research (*1*) recommending the need for measures of service reliability other than the traditional Terminal On-Time Performance (OTP). Extensive research was conducted to understand transit service reliability from the passenger and the transit manager's perspectives (*2*), building on prior models of headway variance (*12*). The operational and public impacts of using different reliability indices was explicitly assessed using Monte Carlo simulations (*3*).

Statistical measures of service reliability, such as root-mean-squared average passenger wait time (*14*), were considered too complex for use as public measures. NYCT developed simplified version of the algorithms that are more easily understood. The result was the establishment of three main performance indicators: Headway Regularity, En-route Schedule Adherence (ESA), and Wait Assessment (WA). This design of the PI program achieves a duality of purposes:

1. To provide the public with measurements that are clearly defined, easily verifiable, readily understandable, and realistically represent the many factors that impact their riding experience, such that NYCT can be held accountable to its core mission.
2. To provide quantitative information to operating personnel that can be used to diagnose and correct service problems, and to improve overall performance.

PROGRAM DESIGN

NYCT essentially split the audience into two broad categories, internal operating departments and the external public at large. These two groups have quite different needs and expectations regarding monitoring system performance. To establish the indicators, NYCT Operations Planning consulted with both stakeholders.

For the public, focus groups were held to determine issues that were most important to the ridership. Measurements at odds with customer's perceptions of service provision (either too high or too low) would not be readily accepted. NYCT regularly shares its PI results with rider advocacy groups, thus verifiability and believability are important criteria. ESA and WA were designed to be easily understandable to the public.

For operations management, the PI program should supplement the operational reporting issued by Departments of Buses (DOB) and Subways (DOS). The information must be readily scalable to the system, division, route, timepoint, and survey assignment level. Importantly, operations management are assured that these statistical measurements are not a mechanism for grading *individual* dispatcher or operator performance, but are used as a resource for monitoring and improving service delivery. Regularity was primarily designed to be a diagnostic tool for internal use.

Based on ridership volumes and data collection cost constraints, NYCT selected 23 subway routes and 42 borough-representative principal bus routes for performance analysis. A stratified statistical sample, designed to prevent undesirable sample bias, is generated using a fully automated system and achieves an accuracy of 95 ± 5% at the route level.

Route Selection

Ideally, performance of all routes should be analyzed; however, this is costly and impractical. A sample of principal routes is selected to represent the greatest ridership, demographic, and geographic areas of the city.

Every NYCT subway route (except the Rockaway Park S Shuttle and Franklin S Shuttle) carried in excess of 80,000 revenue weekday riders in 2006. 6 is the busiest route with 700,000 weekday passengers. Together the 23 major routes carry in excess of five million riders on a weekday. Due to high passenger volumes, all 23 major subway routes and the Grand Central S Shuttle are monitored.

For buses, route selection was complicated because the 193-route NYCT bus system (excluding Express Bus) is too large to measure all routes. The system carries 2.3 million daily riders, more than 50% of which are on the top 45 routes whose weekday revenue ridership ranged between 60,000 and 16,000. The heaviest bus route was the M15/M15 Limited combination operating on 1 and 2 Avenues in Manhattan's Upper East Side.

Approximately 45 bus routes could be monitored with the available surveying resources. However, simply selecting the top overall ridership routes would exclude all Staten Island routes, and de-emphasize Queens. To ensure coverage in all five boroughs, top ridership routes by borough were selected. Additionally, certain key routes were added to the monitoring program because of their prominence (M27/M50) or importance to the community (M18, BX55, B63).

By combining the monitoring efforts for routes that share a trunk section, 42 borough-representative route groups (Table 1) were selected for monitoring. The groups (including 42 routes, 10 associated limited-stop routes, plus 22 route variants) represent just under half of all weekday bus ridership.

Routes indicated as "Limited" operate limited-stop service, essentially skipping lightly used stops to provide shorter running times. Limited buses generally operate exclusively during the day. Express bus services are not measured due to their generally narrow span of service (peak hours only).

Overall, the selected bus routes generate 25% of boardings compared to the subway system, but consume twice the survey resources. Routes with least ridership (and consequently low service frequencies) consume the most survey resources.

Table 1
Bus Routes Selected for Monitoring Under NYCT's Performance Indicator Program

Borough	Route	Description	2006 Weekday Revenue Ridership	System-wide Rank
Bronx	BX1/BX2, BX1/BX2 Limited	Grand Concourse/Melrose Av	42,531	5
	BX9	B'way/Kingsbridge Rd	27,119	18
	BX12, BX12 Select	Pelham Pkwy/Fordham Rd	42,410	6
	BX19	Southern Blvd/E 149 St	34,091	11
	BX36	E 180 St/E 174 St	30,416	12
	BX40/BX42	Tremont Av Crosstown	28,199	16
	BX41	Webster Av/White Plains Rd	29,110	14
	BX55 Limited	3 Avenue	16,942	41
Brooklyn	B6, B6 Limited	Bay Pkwy/Flatlands Av	43,374	3
	B15	New Lots Av	22,307	25
	B35, B35 Limited	Church Av/39 St	38,820	9
	B41, B41 Limited	Flatbush Av	40,951	7
	B44, B44 Limited	Nostrand Av	42,670	4
	B46, B46 Limited	Utica Av/Malcolm X Blvd	52,974	2
	B63	5 Av/Atlantic Av	15,332	53
Manhattan	M1	5 Av/Madison Av	16,241	46
	M2	5 Av/Adam Clayton Powell Blvd	15,544	52
	M3, M18	5 Av/St Nicholas Blvd Convent Av Shuttle	19,021 901	35 186
	M4	5 Av/Broadway	23,887	22
	M7	6 Av/7 Av/Lenox	16,855	42
	M10, M20	8 Av/Frederick Douglass Blvd 7 Av/8 Av/Hudson St	11,334 5,731	82 133
	M14A/M14D	14 St Crosstown	40,479	8
	M15, M15 Limited	1 Av/2 Av	59,708	1
	M27 M50	49/50 Sts Crosstown	4,027 4,580	153 146
	M31	57 St/York Av	14,923	57
	M66	66/67 Sts Crosstown	14,870	58
	M86	86 St Crosstown	29,036	15
	M101, M101 Limited M102 M103	3 Av/Lex Av/Amsterdam Av 3 Av/Lex Av/Lenox Av 3 Av/Lexington Av	37,607 18,400 16,516	10 38 44
	M104	Broadway/42 St	24,264	20
Queens	Q43	Hillside Av	15,936	49
	Q20 Q44/Q44 Limited	Main St/Union St Main St/Cross Bronx Expwy	22,962 15,002	24 55
	Q46	Union Tpk	21,737	29
	Q58	Fresh Pond Rd/Corona Av	27,363	17
	Q83	Liberty Av/Murdoch Av	9,957	88
	Q85	Merrick Blvd/Conduit Av	12,919	72
Staten Island	S44/S94	Richmond Av	7,389	113
	S48/S98	Forest Av	8,170	106
	S53/S93	Clove Rd/Verrazano Narrows Br	9,093	96
	S74/S84	Richmond Rd/Arthur Kill Rd	5,990	129
	S76/S86	Richmond Rd/New Dorp La	5,005	140
	S78	Hylan Blvd	7,104	115
	S79	Hylan Blvd/Verrazano Br	8,404	102

MEASUREMENT STANDARDS

Initially, two performance measures were established in 1994: Headway Regularity and En-route Schedule Adherence (ESA). Responding to rider advocacy group concerns that the Regularity indicator was confusing, an additional Wait Assessment (WA) indicator was implemented in 2000, to improve public understanding. In spirit, WA is a "Customer Bill of Rights" type measurement, representing the percentage of time-intervals in which customers did not endure an excessively long wait (i.e. that NYCT met the *reliability* component of its core mission).

The same raw data (fleet number and departure time of all vehicles passing a location) is collected for all three measurements, minimizing resource consumption. Indicators result from different analytical algorithms applied after data entry. All three indicators use data collected at en-route timepoints (rather than terminals), because the majority of riders enter and depart the system at intermediate stops along the route. The traditional Terminal On-Time Performance (OTP) is more appropriate for commuter rail, where most inbound riders disembark at terminal locations in the Central Business District (CBD). Also, various operating factors (such as terminal congestion) affect OTP in a way that does not represent most customers' perception of service.

Headway Regularity

Regularity essentially measures how evenly distributed actual bus or subway service is in relation to scheduled service. Bunched service (when vehicles closely follow one another) or gaps in service (widely spaced intervals between departures) are penalized by this measure. The tendency for buses and trains to form pairs is well documented in the literature (*4*). Regularity is designed in part to evaluate the extent of this problem, measured on a route-by-route basis. The Regularity indicator is defined in Table 2.

Regularity is measured when service is more frequent. During those periods, maintaining even headways is more important than adhering to predetermined schedules. Assuming random passenger arrival rates, maintaining even headways tends to minimize average passenger wait time and ensures even passenger loads (*5, 6*).

