RAIL TRANSIT OPERATIONS ANALYSIS: FRAMEWORK AND APPLICATIONS

by

Adam B. Rahbee
B.S., Northeastern University (1999)

SUBMITTED TO THE DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN TRANSPORTATION

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2001

Copyright © Massachusetts Institute of Technology 2001
All rights reserved

Signature of Author
Department of Civil and Environmental Engineering
August 13, 2001

Certified by

Professor Nigel H.M. Wilson
Thesis Supervisor

Accepted by

Oral Buyukozturk
Chairman, Departmental Committee on Graduate Students
RAIL TRANSIT OPERATIONS ANALYSIS: FRAMEWORK AND APPLICATIONS

by

Adam B. Rahbee

Submitted to the Department of Civil and Environmental Engineering on August 13, 2001 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Transportation

ABSTRACT

This thesis investigates strategies to improve service quality performance on a rail rapid transit line. Using data sources now commonly available, such as databases in Automatic Transit Supervision systems, an approach is suggested that uses available data and continuing input from agency personnel. The approach aims to draw out local operational and institutional conditions affecting service quality performance, and aims to identify improvement strategies that are likely to be implemented.

There are three ways of defining operational problems. The first is a failure in meeting performance measures. The second is a failure of agency staff in making decisions given agency objectives and constraints. Areas of decision-making are defined as line characteristics, operating plan and service management. The third way of defining an operational problem is when agency objectives are problematic for overall good performance.

The recommended methodology for studying rail transit line performance attempts to highlight problems of the second and third types. Since the operational practices of rail transit lines are not currently well documented, a rich description of line operations is needed to identify these problems.

Because agency objectives and constraints that enter into decision processes are complex, an analyst must be able to draw them out of other agency staff. Although this may be accomplished through interviews alone, it may be more useful to prepare representations of system behavior (such as time-space plots or other illustrations) and discuss these specific cases with key staff in interviews. In this way, agency staff can provide more detailed information to the analyst to document and use.

This methodology was applied in a case study of the Massachusetts Bay Transportation Authority Red Line. The study resulted in series of recommendations, one of which has led to a reduction in passenger waiting time without any additional costs being incurred. Further modifications to the operating plan as warranted by the findings have the potential to further reduce passenger waiting time at no operational or capital cost. Reconfigurations to the signal design may reduce minimum headways significantly.

Thesis Supervisor: Nigel H.M. Wilson
Title: Professor of Civil and Environmental Engineering
Acknowledgements