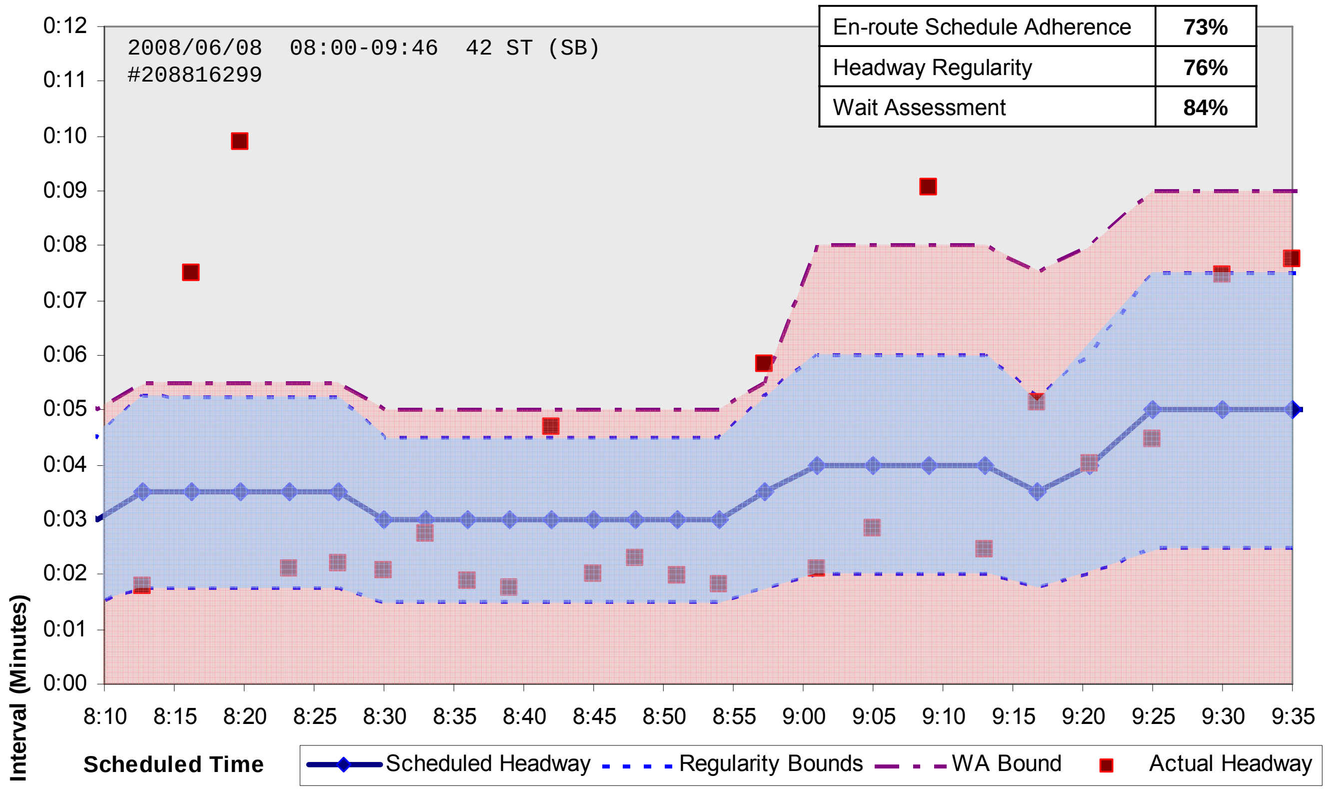
Each <u>scheduled</u> interval is matched to an actual observed headway. Only one observed headway is compared with each scheduled interval. If two actual headways (i.e. three consecutive departures, due to bunching) match one single scheduled interval, and neither could be matched to any other scheduled intervals, the latter observed headway is discarded and not credited. Scheduled intervals that contain no actual service are considered automatic failures ("autofails").

Figure 1 illustrates an example of analysis results presented graphically, listing scheduled and corresponding actual intervals. The blue shading represents the passing region for Regularity. Figure 2 illustrates the precise matching algorithm.

Wait Assessment (WA)

Replacing Regularity as the public *service reliability* measure, the WA standard sets a fixed threshold of what constitutes acceptable service, and enables passengers to hold the agency accountable. Bunching (providing too much service within a short span of time) and unevenly spaced service is not explicitly penalized in WA, only long service gaps trigger penalty, due to excessive wait. As defined in Table 2, a scheduled interval fails the WA Criteria if the time between departures is greater than the maximum permissible wait time (minutes above scheduled interval). In Figure 1, the red region passes the WA Criteria. Figure 2 also shows the WA algorithm.

Figure 1
Performance Analysis Graph



En-route Schedule Adherence (ESA)

ESA assesses timekeeping and schedule accuracy along en-route timepoints. While Regularity and WA analyze scheduled intervals as their basic unit of measurement, ESA assesses the performance of each scheduled trip as observed, independently of any other trip. Regularity and WA are relative performance measures, whereas ESA is an absolute measure. Bunching is penalized to the extent that not all bunched vehicles are adhering to schedule.

ESA was initially measured in the evening and overnight period (7 p.m. – 6 a.m.), when service frequency decreases and customers rely more on published schedules than simply waiting for the next available departure. In 2008, recognizing 10+ years of incremental improvements in service frequencies during the late evening hours (9 p.m. – midnight), WA reporting was extended to cover the entire day (6 a.m. – midnight) and ESA was de-emphasized as a public indicator. Figure 3 compares the three performance indicators measured by NYCT.

Table 2
Comparing NYCT Performance Indicator Standards

Indicator	Definition	Examples	Applicable Time Period	Indicator Publicly Reported
Headway Regularity	% of scheduled intervals passing the Regularity Criteria. A scheduled interval passes the Regularity Criteria if and only if it: 1. contain an actual vehicle departure (i.e. train or bus leaving the timepoint); and 2. the actual service interval between that departure and *the following* actual departure fall within ±50% or five minutes of the scheduled interval, whichever is less.	For a scheduled headway of 4 minutes, an actual headway of between 2~6 minutes (–50% to +50% of four minutes) would be permissible. For a scheduled headway of 15 minutes, an actual headway of 10~20 minutes (–5 to +5 minutes) would be permissible.	6 a.m. – 9 p.m. (expanded to 6 a.m. – midnight in 2008 internal to departments for diagnostics)	1994~2000
Wait Assessment (WA)	% of scheduled intervals that passing the WA Criteria. A scheduled interval passes the WA Criteria if and only if it: 1. contain an actual departure; and 2. the actual interval between that departure and *the following* actual departure is less than the time-period and mode dependent maximum acceptable wait times of scheduled headway +*n* minutes: Subway: +2 (Peak), +4 (Off-peak) Bus: +3 (Peak), +5 (Off-peak).	For a subway line with off-peak headway of 8 minutes, a maximum wait time of 12 minutes (8 +4 minutes) is permissible. For a bus scheduled to operate every 10 minutes during the peak period, the maximum allowable headway is 13 minutes (10 +3).	6 a.m. – 9 p.m. (expanded to midnight in 2008) Peak: 6 a.m. – 9 a.m. 4 p.m. – 7 p.m. Off-peak: 9 a.m. – 4 p.m. 7 p.m. – 9 p.m.	2000~present
En-route Schedule Adherence (ESA)	% of scheduled trips that passing the ESA Criteria. A scheduled trip passes the ESA Criteria if and only if the scheduled trip is observed in service between one minute before and five minutes after the scheduled time.	The regularly scheduled Manhattan-bound (A) train leaves the Far Rockaway Terminal in Queens at 22:11, and is scheduled to depart 125 Street (an intermediate en-route timepoint) northbound at 23:37. If this train is observed leaving 125 Street between 23:36:00 and 23:42:00, it is considered on-time.	7 p.m. – 6 a.m. (changed to 6 a.m. – midnight in 2008 internal to departments for diagnostics)	1994~2007

Figure 2
Headway Regularity and Wait Assessment Algorithms

Regularity — **Pass:** Y within Range **R**
Fail: Y out of Range **R**
Reg Criteria: *± 5 min of scheduled intervals equal to or greater than 10 min.*
± 50% of scheduled intervals that are less than 10 min.

Wait Assessment — **Pass:** Y ≤ **W**
Fail: Y > **W**
BUS: *+3 min peak (7am-9am; 4pm-7pm)*
+5 min off-peak (9am-4pm; 7pm-midnigh
SUB WAY: *+2 min peak (6am-10am; 3pm-7pm)*
+4 min off-peak (10am-3pm; 7pm-midnight)

INSTRUCTIONS

A= Scheduled time
B= Next scheduled time (sched. Max)
C= Actual time
D= Next actual time
A-(B-A)= Sched. Min

X= Scheduled time interval **(B-A)**.
Y= Actual time interval **(D-C)**.
R= ± 50% scheduled interval X **(0.5X to 1.5X)** **OR** **X ± 5 min.** for scheduled interval for X >10 min.
W= add 3 or 5 min to scheduled interval X **(X+3 or 5)**.

Auto fail: *There's no next actual time within scheduled range between A-(B-A) and B, or the system find an actual time but was already used and passed* ***(Both Reg. and Wait FAIL).***
Repeat match: *If the next actual time is not within A-(B-A) and B, and a previous actual time was already used and fail, it repeats the match.*
Skipped: *If an actual time does not fall within A-(B-A) and B, system skips that point.*
Computer Output (ATC Error Log) calculates a scheduled interval (X), an actual interval (Y), a range to review actual data (s.min [A-(B-A] to s.max [B]), and a pass/fail range at scheduled time.

Steps in setting up and calculating Regularity and Wait Assessment (Algorithm Detail).

1. Set up column A by lining up all the Scheduled Time within the Survey Start Time and End Time.
2. Set up column B by lining up just as column A with a next scheduled time A, and also skip all scheduled time with CH.MISS (such as comfort relief), such as on line 29 below.
3. Set up column A-(B-A) by calculating from columns A and B.
4. Select a first actual time in column C to fit into the a first available scheduled range between columns A-(B-A) and B.
5. Select a next actual time in column D (same row as column C).
6. Set up all the actual time in columns C and D by applying the rules of **auto fail, repeat match, and skipped**.
7. Continue selecting a next available actual time in column C (to fit in columns A(B-A) and B) and a next actual time in column D until the end of the survey data.
8. Calculate columns X, Y, R, W, REG, and WAIT.