The author would like to thank several individuals. In particular, Nigel H.M. Wilson of MIT and Peter G. Furth of Northeastern University. This research was fostered through Anne Herzenberg, Chief Operating Officer of the Massachusetts Bay Transportation Authority (MBTA). Several people at the MBTA who helped significantly were Henry Mar, Glenn Camelien, and Frank Feltes of the MBTA OCS, O Nna and Bob Swirbalus of the MBTA Signal Department, Paul Strobis and Robert Gatulis of the Scheduling Department, Maureen Shirkus and Sean McCarthy of the OCC, Red Line Chief Tony Kinehan, former Red Line superintendent Joseph Casey, the Red Line train starters, and Jonathan Belcher of CTPS. Special thanks to Patrick Amory and Bill Andretti for their editorial comments.
**TABLE OF CONTENTS**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1: Introduction</td>
<td>5</td>
</tr>
<tr>
<td>1.1 Background and Motivation</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Prior Research</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Research Approach</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Thesis Content and Organization</td>
<td>14</td>
</tr>
<tr>
<td>Chapter 2: A Framework for Operations Analysis</td>
<td>15</td>
</tr>
<tr>
<td>2.1 Problem Definition</td>
<td>15</td>
</tr>
<tr>
<td>2.2 Literature Review on Performance Measures</td>
<td>16</td>
</tr>
<tr>
<td>2.3 Problem Identification Process</td>
<td>18</td>
</tr>
<tr>
<td>2.4 Line Characteristics</td>
<td>21</td>
</tr>
<tr>
<td>2.5 Operating Plan</td>
<td>36</td>
</tr>
<tr>
<td>2.6 Service Management</td>
<td>43</td>
</tr>
<tr>
<td>2.7 Discussion</td>
<td>46</td>
</tr>
<tr>
<td>Chapter 3: Data and Methods</td>
<td>48</td>
</tr>
<tr>
<td>3.1 Data</td>
<td>48</td>
</tr>
<tr>
<td>3.2 Generated Information and Associated Methods</td>
<td>54</td>
</tr>
<tr>
<td>3.3 Representations</td>
<td>61</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>64</td>
</tr>
<tr>
<td>Chapter 4: MBTA Red Line Application</td>
<td>65</td>
</tr>
<tr>
<td>4.1 MBTA Red Line Description</td>
<td>66</td>
</tr>
<tr>
<td>4.2 Data Availability</td>
<td>71</td>
</tr>
<tr>
<td>4.3 Problem Identification and Resolution</td>
<td>72</td>
</tr>
<tr>
<td>Chapter 5: Summary and Conclusions</td>
<td>106</td>
</tr>
<tr>
<td>5.1 Summary and Conclusions</td>
<td>106</td>
</tr>
<tr>
<td>5.2 Future Research</td>
<td>109</td>
</tr>
<tr>
<td>Appendix A: Description of OCS Database and Other Data</td>
<td>111</td>
</tr>
<tr>
<td>Appendix B: Supplementary Figures</td>
<td>117</td>
</tr>
<tr>
<td>References</td>
<td>126</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

This thesis proposes a methodology for investigating the performance of rail lines, focusing on service quality and cost efficiency. The proposed methodology is demonstrated through application to a specific rail line. The application demonstrates that the proposed approach can be a useful means to identify areas of poor performance and help define strategies for improvement.

1.1 Background and Motivation

A significant amount of funding from research institutions, transit agencies and government grants has been devoted to research into improving transit performance. This is especially true in light of the surge of interest in new technologies. High expectations exist for the role of new technology in improving service quality and cost efficiency in transit systems. Some of these technologies may in fact have an important role to play in operations analysis, which is an area of work in which transit agencies have always invested some level of time and effort.

Much of this funding goes toward independent research by corporate consultants or researchers at academic institutions, particularly when resources are not available within the agency. Only rarely does the transit agency have a staff devoted to research and development. Operating departments at transit agencies are preoccupied with the daily demands of providing service. Independent research can allow more a thorough and detailed analysis to be conducted.

In addition to providing a less constrained research environment, independent research offers additional benefits. It can be instrumental in resolving differences among parties within the transit agency that may have different views on how to improve performance. It also tempers the problems
raised by the chief traditional mechanism for problem identification on rapid transit lines: comments passed up the chain of command. Thus operations analysis has been mainly reactive, dealing with the filtration of comments and the resolution of identified problems. Very few transit employees have the time, knowledge or data to investigate areas beyond their own area of responsibility. Everyone may have an opinion on how to improve performance, but few are in a position to undertake the broader types of analysis for making effective strategic decisions.

Independent research has these particular advantages because transit agencies are often preoccupied with operational necessities. At the same time, independent research is also very difficult to make credible within the transit agency.

Despite the typically short life of an independent research project, there is a strong incentive to produce a neatly packaged result, including, typically, a model and some results from the application of the model. Unfortunately there is often not enough time to fully research and understand the real agency objectives and constraints. This often results in a greatly simplified model that the agency staff is likely to dismiss as impractical.

Although it may be a legitimate technique consciously to simplify a problem in order to model it, to create a model in ignorance would be a mistake. To produce useful and implementable results, the researcher must be intimately familiar with the system being modeled. The pressure to produce quick results is often too great, which may cause the researcher to overlook problems in their modeling assumptions.

With electronic data collection systems becoming increasingly available, researchers currently have more observational data available for analysis than ever before. System data is becoming increasingly standardized and more is
available in electronic form than before. Increased data availability presents many opportunities for conducting useful analysis.

1.2 Prior Research

Bauer (1981) proposes a framework for the short-term operations planning process. Her research methodology consisted of a review of current practice in short-term operations planning, and then a suggested planning process. Her suggested process consisted of two steps: problem identification and problem resolution. She argued that the effectiveness of problem identification and resolution depends on the availability and dissemination of data. The operations planning activities her framework includes were focused on bus systems, using manually collected data. The planning decisions included in her suggested process were in the areas of schedule and route design. Aside from Bauer (1981), most of the prior research to date has been on analysis techniques and model development.

Krstanoski (1996), in his dissertation on rail line performance analysis, suggested a methodology for identifying problems with rail line performance. His method involved creating a simulation model of a hypothetical rail line based on some component models. He modified the inputs to the simulation and observed the change in the line’s performance. The simulation modeled dwell time variables along a line and did not include the terminal process and routine service management at the terminal or elsewhere. He believed that line performance could be captured through a dwell time model and a running time model. In the dwell time model he used, dwell times are determined solely through passenger flows at the door level and a random seed. The simulation did not account for cycles of train movements throughout the day. The dwell time model was calibrated using very detailed data on passenger flows captured by video at a single station in Toronto (see Mori (1988)). All other data was generated for the simulation using reasonable values. Krstanoski’s aim was to “capture the overall RT system performance variable properties in time and
distance" (p. 188). Krstanoski believed that the simulation was insightful enough to "reveal new RT line operational properties that have been overlooked or not precisely described by the existing literature in the field" (p. 1).

Krstanoski's research illustrates the tendency to begin by making strong assumptions about the properties of rail lines, rather than framing the research as an investigation of these properties. He labels his research an investigation into the properties of rail lines—when it is actually an investigation into the interactions between the assumed properties. In addition to problems in the methodological approach, his assumed properties of rail lines are not reflective of what is currently known about rail line operational characteristics. In this case, the omissions were not made because more realistic models were intractable, but because the models did not yet exist. These omissions are symptomatic of the methodology.

For example, on some lines, disturbances will last for many downstream trips. Most rapid transit lines do not have terminals that can absorb any amount of delay. On many lines, dwell times are strongly influenced by other systematic factors besides passenger flows and vehicle characteristics. One dwell time model may not be suitable for all stations on a line. Operations control techniques, which are dependent on the line characteristics and the operating plan, will determine how long disturbances will last throughout the day. For some lines, the terminal turnaround process (vehicle and crew management) may contain the most critical decisions affecting line performance. Since these techniques may strongly affect performance, they need to be included in the analysis of line performance. Comprehensive line-wide models for rail service management techniques have not been found in the available research literature.

The system behavior characteristics in Krstanoski's simulation are limited to lines that have operational characteristics consistent with his assumptions. For lines which have properties omitted from the simulation, the
questions posed must not require observing more than the effects in a localized area. An example of this would be the determination of maximum throughput passing a point or series of points along the line, with assumptions on the capacity of the track segments on either side of that segment. On some lines, dwell time variability may well be responsible for a great deal of the operating difficulty. This is more likely to be the case in lines with great numbers of stations and where dwell times are very sensitive to passenger flows, as is the case with on-board fare collection.

Krstanoski’s research illuminates an observation on the development of methodology itself. To be able to create and verify a line-level simulation model capturing system performance variables within component models, the researcher is required a priori to have observed the properties of the system. In order to do this effectively, there must be some amount of data analysis and abstraction. This initial data analysis for constructing a realistic simulation model could itself identify many types of performance problems on the line being studied, and reveal operational properties not discussed in the literature. Krstanoski’s only observational data was dwell time and passenger flows at a single station. Had Krstanoski used a set of data for creating other component models besides dwell time, he might have realized that the analysis of train movement data could provide transit agencies with valuable problem identification tools. This realization in turn might have changed his methodological approach from one in which the outcome would have been observations on the interactions between assumed properties to one in which the outcome would have been techniques for problem identification and resolution. A recent survey found that automatically collected real-time train movement data is available at many transit agencies worldwide (NYCTA, 2001). Even in the absence of real-time data, properties of rail lines such as service management can be gainfully studied using manually collected data.
1.3 Research Approach