Computer Output Generated from ATC (Regularity Error Log)

BUS	Sched	Prev	Next	Actual	Next	
	1657					sched rec not proc, sched_time < act_time[0]
	1700					sched rec not proc, sched_time < act_time[0]
1	1705	1657	1713	1705	1723	matche 8 18 N
2	1713	1711	1715	1723		sched rec auto fail
3	1715	1709	1721	1723		sched rec auto fail
4	1721	1713	1729	1723	1724	matche 8 1 N
5	1729	1728	1730	1724		is skipped 1726 is skipped 1745 sched rec auto fail
6	1730	1723	1737	1724	1726	matche 7 2 N
7	1737	1737	1742	1726		is skipped 1745 sched rec auto fail
8	1742	1740	1744	1726		is skipped 1745 sched rec auto fail
9	1744	1737	1751	1726		is skipped 1745 1747 matched 7 2 N
10	1751	1745	1757	1747	1749	matche 6 2 N
11	1757	1755	1759	1749		is skipped 1752 is skipped 1802 sched rec auto fail
12	1759	1753	1805	1749		is skipped 1752 is skipped 1802 1816 matched 6 14
13	1805	1803	1807	1816		sched rec auto fail
14	1807	1800	1814			act=180 repeat matche 7 14 N
15	1814	1813	1815	1816		sched rec auto fail
16	1815	1808	1822	1816	1826	matche 7 10 Y
17	1822	1819	1825	1826		sched rec auto fail
18	1825	1822	1828	1826	1828	matche 3 2 Y
19	1828	1821	1835	1828	1828	matche 7 0 N
20	1835	1834	1836	1828		is skipped 1829 is skipped 1830 is skipped 1840 sched
21	1836	1828	1844	1828	1829	matche 8 1 N
22	1844	1836	1852	1829		is skipped 1830 is skipped 1840 1840 matched 8 0
23	1852	1844	1900	1840		is skipped 1843 is skipped 1845 1901 matched 8 16
24	1900	1852	1908	1901	1902	matche 8 1 N
25	1908	1900	1916	1902	1904	matche 8 2 N

Algorithm Detail

BUS	A	B	A-(B-A)	C	D	X	Y	(R)		W	REG	WAIT
	16:57											
	17:00											
1	17:05	17:13	16:57	17:05	17:23	8	18	4	12	11	F	F
2	17:13	17:15	17:11	17:23	auto fail						F	F
3	17:15	17:21	17:09	17:23	auto fail						F	F
4	17:21	17:29	17:13	17:23	17:24	8	1	4	12	11	F	P
5	17:29	17:30	17:28	17:45	auto fail						F	F
6	17:30	17:37	17:23	17:24	17:26	7	2	4	11	10	F	P
7	17:37	17:42	17:32	17:45	auto fail						F	F
8	17:42	17:44	17:40	17:45	auto fail						F	F
9	17:44	17:51	17:37	17:45	17:47	7	2	4	11	10	F	P
10	17:51	17:57	17:45	17:47	17:49	6	2	3	9	9	F	P
11	17:57	17:59	17:55	18:02	auto fail						F	F
12	17:59	18:05	17:53	18:02	18:16	6	14	3	9	9	F	F
13	18:05	18:07	18:03	18:16	auto fail						F	F
14	18:07	18:14	18:00	repeat	repeat	7	14	4	11	10	F	F
15	18:14	18:15	18:13	18:16	auto fail						F	F
16	18:15	18:22	18:08	18:16	18:26	7	10	4	11	10	P	P
17	18:22	18:25	18:19	18:26	auto fail						F	F
18	18:25	18:28	18:22	18:26	18:28	3	2	2	5	6	P	P
19	18:28	18:35	18:21	18:28	18:28	7	0	4	11	10	F	P
20	18:35	18:36	18:34	18:40	auto fail						F	F
21	18:36	18:44	18:28	18:28	18:29	8	1	4	12	11	F	P
22	18:44	18:52	18:36	18:40	18:40	8	0	4	12	11	F	P
23	18:52	19:00	18:44	18:45	19:01	8	16	4	12	11	F	F
24	19:00	19:08	18:52	19:01	19:02	8	1	4	12	13	F	P
25	19:08	19:16	19:00	19:02	19:04	8	2	4	12	13	F	P

Figure 3
Hypothetical Example Comparing NYCT Performance Indicators

SAMPLING

The PI sample size is dictated by required accuracy and precision, and available survey resources. A random sample is automatically generated subject to stratification and evenness constraints to minimize undesirable bias. NYCT's combined sample creation and scheduling algorithm maximizes resource utilization while meeting statistical and operational requirements.

Confidence Levels and Accuracy

A 95 ± 5% confidence level was chosen for the base level statistics (by route) in each reporting period. Thus, a minimum performance change of 5% is required before NYCT can be 95% sure that change actually occurred. The binomial sampling formula was used for sample size determination:

$$n = \frac{Z^2 p(1-p)}{d^2}$$

where

n = sample size,
Z = confidence limit (1.960 for 95%, two-tailed),
p = probability of successful (passing) trips, and
d = expected margin of error (5%).

Assuming independent observations, this confidence level is achieved by sampling 384 bus or train departures per route in each reporting period. Sampling is done at a base probability level

(p = 0.50) to ensure a maximum and consistent sample size. A 10% ‘spare’ factor was added to account for missing or unusable data, raising the base per-route sample size to 422 observations.

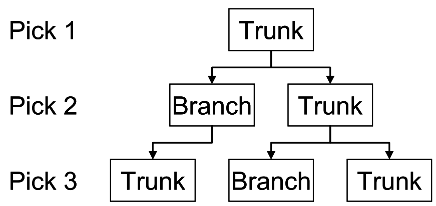
Route level data is aggregated to report borough, division, and systemwide results. Each route is equally represented, not weighted by passenger volume. 95 ± 2% confidence is achieved at the borough/division level, and 95 ± 1% for systemwide average.

Sample Design

The day was divided into five shifts, each representing a six-hour assignment for one surveyor (05:00–11:00, 08:45–14:45, 12:15–18:15, 16:00–22:00, and 19:30–01:30). Since surveyors report and return to a central location, the overlap in blocks accounts for travel time (approximately 1:45 per shift). Surveyors select one of the five fixed shifts in each ‘pick’, normally once a quarter. Once picked, fixed shift report times do not change, requiring a consistent number of assignments per day in each shift for staffing purposes. Vacation coverage is provided by ‘variable’ shift positions which can cover multiple fixed shifts as required. These ‘variable’ surveyors are assigned on a weekly basis depending on the week’s surveying needs.

Scheduled trips per shift (less travel time) are calculated for each time period-route combination using a travel time table and the schedule in effect. The assignments required per block to gather 422 observations are then determined. The sample is picked by route, location, and block. All timepoints are included in the sample pool. Since a pure random sample would allow skewing by direction or time of day, constraints are built-in to maintain sample evenness and address practicalities of data collection:

- Assignment blocks per month and per direction must be equivalent for each route.
- No assignment may be duplicated by day, location, block and direction.
- Total assignments per day must be equivalent in each shift.
- Subway assignments are divided into two halves of two hours each, to minimize surveyor detection.
- Multiple train lines are sampled together at shared track locations. Because of track sharing, trunk sections have a greater chance of being selected than peripheral locations, allowing the data to emphasize busiest parts of the system.
- Branch locations on a bus route are picked with less frequency than trunk locations, representing the overall ridership pattern. If a branch location is selected, a truck location must be selected for the next pick. If a trunk location is picked, either a branch or trunk location is selected next. This branching progression of picks, which interestingly creates a Fibonacci series, ensures that a route is represented primarily by trunk locations which service the maximum number of customers.



DATA PROCESSING

NYCT developed significant technical infrastructure to collect, analyze, and report on the performance indicators. The backend engine for collecting and analyzing data is the Automated Traffic Clerking (ATC) client-server system. Originally developed by MultiSystems in 1993, NYCT has substantially enhanced it over time to meet changing data collection needs. ATC data is housed on an Oracle 9i database, with a scheduled upgrade to Oracle 10g forthcoming.

A separate sampling, scheduling, and inventory system (SSIS) was developed in Visual Basic and SQL, embedded within a Microsoft Access database, interfacing with ATC via Open Database Connectivity (ODBC) links. Field data collection is scheduled, tracked and inventoried using these automated processes.

Sampling, Scheduling and Inventory System (SSIS)

The sampling program generates sufficient assignments to fully staff each of the five dedicated shifts for the quarter while also meeting the requirements for statistical accuracy. The scheduling module tracks field surveyors' leave balances and future availability, using this information to create daily work schedules. A reporting module creates forms (paper and electronic) for field data collection and for supervisors at the survey control desk. The inventory module tracks data from sample creation through analysis and returns missed or unusable assignments to the survey request pool for automatic re-scheduling.