To create results more useful than in the prior efforts, this thesis first suggests a new methodology for analyzing rail line performance. The methodology is based upon several premises. The first is that research into the properties of rail transit lines should begin with a grounding in observational data. This can be accomplished by examining the rail line’s characteristics using data sources, in conjunction with ongoing discussions with decision-makers within the transit agency. It is very important to engage in discussion with agency staff. There is a wealth of knowledge existing within the transit agency at many levels of decision-making, ranging from signal design to schedule design to service management. Each decision-making area involves a complex set of objectives and constraints. It is the researcher’s responsibility to understand these objectives and constraints as they exist rather than to make assumptions about them. This can point out discrepancies in the way people at the “bottom” and people at the “top” think they’re dealing with a problem.

The second premise of the methodology is that the research question should be placed in context with the operations planning process. Each existing line has its own characteristics that will have a significant influence on its performance. Track layout at terminal stations, passenger demand, and fleet size are examples of some critical line characteristics. Some characteristics such as signal design are relatively easy to change, given time and resources, while others such as passenger arrival rates are not easily controllable. These line characteristics establish the constraints for the design of the line’s nominal schedule, known as the operating plan. Key schedule design components are service frequencies, train lengths, running times, vehicle recovery times, and crew layover times. In turn, the line’s characteristics and operating plan set up the context for service controllers, whose responsibility it is to make service management decisions in real-time. Service management is also known as operations control or service control.
Basic service management techniques used in many different contexts include, for example: holding, short-turning, and re-sequencing of trains, as well as skipping stops, reassigning crews, and rerouting trains. These three areas of line characteristics, operating plan and operations control will be referred to repeatedly throughout this thesis.

There are several different ways in which this framework can be useful. First, in an ideal world the operating plan is based firmly on the line characteristics with the service management function applied effectively to deal with the inevitable occasional failures in being able to execute the operating plan exactly. Second, this hierarchy is necessary to compensate in service management for what's problematic in the operating plan or in line characteristics. On lines with insufficient running time, service controllers may give up trying to get a schedule change and instead manage with what they have—for example, using deadheading more frequently as in the recent past on the MBTA Blue Line, and crew swapping as in the MBTA Red Line at present—to get trains back on schedule.

The first set of questions posed by this methodology make up the problem identification phase of operations planning and analysis activities. They ask whether operations control decisions are consistent with the agency's complex set of objectives and constraints, based on the operating plan and line characteristics. Similarly, they ask whether the operating plan is consistent with the agency's complex set of objectives and constraints, based on the line characteristics. A second set of questions may ask whether the line characteristics should be changed, and if so how and what the expected benefit would be. It is finally valuable to ask whether the agency's complex set of objectives and constraints in any of these decision areas should be redefined.

The coping mechanisms for failures in the framework are also likely to be valuable themselves. A researcher investigating problems through application of the framework needs to find out how the system has adapted to its existing
problems, not just where the problems exist. It is important to find out what are the transit agency’s areas of flexibility and responsiveness in adapting to problems. Finding out the agency’s strong points is useful because this is what the researcher has to work with in suggesting changes or improvements.

In the MBTA Red Line case, the greatly reduced manpower availability has resulted in having a highly developed and skilled base of supervisors for managing crew personnel in real-time with little resources. In the context of other lines, no such highly developed skill base is fostered. This is an important observation, as changes within the framework that may be viewed as improvements may actually be counterproductive. Fellows (1990) pointed this out in the context of the MBTA Green Line, where the MBTA faced a decision whether to centralize or decentralize Green Line service control—a decision that could have resulted in the agency losing the skills of experienced service managers who would be replaced by new central control dispatchers.

In the Chicago Transit Authority, supervisors responsible for respacing intervals at terminals don’t have real-time information on train location, having only a radio link through which they can ask the control center where trains are. A line manager argued strongly with the author that this is preferable to the supervisor having a workstation with real-time information, because it provides incentive for the supervisor to listen extremely carefully to any radio communication—a practice that is beneficial due to all of the other important information being transmitted via radio that could otherwise be missed.

In New York City Transit, tower operators responsible for service management distributed throughout the system have a limited view of the line, while central controllers responsible for incident management have no view of the line at all. This organizational structure, which has been criticized, allows the tower operators to concentrate on what they do—service management—even when incidents are occurring. In other systems, service management can
suffer because of the level of information and the level of responsibility in a centralized structure without dedicated tower operators.

Not all coping mechanisms are for the best, however. In the Chicago Transit Authority Red Line, long dwell times at the Howard terminal result in queuing delays. The schedules under design take into account the operational issue causing the delays—that it takes a long time to clear an 8-car train of passengers before the train may be taken out of service or moved through the loop before making its next trip. In this case, the better action might be to increase the level of personnel available to check the trains, thus reducing the dwell times.

The next phase of operations planning and analysis activities is the analysis of alternatives. These research questions may inquire into the performance impacts of changes in three areas: line characteristics, operating plan and operations control. These two phases of problem identification and alternatives analysis are activities that transit agencies need to carry out, either on a routine or periodic basis. Some types of analysis, given their level of difficulty to conduct, may not be necessary. To what level of depth these analysis are developed will depend on the management. The question of the appropriate depth of analysis for each type of operational problem is beyond the scope of this thesis.

Some of the questions mentioned above have never been fully addressed in the literature. Therefore, a significant portion of this thesis is composed of new analysis techniques that could be applied to rail lines. The methodology described here was applied to an existing rail transit line, the Massachusetts Bay Transportation Authority (MBTA) Red Line. Examples of both problem identification and alternatives analysis are included in the application. Problem identification is focused on several issues within the three areas of line characteristics, operating plan and operations control. Within line characteristics, potential problems are examined in the design of the Red
Line's ATC track circuit layout and speed command logic. Within the operating plan, some potential problems were examined in the setting of scheduled end-to-end running times, train recovery time, crew layover time and train length. Within the operations control area, potential problems in the re-spacing of intervals at branch terminals were examined. Some issues examined lay in more than one area, such as delays due to terminal congestion and the intertiming of branch services.

1.4 Thesis Content and Organization

Chapter 2 presents the framework for analyzing rail system performance. This begins with a discussion of what constitutes an operational problem. Then it describes a general process for the identification of these problems. A description of the elements of operations within the framework of line characteristics, operating plan and service management is then presented.

Chapter 3 discusses data and methods that can be used in the process of identifying problems. The availability of data, especially automatic data collection systems on train movement and other system behavior are important means of building an operations knowledge base and of identifying problems in operations.

Chapter 4 represents the results of a study of various operational aspects of the Massachusetts Bay Transportation Authority's Red Line. One specific result achieved without any cost was a reduction of passenger waiting time stemming from fine-tuning of the operating plan. Further redesign of the operating plan could reduce weekend passenger waiting time significantly. Recommended reconfiguration of a small area of the Red Line's ATC system could alleviate a major bottleneck on the line with minimal investment.

Finally, Chapter 5 summarizes the methodology and makes recommendations for future research on the topic of rail transit operations.
Chapter 2
A Framework for Operations Analysis

This chapter presents the elements of the three-tiered framework introduced in Chapter 1. The framework serves to place each type of decision affecting operations in context, thereby giving the analyst perspective on how relevant and useful their analysis may be in improving rail line performance. First, this chapter briefly discusses what may constitute a ‘problem’ in operational performance. The universal questions asked in any area are whether problems exist, and what would be the effectiveness of making various changes. Then, this chapter describes the three areas of line characteristics, operating plan and service management.