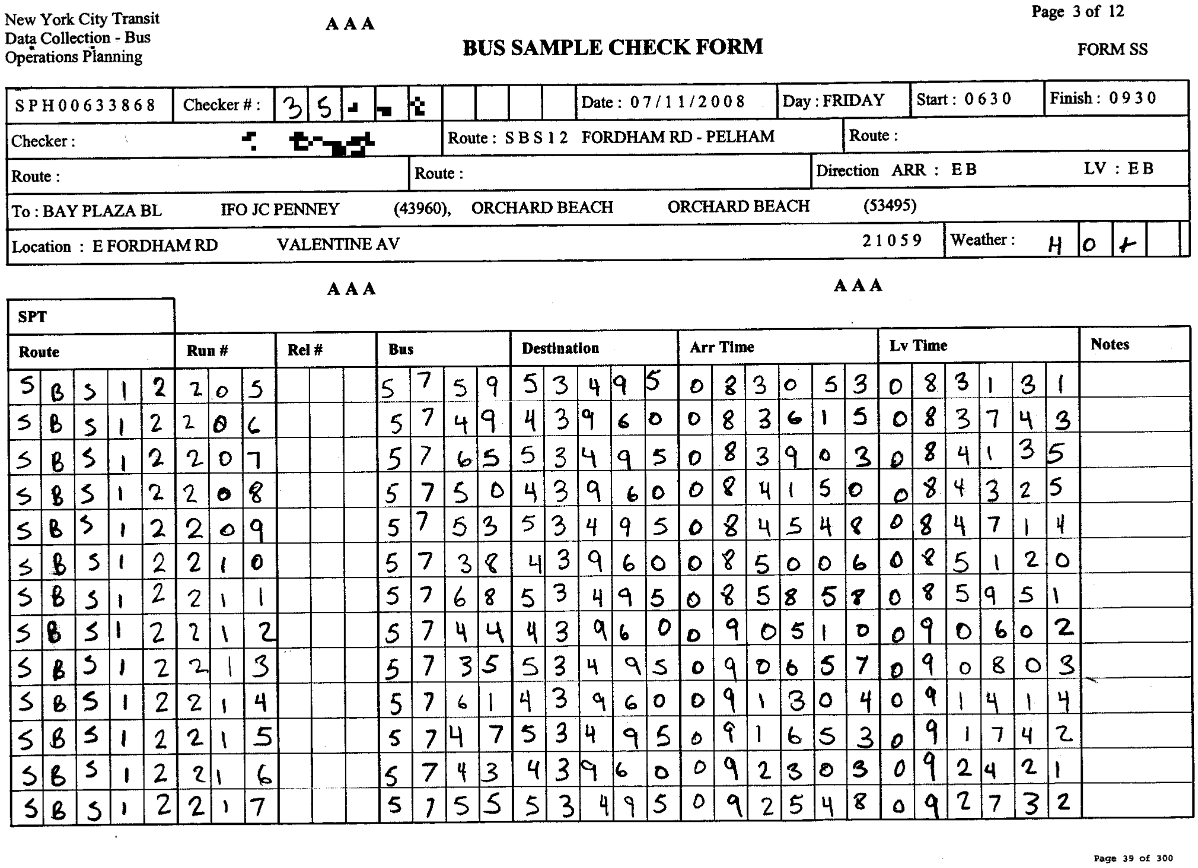
The application also provides route level and shift level sample completion rates, to permit the scheduling of extra work as required. Analysts use SSIS to monitor causes of surveyor absenteeism and missed or lost data. This information is invaluable in projecting future work completion rates. Supervisory personnel also use it to improve the efficiency of field data collection.

Data Collection

Traditionally PI data was manually collected by field surveyors using pre-printed data collection forms. Top part of Figure 4 shows the bus data collection form. The form for subways is similar. The pre-printed header information tells surveyor the location, routes, direction, and date to survey. Surveyors collect information for each bus/train they observe. Upon completion, forms are returned to a central location. Analysts verify the survey's location, date and time and the data is checked for blank fields, obvious gaps, and missing data. Incorrectly completed surveys are immediately rescheduled using the SSIS.

Forms were sent to an outside firm for data entry, converting the data into machine readable format. As data volumes increased, it became increasingly clear that the data entry contract was not cost efficient. NYCT purchased a heavy duty scanner and associated Optical Character Recognition (OCR) software. The software, specifically designed to read survey forms, allows for validation rules to be created, to easily translate scanned data into machine-readable format.

Figure 4
Comparing Scanned Paper Forms with PDA Screen Design



New York City Transit
Data Collection - Bus
Operations Planning

AAA

Page 3 of 12

BUS SAMPLE CHECK FORM

FORM SS

SPH00633868 | Checker # : 35 | Date : 07/11/2008 | Day : FRIDAY | Start : 0630 | Finish : 0930

Checker :

Route : SBS12 FORDHAM RD - PELHAM

Direction ARR : EB LV : EB

To : BAY PLAZA BL IFO JC PENNEY (43960), ORCHARD BEACH ORCHARD BEACH (53495)

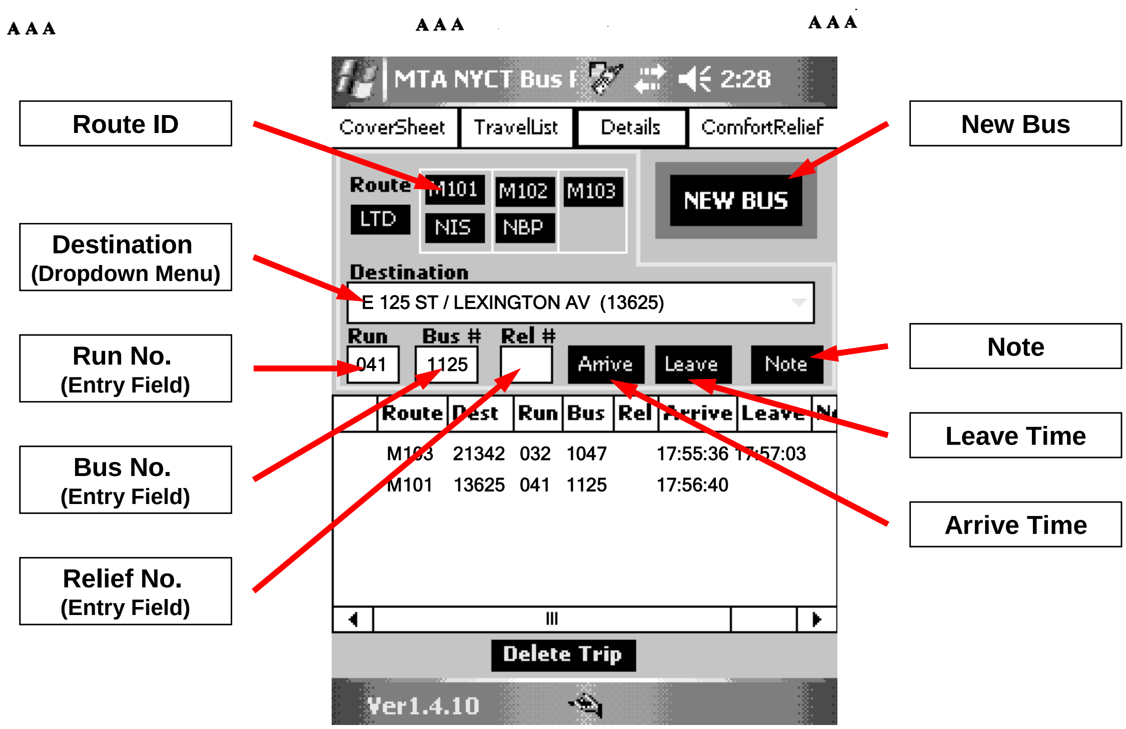
Location : E FORDHAM RD VALENTINE AV 21059 Weather : HOT

AAA AAA

SPT

Route	Run #	Rel #	Bus	Destination	Arr Time	Lv Time	Notes
SBS12	205		5759	53495	083053	083131	
SBS12	206		5749	43960	083615	083743	
SBS12	207		5765	53495	083903	084135	
SBS12	208		5750	43960	084150	084325	
SBS12	209		5753	53495	084548	084714	
SBS12	210		5738	43960	085006	085120	
SBS12	211		5768	53495	085857	085951	
SBS12	212		5744	43960	090510	090602	
SBS12	213		5735	53495	090657	090803	
SBS12	214		5761	43960	091304	091414	
SBS12	215		5747	53495	091653	091742	
SBS12	216		5743	43960	092303	092421	
SBS12	217		5755	53495	092548	092732	

Page 39 of 300



Paperless Data Collection (PDA Application)

In August 2008, paper forms for bus data collection were replaced with a Personal Digital Assistant (PDA) application. Surveyors are given Compaq iPAQ PDAs running an in-house application built using Microsoft Visual Studio 2005 Compact Framework for the front-end and an Oracle Lite database as data repository. Oracle Lite allows for a seamless integration with the ATC Oracle database. Once the PDA is cradled, data is automatically synched with the backend database for processing by the SSIS inventory system. Feedback from both surveyors and analysts has been extremely positive for the application. A similar application for subway data collection is currently under development. The bottom part of Figure 4 shows the main data collection screen for buses.

Data Validation and Analysis

Data validation and analysis is performed on the ATC system. The automated analysis process is different for buses and subway because significant differences exist in data collection methodologies. Prior to PDA implementation, bus data was scanned, validated, and imported into ATC for analysis. Today bus data is synched, validated, and then analyzed. Subway data continues to be manually entered and analyzed in a spreadsheet environment, and then imported into ATC for archiving and reporting.

Data Validation

Prior to accepting the imported data, the ATC software performs checks to guard against common data entry and import errors:

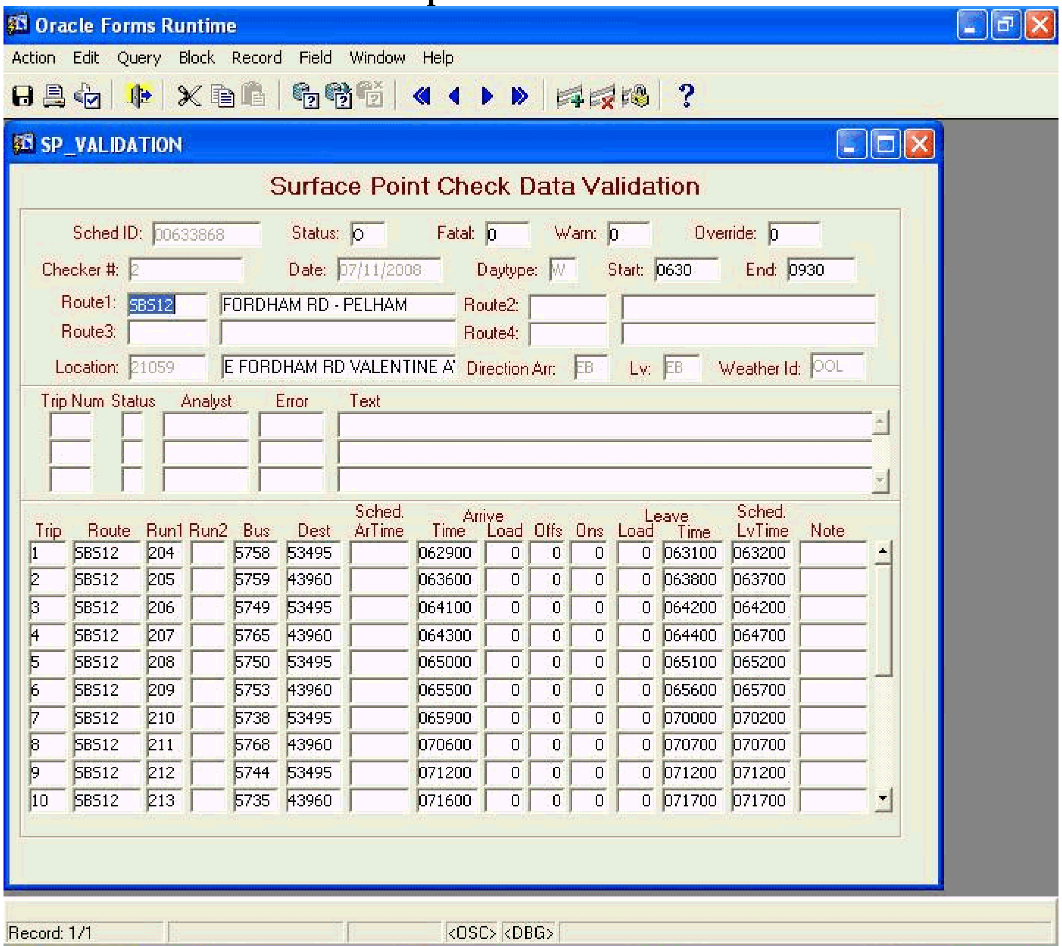
- **Field Validation:** While importing bus data, ATC performs rule-based validation. For example, 27:93:85 would be rejected as an invalid time value. PDAs eliminate this process by preventing invalid data entry.
- **Consistency Check:** ATC flags common errors such as times recorded out of order. Since paper forms are completed sequentially, the observed arrival and departure times should be recorded in ascending order. The majority of consistency check failures are the result of imperfect optical character recognition.

Bus Analysis

Also known as a completeness check, the first stage of data analysis in ATC matches actual observations (downloaded from PDAs) to the corresponding scheduled trip, using Run Number as key. Bus schedules are generally updated quarterly and schedules are automatically uploaded into ATC. Once matched, ATC automatically computes Regularity, WA, and ESA. Regularity and WA use the schedule data only to determine trip intervals. Performance is separately determined for each schedule record according to the measurement standards, and then tallied to determine the pass rate (ratio of passing schedule records to total).

The ATC application allows the analyst to view the automatically matched schedule and actual times for each trip. Figure 5 shows the main validation screen. For missed trips, analysts must make the determination whether surveyor missed the trip (performance should not be penalized) or the bus was actually not in service (results in performance penalty). Surveyors are instructed to note on data collection form when they have departed from location for any reason.

Figure 5
Oracle Completeness Validation Screen



Subway Analysis

The lack of Run Numbers on subway trains requires analysts to match the Lead Car Number to the daily train register to determine scheduled times. Train registers are manually recorded by train dispatchers at terminal locations. Since registers are not imported into ATC, the matching process is done manually on a spreadsheet, while entering actual observations and train register data. The completed analysis is imported into ATC, which is used to generate reports.

One additional complication is that subway lines often operate on a supplemental schedule. Supplements are issued to accommodate necessary construction. Since most track and signal maintenance is done overnight to minimize impact to passengers, the elimination of overnight surveys in 2006 meant that they impact very few surveys. The analysis process accounts for supplements officially issued in advance, but not schedule 'flexing' techniques used by train dispatchers to recover from severe service disruptions.

REPORTING

Initially, the performance indicators were published quarterly (with a two-month lag) in the NYCT Committee Agenda. Although the initial acceptance of the monitoring effort from operations management was slow, the program has developed into an integral resource for operations improvement, with a greater emphasis on timely troubleshooting and diagnostic capabilities. Although public reporting is semi-annual today, a variety of internal diagnostic reports are produced monthly.

Public Reporting

The PI program's main purpose is to monitor how well NYCT is providing service to the public (*9*). Figure 6 represents the 2007 Second Half subway results (*7*). WA and ESA are publicly reported at the systemwide, borough, and route level. These results are routinely used by rider advocacy groups for their annual rating of subway lines. Maintaining a transparent and accountable performance reporting process is critical to achieving public trust in the performance audit infrastructure (*8*).

Internal Supplemental Reporting

Years of performance data are stored in the ATC system, allowing detailed analysis of performance trends. Operational data is made available internally at a higher level of disaggregation, to facilitate NYCT corrective actions. Analytical reports developed in partnership with operating departments are issued regularly. Two examples are described:

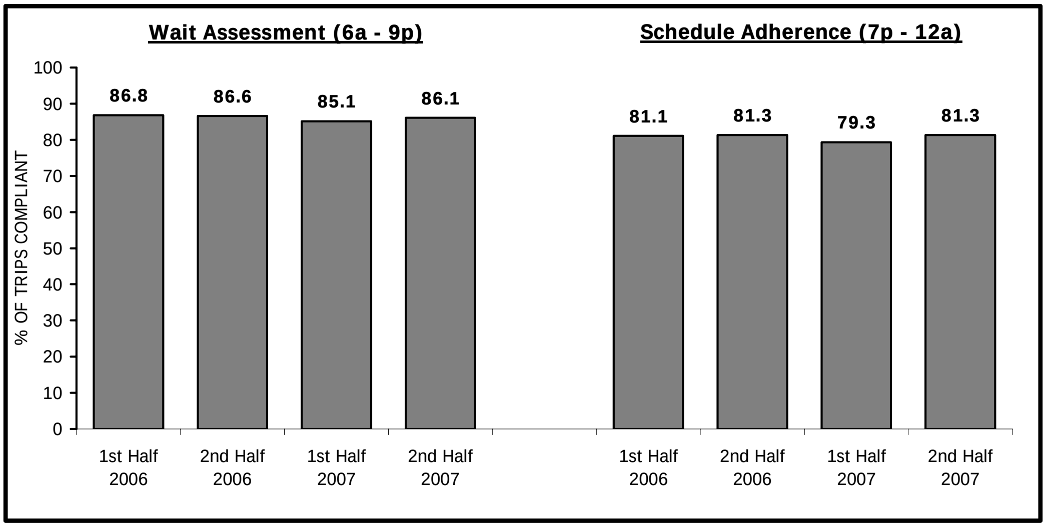
Route Profile Reports

Profile reports aggregate observations made over a longer time period (default is one year), to ascertain if specific en-route timepoints are consistently early or late. The longer timeframe is important, to gather sufficient observations for a meaningful trend analysis and to reduce the impact of unusual circumstances. Figure 7 presents a profile report route summary, showing the per-minute distribution of early and late trips by time period, timepoint, and direction. The histogram shows data at one specific timepoint. Operations management troubleshoots delay issues using this report.

Survey Level Diagnostic Reports

The survey level diagnostic report was developed at operating department request to identify performance declines. Operations management proactively reviews the results to determine major factors impacting overall performance. Distributed bi-weekly, this report shows observation date, location, direction, route, and the corresponding Regularity, WA, and ESA results for each survey assignment.

Figure 6
Subway Performance Indicators



Definition

Wait Assessment is measured during the day (6:00 a.m. – 9:00 p.m.), when service is relatively frequent. It is defined as the percentage of actual service intervals that are no more than the scheduled interval plus 2 minutes during peak (6 a.m. – 9 a.m., 4 p.m. – 7 p.m.) and plus 4 during off-peak (9 a.m. – 4 p.m., 7 p.m. – 9 p.m.).

Schedule Adherence, which is assessed in the evening (7:00 p.m. – 12:00 a.m.), is defined as the percentage of trips departing from all scheduled enroute timepoints between 1 minute before and 5 minutes after their scheduled departing time.

The results presented are for all subway lines.

2007 Annual Goals: Wait Assessment: 86.8% Schedule Adherence: 75.6%

Semi-Annual Results:

Wait Assessment		Schedule Adherence	
2nd Half 2007	86.1%	2nd Half 2007	81.3%
1st Half 2007	85.1%	1st Half 2007	79.3%
2nd Half 2006	86.6%	2nd Half 2006	81.3%
1st Half 2006	86.8%	1st Half 2006	81.1%

Discussion of Results: an increase/decrease of less than 1% is statistically unchanged.

2nd Half 2007 vs. 2nd Half 2006: The differences in "Wait Assessment" and "Schedule Adherence" are less than 1%.

Operating areas have responded favorably to this type of detailed reporting. It has also prompted operations management to become more closely involved, whether by allocating roving dispatchers to monitor problem areas, or by earmarking surveys for rescheduling. Surveys are generally rescheduled if documented circumstances outside NYCT's control (such as police or fire department investigations, emergency construction reroutes, or special events) adversely affected a route's performance on one specific day, because such events violate the sampling assumption of independent sequential observations. Surveys were rescheduled for everything from the July 18, 2007 underground steam pipe explosion near Grand Central to the March 14, 2008 police blockade and bus detours for a Presidential visit to the United Nations.