2.1 Problem Definition

For the purposes of this discussion, there are three types of problems. One type of problem exists when the needs of the population are not being met—either the service quality needs of the passengers or the cost-efficiency needs of the transit agency. This type of problem tends to be highlighted by the use of performance measures within the transit agency. Trend analysis and peer comparison may be two ways to make use of these measures.

A second type of problem exists when the transit agency’s own objectives are not being carried out in practice. For example, a transit agency may wish to set scheduled crew arrival times at the end of a shift to be equivalent to their typical value, in order to minimize overtime pay and avoid crew dissatisfaction. Analysis of data would show whether this objective is being met or not.

A third type of problem exists when the decision-maker’s objectives and constraints are not properly set. In some cases, the objectives and constraints may be problematic because they may lead to poor service quality
or poor cost-efficiency in the long term. In another possibility, the objectives and constraints for a particular decision (for example, amount of scheduled crew layover time) from one part of the transit agency may be different from another affected part of the agency. When these differences result in the prevailing of one party’s influence, it may result in intra-agency conflicts.

2.2 Literature Review on Performance Measures

The first type of problem mentioned above is whether the population needs are being met—either the service quality needs of passengers or the cost-efficiency needs of the transit agency. The literature contains some discussions of the passenger’s perspective in measuring service quality. du Plessis (1984) places operational service quality in perspective with other aspects of service quality such as the condition of facilities. Sparberg (2001) discusses the setting of standards for rail service delays given the delay causes. Bowman and Turnquist (1981), Rudnicki (1997) and Muller (1999) discuss the differing expectations of passengers timing their arrivals to schedules versus arriving randomly as stations, and propose measures for regularity and punctuality. Henderson, Adkins and Kwong (1991) propose measures of travel time variability.

Periodic performance reports can be cited for many systems, including those for the New York City Transit Authority (2000a), the MBTA by the MBTA Advisory Board (1991, 1995) (including Tsihilis, 1982) and by MacDorman and Associates (1995). The headway regularity index used at NYCTA is discussed by Henderson, Kwong and Adkins (1991) A recent change in its definition was very controversial (see New York Times article by Kennedy (2000)). Comparing actual service quality to an ideal standard was discussed by Wilson et al. (1992) in the context of a now defunct MBTA system for monitoring violations of the standard in real-time. BART (Bay Area Rapid Transit) produces a daily performance report with measures of on-time performance, throughput and
service delays and generates a rating of the overall performance (BART, 2000). BART uses passenger-based measures as described by Buneman (1984). The TTC (Toronto Transit Commission) produces a daily report showing measures of excess wait time, throughput and regularity (TTC, 1991). The Docklands Light Railway’s operation is contracted with a requirement on service quality performance, as defined by a measure of system service availability. This measure is compared with other measures in terms fairness to passengers and the transit agency, by Gerrish and Hodgson (1998). Elms (1998) discusses definitions of system service availability as a performance measure. Tober (1977) discusses the establishment of service quality and efficiency standards at the MBTA as a way of identifying substandard performance. The level-of-service concept, as borrowed from research into the performance of highways, has been discussed by several authors including Murugesan and Moorthy (1998), Zhao et al. (1997) and the Transit Capacity Manual (TCRP, 1999). Authors Froloff, Rizzi and Saporito (1989) of RATP propose indicators of regularity, punctuality, personnel management and bypassed passengers. They argue that these measures do not satisfy the requirement that performance indicators be consistent with the objectives of service management. The authors go to great length to describe the conflicting objectives of service management that exist in any transit line, and even construct descriptions of which objectives apply during certain times of day on actual RATP bus routes.

Fielding (1987) proposes measures of cost efficiency and service utilization, but does not discuss service quality. Tomazinis (1975) discusses service quality and safety measures. The Canadian Transit Handbook (1993) proposes a list of common performance indicators used for planning, operations and administration, but does not frame these indicators as measures. The Railway Technology Strategy Centre (2000) surveyed CoMET (Community of Metros) agencies asking for values of various performance indicators related to service quality, cost-efficiency and other areas such as safety. New York City Transit (2000b) surveyed eleven transit agencies in the US and three in Europe for definitions of service reliability measures used by management. Passenger
survey-based research on passenger expectations has been reviewed in Gray and Hoel (1992, p. 622).