Figure 7

M15 - Local SERVICE -- KEY LOCATIONS USED IN ANALYSIS

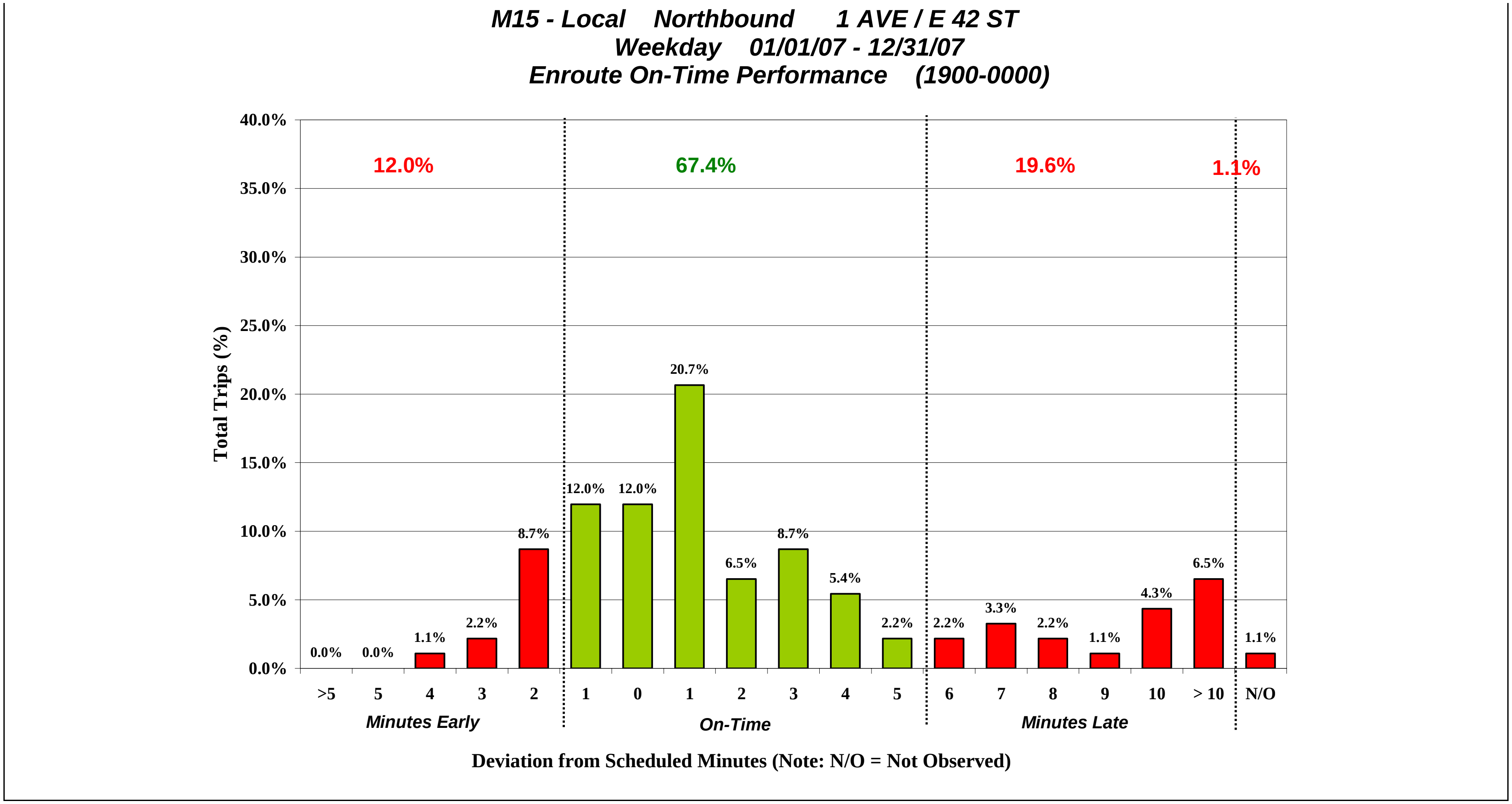
Deviation from Scheduled Minutes

-- 01/01/07 to 12/31/07 --

			Early					*On - Time*							*Late*							*TOTALS*		
Time	**Dir.**	**Location (s)**	>5	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	10	> 10	N/O	Early	On-Time	Late
700 to 1000	NB	1 AVE / E 1 ST	1.5%	0.0%	1.5%	3.0%	1.5%	9.1%	9.1%	16.7%	13.6%	6.1%	7.6%	4.5%	6.1%	1.5%	4.5%	4.5%	3.0%	4.5%	1.5%	7.6%	66.7%	24.2%
		1 AVE / E 14 ST	1.9%	0.0%	0.0%	0.0%	0.0%	5.6%	7.4%	20.4%	13.0%	3.7%	7.4%	7.4%	5.6%	11.1%	3.7%	0.0%	7.4%	1.9%	3.7%	1.9%	64.8%	29.6%
		1 AVE / E 42 ST to 1 AVE / E 57 ST	1.9%	0.0%	0.0%	3.8%	7.7%	15.4%	13.5%	5.8%	9.6%	3.8%	13.5%	0.0%	7.7%	1.9%	1.9%	3.8%	3.8%	5.8%	0.0%	13.5%	61.5%	25.0%
		1 AVE / E 79 ST to 1 AVE / E 97 ST	2.9%	4.3%	2.9%	14.3%	14.3%	14.3%	5.7%	11.4%	7.1%	2.9%	2.9%	4.3%	4.3%	1.4%	1.4%	0.0%	1.4%	2.9%	1.4%	38.6%	48.6%	11.4%
	SB	2 AVE / E 96 ST to 2 AVE / E 79 ST	0.0%	0.0%	1.7%	3.4%	0.0%	5.1%	13.6%	18.6%	11.9%	8.5%	8.5%	8.5%	5.1%	3.4%	3.4%	1.7%	1.7%	3.4%	1.7%	5.1%	74.6%	18.6%
		2 AVE / E 68 ST to 2 AVE / E 34 ST	0.0%	0.0%	0.0%	2.9%	8.6%	8.6%	5.7%	8.6%	10.0%	10.0%	7.1%	7.1%	7.1%	2.9%	4.3%	4.3%	2.9%	10.0%	0.0%	11.4%	57.1%	31.4%
		2 AVE / E 14 ST	0.0%	23.3%	6.7%	20.0%	6.7%	16.7%	13.3%	10.0%	0.0%	3.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	56.7%	43.3%	0.0%
		ALLEN ST / E HOUSTON ST	0.0%	0.0%	0.0%	2.0%	0.0%	2.0%	14.0%	20.0%	10.0%	6.0%	0.0%	6.0%	4.0%	6.0%	2.0%	2.0%	0.0%	26.0%	0.0%	2.0%	58.0%	40.0%
1000 to 1500	NB	1 AVE / E 14 ST to 1 AVE / E 23 ST	0.5%	0.5%	1.6%	2.7%	7.7%	7.7%	14.3%	12.6%	10.4%	8.8%	10.4%	4.4%	3.8%	3.3%	0.5%	1.6%	1.1%	4.9%	2.7%	13.2%	68.7%	15.4%
		1 AVE / E 57 ST	1.7%	1.7%	0.0%	0.0%	3.4%	8.5%	27.1%	18.6%	11.9%	3.4%	3.4%	1.7%	5.1%	0.0%	1.7%	3.4%	3.4%	3.4%	1.7%	6.8%	74.6%	16.9%
		1 AVE / E 79 ST	0.0%	0.0%	0.0%	3.1%	0.0%	0.0%	12.5%	6.3%	3.1%	0.0%	6.3%	12.5%	3.1%	3.1%	0.0%	3.1%	6.3%	37.5%	3.1%	3.1%	40.6%	53.1%
		1 AVE / E 97 ST	1.1%	0.0%	4.5%	4.5%	19.1%	19.1%	10.1%	5.6%	4.5%	5.6%	6.7%	6.7%	0.0%	2.2%	0.0%	1.1%	1.1%	3.4%	4.5%	29.2%	58.4%	7.9%
	SB	2 AVE / E 96 ST to 2 AVE / E 79 ST	0.0%	0.0%	0.0%	4.7%	9.3%	18.6%	7.0%	16.3%	2.3%	11.6%	14.0%	7.0%	4.7%	2.3%	0.0%	0.0%	0.0%	0.0%	2.3%	14.0%	76.7%	7.0%
		2 AVE / E 68 ST	0.0%	0.0%	0.0%	0.0%	4.9%	7.3%	7.3%	4.9%	4.9%	4.9%	14.6%	4.9%	7.3%	2.4%	4.9%	2.4%	2.4%	24.4%	2.4%	4.9%	48.8%	43.9%
		2 AVE / E 50 ST	0.0%	1.1%	3.4%	6.8%	5.7%	11.4%	17.0%	11.4%	10.2%	3.4%	3.4%	8.0%	2.3%	2.3%	1.1%	2.3%	2.3%	5.7%	2.3%	17.0%	64.8%	15.9%
		2 AVE / E 34 ST to 2 AVE / E 14 ST	5.3%	1.8%	1.8%	1.8%	1.8%	5.3%	15.8%	12.3%	3.5%	7.0%	3.5%	0.0%	1.8%	7.0%	1.8%	3.5%	8.8%	17.5%	0.0%	12.3%	47.4%	40.4%
		ALLEN ST / E HOUSTON ST to MADISON ST / ST JAMES PL	2.2%	1.1%	1.1%	1.1%	3.4%	7.9%	11.2%	12.4%	6.7%	4.5%	4.5%	7.9%	1.1%	9.0%	2.2%	0.0%	3.4%	19.1%	1.1%	9.0%	55.1%	34.8%
1500 to 1900	NB	MADISON ST / ST JAMES PL to 1 AVE / E 14 ST	1.5%	0.0%	1.5%	1.5%	8.8%	8.8%	13.2%	11.8%	11.8%	10.3%	7.4%	4.4%	7.4%	1.5%	2.9%	2.9%	0.0%	2.9%	1.5%	13.2%	67.6%	17.6%
		1 AVE / E 23 ST to 1 AVE / E 57 ST	1.1%	1.1%	0.0%	4.2%	2.1%	10.5%	12.6%	5.3%	6.3%	12.6%	4.2%	2.1%	1.1%	3.2%	3.2%	2.1%	4.2%	24.2%	0.0%	8.4%	53.7%	37.9%
		1 AVE / E 79 ST	2.0%	0.0%	2.0%	2.0%	2.0%	6.0%	10.0%	14.0%	4.0%	6.0%	4.0%	8.0%	6.0%	4.0%	4.0%	6.0%	2.0%	18.0%	0.0%	8.0%	52.0%	40.0%
		1 AVE / E 97 ST	2.3%	0.0%	0.0%	0.0%	7.0%	23.3%	23.3%	7.0%	2.3%	7.0%	2.3%	2.3%	0.0%	4.7%	2.3%	4.7%	0.0%	9.3%	2.3%	9.3%	67.4%	20.9%
	SB	2 AVE / E 79 ST to 2 AVE / E 50 ST	1.3%	3.9%	1.3%	0.0%	7.8%	9.1%	18.2%	7.8%	6.5%	5.2%	6.5%	6.5%	3.9%	1.3%	6.5%	3.9%	0.0%	9.1%	1.3%	14.3%	59.7%	24.7%
		2 AVE / E 34 ST to 2 AVE / E 14 ST	0.0%	2.0%	2.0%	0.0%	4.0%	2.0%	4.0%	4.0%	4.0%	2.0%	10.0%	2.0%	2.0%	8.0%	6.0%	2.0%	0.0%	42.0%	4.0%	8.0%	28.0%	60.0%
		ALLEN ST / E HOUSTON ST	0.0%	0.0%	0.0%	0.0%	1.8%	1.8%	1.8%	3.5%	5.3%	3.5%	3.5%	8.8%	7.0%	1.8%	3.5%	1.8%	1.8%	54.4%	0.0%	1.8%	28.1%	70.2%
		E BROADWAY / CATHERINE ST to MADISON ST / ST JAMES PL	0.0%	0.0%	5.0%	7.5%	5.0%	0.0%	2.5%	5.0%	10.0%	7.5%	7.5%	2.5%	0.0%	2.5%	5.0%	2.5%	5.0%	25.0%	7.5%	17.5%	35.0%	40.0%
1900 to 0000	NB	MADISON ST / ST JAMES PL	0.0%	0.0%	0.0%	0.0%	1.9%	5.6%	14.8%	14.8%	11.1%	14.8%	7.4%	11.1%	5.6%	0.0%	3.7%	5.6%	0.0%	1.9%	1.9%	1.9%	79.6%	16.7%
		1 AVE / E 1 ST	0.0%	0.0%	1.3%	0.0%	2.6%	14.5%	10.5%	18.4%	15.8%	5.3%	6.6%	3.9%	3.9%	3.9%	3.9%	0.0%	0.0%	2.6%	6.6%	3.9%	75.0%	14.5%
		1 AVE / E 14 ST	0.0%	0.0%	0.0%	0.0%	4.4%	10.3%	13.2%	10.3%	11.8%	8.8%	5.9%	8.8%	4.4%	1.5%	1.5%	2.9%	4.4%	10.3%	1.5%	4.4%	69.1%	25.0%
		1 AVE / E 23 ST	1.3%	0.0%	1.3%	3.8%	7.7%	11.5%	11.5%	14.1%	7.7%	2.6%	7.7%	2.6%	2.6%	7.7%	2.6%	2.6%	0.0%	10.3%	2.6%	14.1%	57.7%	25.6%
		1 AVE / E 42 ST	0.0%	0.0%	1.1%	2.2%	8.7%	12.0%	12.0%	20.7%	6.5%	8.7%	5.4%	2.2%	2.2%	3.3%	2.2%	1.1%	4.3%	6.5%	1.1%	12.0%	67.4%	19.6%
		1 AVE / E 57 ST	0.0%	0.0%	0.0%	2.6%	7.7%	7.7%	12.8%	7.7%	7.7%	15.4%	7.7%	7.7%	0.0%	12.8%	0.0%	2.6%	2.6%	5.1%	0.0%	10.3%	66.7%	23.1%
		1 AVE / E 79 ST	0.0%	0.0%	0.0%	0.0%	2.2%	6.7%	13.3%	13.3%	6.7%	13.3%	4.4%	2.2%	6.7%	4.4%	11.1%	4.4%	2.2%	8.9%	0.0%	2.2%	60.0%	37.8%
		1 AVE / E 97 ST	1.9%	0.0%	0.0%	5.7%	15.1%	15.1%	17.0%	11.3%	3.8%	1.9%	3.8%	1.9%	1.9%	1.9%	0.0%	0.0%	5.7%	13.2%	0.0%	22.6%	54.7%	22.6%
	SB	2 AVE / E 96 ST to 2 AVE / E 68 ST	0.0%	0.0%	1.5%	0.0%	3.0%	17.9%	23.9%	23.9%	6.0%	4.5%	3.0%	4.5%	1.5%	7.5%	1.5%	0.0%	0.0%	0.0%	1.5%	4.5%	83.6%	10.4%
		2 AVE / E 50 ST	0.0%	0.9%	2.6%	6.8%	7.7%	12.0%	15.4%	13.7%	6.8%	4.3%	1.7%	6.0%	3.4%	5.1%	3.4%	0.9%	0.9%	7.7%	0.9%	17.9%	59.8%	21.4%
		2 AVE / E 34 ST to 2 AVE / E 14 ST	0.0%	1.3%	2.6%	3.9%	11.7%	5.2%	13.0%	11.7%	13.0%	6.5%	2.6%	1.3%	5.2%	3.9%	1.3%	3.9%	0.0%	13.0%	0.0%	19.5%	53.2%	27.3%
		ALLEN ST / E HOUSTON ST to E BROADWAY / CATHERINE ST	1.6%	1.6%	1.6%	1.6%	1.6%	3.2%	9.7%	1.6%	14.5%	3.2%	14.5%	4.8%	4.8%	3.2%	8.1%	8.1%	1.6%	11.3%	3.2%	8.1%	51.6%	37.1%
		MADISON ST / ST JAMES PL	2.3%	2.3%	2.3%	10.2%	3.4%	12.5%	12.5%	10.2%	12.5%	8.0%	6.8%	3.4%	1.1%	0.0%	1.1%	3.4%	3.4%	4.5%	0.0%	20.5%	65.9%	13.6%

Light orange shade represents locations or group of locations with over 20% earlies or lates.
Dark orange shade represents locations or group of locations with an On-time performance equal to or below 50%.

Figure 7 (Continued)



OPERATIONAL USES OF PERFORMANCE INDICATORS

In its capacity as an oversight mechanism, the PI program achieves continual performance improvements through the usual goal-setting and management processes. Operations managers generally devote more energy to maintaining service regularity when independent performance measurements are used to set and monitor departmental goals. In addition to this auditing function, performance indicator data is now used more frequently to specifically identify operations issues.

Initially, there was resistance towards the PI program because of its monitoring role. With strong support and encouragement from NYCT's senior management, analytical and operating units forged a partnership. Field managers now routinely request and utilize performance data to evaluate their operation and to provide more efficient service. Special programs and initiatives use the PI framework to assess and modify service or to pinpoint problems. Some examples follow:

Scheduled Running Time Recalibration

Running times require periodic adjustments to reflect traffic conditions and to ensure it is achievable. Calibration of running times is an important part of transit scheduling (*10*). For the M15 (1/2 Avenues), profile reports indicated the first timepoint at 79 Street/2 Avenue had a large percentage of early southbound arrivals, while timepoints further south had a large percentage of late arrivals. If an en-route timepoint near the beginning of a route is consistently early while a subsequent one is consistently late, one conclusion that could be drawn is that there is room for improvement in the relative distribution of running time along the route. Running times were reallocated and schedule adjusted accordingly. As a result, ESA improved by 9% within one quarter.

Dispatcher Programs

NYCT initiates surface dispatcher programs in response to steep declines in performance indicators, or community concerns about service reliability. Extra dispatchers are allocated to specific routes to monitor field performance and recommend operational changes if necessary. One dispatcher program on a key Queen-Bronx interborough route led to a mid-pick rewrite of operating schedule (implemented as a supplement, under supervision by field dispatchers) that improved reliability substantially while the Bronx-Whitestone Bridge underwent unavoidable major construction work.

Another dispatcher program monitored outbound Manhattan-Staten Island express routes with a view to improve reliability. Analysis of special PI data concluded that traffic congestion had caused major delays for inbound daytime Staten Island-Manhattan express buses. This, combined with lack of holding or storage space in Manhattan, caused cascading outbound delays that could not be resolved by dispatchers alone. To improve performance, NYCT worked with MTA **Bridges and Tunnels** and New York City Department of Transportation (NYCDOT) to reconfigure street access to the Brooklyn-Battery Tunnel.

Pilot Program Monitoring

NYCT's planning and operations staff proactively implement new operating plans under pilot initiatives sponsored by community leaders, line managers, and the capital program. The PI framework is used to monitor the impact of these pilot programs or major changes in operations. Recently, NYCT implemented the BX12 +select**bus**service in conjunction with NYCDOT. The BX12 is the first bus rapid transit (BRT) corridor in NYCT's history, with off-board fare collection, dedicated bus lanes and bus fleet, and traffic signal priority. In partnership with operations management, bus planning, and bus schedule units, PI data is currently being collected and analyzed to monitor the reliability of this new service.

NYCT is currently adapting automatically-collected signal system data for monitoring the impact of trial service improvement packages on the subway system. PI data and methodologies were part of an overall evaluation plan to determine the efficacy of corridor-based scheduling initiatives on the Lexington Avenue Line, a new platform-docking plan at a major terminal, and increased service frequency resulting from a signal capacity capital project.

Dispatching Strategy Evaluation

Poor performance on rapid transit lines typically relates to congestion, construction delays, and unusual incidents. Occasionally, PI statistics alert line management to suboptimal dispatching strategies. Early train departures, particularly in the evening and off-peak, can result in missed connections, uneven spacing, and generally poor performance. After reviewing PI profile reports showing many early trains during off-peak hours, the district management instituted an initiative to hold all trains to time at critical timepoints (i.e. train regulation), even when hold-to-time is not required operationally. This resulted in improved service reliability, which was substantiated by the performance indicators post-implementation.