Some researchers created models that they argued may be used to relate service quality performance to many variables in line characteristics, operating plan and operations control. One such model was proposed, but not implemented, by Heimann (1979). Henderson and Darapaneni (1994) created a model that achieved this result, but only for variables we would describe here as line characteristics. Kimpel et al. (2000) achieved this result but for bus systems.

It can be seen that very little has been done to insure that performance measures are consistent with agency objectives and constraints. Analysis into performance would be incomplete without analysis of how service quality is related to various underlying factors that affect it.

2.3 Problem Identification Process

Rail transit lines operate through many undocumented practices, the underlying folklore of which can be viewed as ‘models,’ in the abstract sense that the decision processes are based on simple relationships between a finite number of inputs. These models, which rely on experience and rules of thumb, are therefore the result of simple trial and error or the subconscious weighing of conflicting objectives. The methodology described here is sensible for documenting these practices, to determine whether they are being carried out in accordance with their intent, and to assess whether the intent is itself effective from a multi-objective systemwide perspective.

This methodology does not formulate a checklist of issues to study. Although most rail lines operate on similar principles, it may not be productive to recommend a small but limited set of analyses to conduct. Local institutional and operating conditions need to be understood on a case-by-case
basis, and each application of the methodology will result in a different set of issues being highlighted. This is also problematic, because the scale in which each issue is presented will influence the response by the transit agency. A large-scale problem-identification type of analysis presented in sufficient detail will almost certainly warrant a response of some sort by the agency.

There are two ways of analyzing the agency’s success in meeting its objectives in a given decision area. The first method should be used when the researcher does not know the exact objectives and constraints as they exist, and does not attempt to assume what they are before beginning the process. First, the researcher creates illustrations of data pertaining to the assumed objective (or one component of the objective, if there are many components to it). Second, the researcher looks for features that appear unusual. Third, the researcher goes to the responsible party within the transit agency, or multiple parties affected or involved, and asks whether the observed behavior is consistent with their complex set of objectives and constraints. Their answer will be either that the observed behavior is intentional or is unintentional. If it is intentional, the underlying reasons as explained by these parties should indicate any additional objectives and constraints. If the result is unintentional, a problem has been identified. Trend analysis and peer comparison, when presented to the agency staff, may also productively elicit the pertinent conditions. If the researcher uses this method, he or she may not need to identify the full set of objectives and constraints before beginning the analysis.

The second method is applicable when the researcher assumes that he or she knows all the existing objectives and constraints. Then the researcher may create a model that attempts to take these objectives and constraints into account. The model may use simplifying assumptions to make it tractable. The model can be run and the results compared to the existing case. A large difference in service quality or cost efficiency performance will indicate a
problem. Often, the research may be framed as an investigation into whether existing resources can be used more efficiently or the investment of new resources justified.

There are three important components to the analysis process. The first component is understanding the data that is available or observable. The second component is the ability to generate new data through estimation, manipulation and reduction. The third and final component is the ability to illustrate the data. These steps have uses in identifying problems and developing and assessing alternatives for improving operational performance.