Data Mining

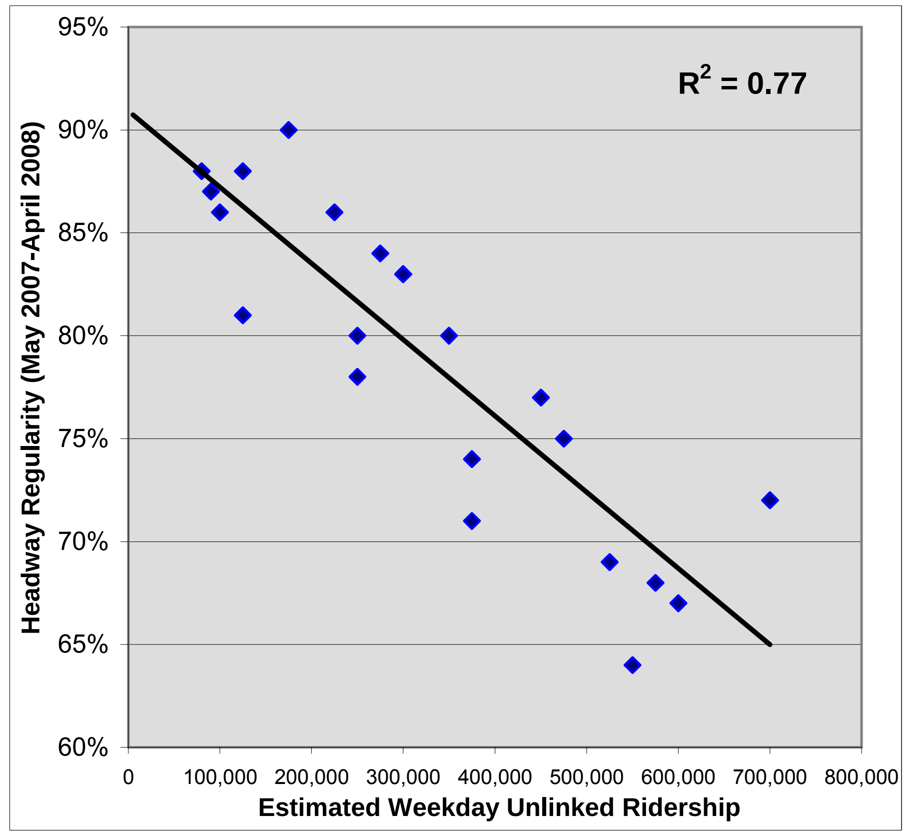
The aggregate PI results are also subject to data mining in search of non-operating factors affecting transit route performance. Figure 8 shows one example of such analysis.

Declines in ESA and other measures of reliability have often been attributed to management actions and internal factors. However, this analysis demonstrates that unlinked ridership (train embarkations) explains 77% of the variation in Regularity as measured by NYCT's standards. The effect of train delays is multiplicative as ridership increases. Trains that become crowded from high ridership incur longer dwell times and induce unproductive passenger behavior such as door holding (*5*). These effects combine to degrade performance and exacerbate train bunching as soon as slight perturbations in running times occur. Worse still, higher frequency corridors also require more exact timekeeping to prevent congestion and maintain throughput.

NYCT has experimented with many different methods of mitigating reliability issues resulting from increased ridership, including the use of platform conductors, service designs to achieve even passenger distribution, publicity campaigns to encourage behavior modification, and programs directly addressing most common delay-causing vehicle factors. Other properties have even proposed measures as drastic as removing seats to relieve overcrowding and maintain better performance (*11*).

Figure 8
Correlation Between Unlinked Ridership (Train Embarkations) and Headway Regularity

Line	Description	Estimated Weekday Ridership	Headway Regularity 5/07-4/08
6 6	City Hall/Pelham Bay Park	700,000	72%
1	South Ferry/Van Cortlandt Park 242 St	600,000	67%
F	Coney Island/Jamaica 179 St	575,000	68%
4	Utica Av Crown Heights/Woodlawn	550,000	64%
A	Far Rockaway/Inwood 207 St	525,000	69%
E	World Trade Center/Jamaica 179 St	475,000	75%
7 7	Times Square 42 St/Flushing Main St	450,000	77%
2	Flatbush Av/Wakefield 241 St	375,000	74%
5	New Lots Av/Nereid Av	375,000	71%
R	Bay Ridge 95 St/Forest Hills 71 Av	350,000	80%
B	Brighton Beach/Bedford Park Blvd	300,000	83%
D	Coney Island/Norwood 205 St	300,000	83%
N	Coney Island/Astoria Ditmars Blvd	275,000	84%
3	New Lots Av/Harlem 148 St	250,000	78%
L	Canarsie Rockaway Pkwy/8 Av-14 St	250,000	80%
Q	Coney Island/7 Av-57 St	225,000	86%
J Z	Broad St/Jamaica Parsons-Archer	175,000	90%
C	Euclid Av/Washington Heights 168 St	125,000	88%
V	2 Av-Houston St/Forest Hills 71 Av	125,000	81%
G	Smith-9 Sts/Forest Hills 71 Av	100,000	86%
S	Times Square 42 St/Grand Central	90,000	
W	Whitehall St/Astoria Ditmars Blvd	90,000	87%
M	Bay Pkwy/Middle Village Metropolitan Av	80,000	88%
S	Prospect Park/Franklin Av	18,000	
S	Rockaway Park/Broad Channel	5,000	



Regardless, the only long-term solution to maintaining reliable performance is to provide sufficient transportation capacity where and when it is needed, and ensuring that crowding does not exceed efficient design levels (*5*). The Second Avenue Subway project, currently under construction, will likely result in substantial performance improvement on the Lexington Avenue Subway 456 once complete.

SUMMARY

NYCT monitors three major performance indicators: Headway Regularity, Wait Assessment, and En-route Schedule Adherence, to provide both their employees and the public with unbiased performance results. This has resulted in:

- Rider advocacy groups adopting NYCT's performance measures as its benchmark of agency performance
- Senior management setting goals for operating areas on the basis of the PI program, supplanting traditional self-reported statistics, providing transparency and public accountability
- Partnership and a spirit of cooperation between Operations Planning and field management in troubleshooting delay issues and continuous quality improvement, with a greater emphasis on the use of data
- Development of extensive technology infrastructure to collect, process, analyze and report large volumes of performance data
- Operations improvement initiatives have tapped the PI infrastructure to monitor the performance of pilot service plans, dispatcher programs, and dispatching strategy

The PI program's evolution demonstrates the long developmental process required to establish an impartial and accountable performance measurement system. Commitment from top management and inter-departmental cooperation is critical in the planning, development and implementation of such a large and multifaceted program. Careful thought must be given to what, when, and how to measure. The resulting indicators should be attainable, understandable, operationally useful and informative to the public.

Future Work

NYCT is committed to continually improving the quality, speed, and accuracy of performance reporting. The next steps for NYCT's performance measurement program include:

- **Full PDA Deployment:** As NYCT gains experience with its paperless data collection applications, further development may allow these applications to be modified for other types of data collection, including subway PI, surface ridechecks, and subway load and dwell time checks.
- **Expanded Use of Automatically Collected Data:** NYCT anticipates more extensive use of automatically-collected data. Algorithms are being developed to correctly interpret signal systems data for performance and run-time monitoring. Preliminary work suggests that toll-road transponder data may be used for monitoring express bus performance.

- **Additional Data Sources:** Automated passenger counter, passenger load sensor, automatic fare collection, and automated vehicle locator data could be used to monitor route compliance and ridership, resulting in additional indicators.
- **Decreased Reliance on Sampling:** Automatic data collection produces extensive data, which may in future allow entire data populations to be evaluated. Work is underway to explore the processing methods and technology that will be required to handle such large amounts of data.

Although much progress has been made, NYCT will continue to investigate new ways to provide New York City with the best transit system in the world.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the assistance of following New York City Transit colleagues during the preparation of this paper: Steven Aievoli, Aaron Berkovich, Wan Chen, Faina Gorokhovskiy, and Michael Kelly. Thanks to Ginny Siggia, Massachusetts Institute of Technology for research assistance. Responsibility for errors or omissions remains with the authors. Opinions expressed in this paper are those of the authors and do not necessarily reflect official policy of Metropolitan Transportation Authority or MTA New York City Transit.

REFERENCES

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

(2) Henderson, G., P. Kwong, H. Adkins. Subway Reliability and the Odds of Getting There On-Time, *Transportation Research Record* **1297**, pp.10-13. (1991)

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

(4) Chapman, R.A. and J.F. Michel. Modeling the Tendency of Buses to Form Pairs. *Transportation Science* **12**, 2, pp.165-175. (1978)

(5) Office of the Inspector General, State of New York. *Measuring New York's Subway Performance During The Rush Hour – An Analysis of the Relationship Between Overcrowding and Headway Evenness and Its Use in Estimating On-Time Performance*. Task 8 of the Passenger Oriented Subway Performance Model, FTA Grant NY-08-0185. New York, N.Y. (September 10, 1992).

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

(7) New York City Transit. *Standard Follow-Up Reports: Service Quality Indicators*, in New York City Transit Committee Meeting Agenda. New York, N.Y. (2008).

(8) Ryus, P. *TCRP Research Results Digest, Number 56*. Summary of TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System. Washington D.C. (2003).

(9) Kittelson & Associates Inc. *TCRP G-6 Interim Report:* A Guidebook for Developing a Transit Performance Measurement System. Washington D.C. (2002).

(10) See, for example, Ceder, A. *Public Transit Planning and Operation: Theory, Modeling, and Practice*. Butterworth & Heinemann/Elsevier, Burlington, Mass. (2007)

(11) Hilkevitch, J. and D. Williams-Harris. CTA to Experiment with Seatless Rail Cars – Brown Line Likely Spot for Trial Run, *Chicago Tribune*, July 17. Chicago, Ill. (2008).

(12) Adebisi, O., A Mathematical Model for Headway Variance of Fixed Bus Routes. *Transportation Research B*, **20B**, No. 1, pp 59-70 (1986)

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