This methodology requires the researcher to decide the level of the analysis. Methods to identify problems in meeting objectives with respect to specific decisions provide varying levels of resolution of underlying cause. The goal of identifying underlying causes is to generate alternative solutions to the problem. Thus, it may be beneficial in some cases to study the root causes at a greater level of resolution, rather than to stop after identifying only its symptoms. This will depend on what level of information management feels is needed before action can be taken (for a more detailed discussion of this subject, see Bauer (1981)). Researchers can always probe more deeply into the cause of a problem, but the depth of the inquiry may need to be limited. True, identifying the root cause with greater precision will help prevent the problem from recurring. For example, a decrease in on-time performance may be due to a long-term increase in dwell times, and the increase in dwell times could have occurred as a result of a decrease in crew size. The oversight in setting running times may result from a lack of time and staff to set the schedule properly, which itself may be due to a departmental budget constraint. One cannot generalize about the appropriate boundary for probing into root causes, as the outcomes of such inquiries are not easily predictable.
The next section describes the proposed organization of analysis around the three areas of line characteristics, operating plan and operations control. These three areas contain many types of decisions, and it would be impractical to describe all of them. This section outlines a general description of some common types of decisions and the structure of their associated objectives and constraints. Literature sources are cited as applicable due to their focus on problem identification, alternatives analysis, and descriptions of objectives and constraints facing decision-makers as they exist in practice.

One of the hypotheses of this thesis is that the local context of any decision-making process in these areas will vary greatly, and thus it is difficult to generalize the existing constraints on one type of decision across all systems. Such generalizations have been made in prior research—where the authors relied upon prior literature as sources and had little additional information from the application context. Nevertheless, this chapter does describe in very general terms what the objectives and constraints may be for many of the decisions in the areas of line characteristics, operating plan and operations control.

2.4 Line Characteristics

Each existing line (and system) has a set of characteristics, some of which can be changed given some investment of time and resources, some of which may be hard to influence. In this section line characteristics are divided into fixed infrastructure, control system settings, energy management, route management, vehicles, work rules, labor contract terms, maintenance, incidents, passenger flows, and supervision. Line characteristics strongly affect service quality performance, as several researchers have argued (Henderson and Darapaneni 1994). Henderson and Darapaneni (1994) found that there was a strong relationship between the service quality performance of New York City Transit’s rail lines and a number of line characteristics. They
found that most of the variation in performance that was observed through routine monitoring is associated with the number of merges, the mean miles between failures, the presence of night work, the scheduled headway (defined here as an operating plan characteristic), passenger crowding levels, and whether school was in session. This supports the claim that the characteristics of each line differ widely and can indeed affect performance significantly. However, since it did not capture some variables that significantly affect performance, it cannot be used credibly as a way to control for the operating difficulties of each line. The authors concluded that reducing the number of merges would improve on-time performance. This conclusion, while it makes sense, would need to be weighed against the other benefits which merges provide: a well-connected system with many types of service provided and easier transfers, without any additional construction costs to build independent lines. Their conclusions may be misleading in that there are ways to improve the performance of poorly performing lines in their study other than eliminating merges. Since the variables in the study did not include some characteristics of terminal operations and recovery times, the effect of changing those factors did not emerge. A critical question is when investment to change the line characteristics is likely to result in sufficient benefit to justify the cost.

2.4.1 Fixed Infrastructure

Fixed infrastructure includes terminal track layout, running track layout, crossovers, route design, station locations, platforms, signal system type and design, and operations control system.

Terminals, running track and crossovers

Changing terminal track layout significantly would be uncommon, but more common would be modifications to the signal system or addition of crossovers that would allow existing track to be used in different ways. The functional
use of terminal tracks clearly affects performance, and this functional use may be changed periodically. There is no literature on how to evaluate potential changes in the functional use of terminal track layout, although such a question could be addressed. A starting point for this research would be to build from the expertise of route planners currently making these decisions.

Loop terminals and relay terminals can be reconfigured to operate as stub terminals as a cost-cutting action, and indeed this type of action is not uncommon in transit agencies. For example, 95th Street terminal on Chicago’s Red Line used stub-end operation rather than relay operation for a short period to save costs. Recently, the MBTA investigated adding a crossover on the Blue Line between State Street and Government Center to end operation there as a stub terminal.

A decision to add a terminal track may be relatively uncommon, but may be investigated as an alternative. The addition of terminal tracks increases capacity, which should reduce delays due to congestion. A third track may be used for trains pulling back into the yard, which may require additional dwell time at the platform to sweep the train of stray passengers. It may also provide the opportunity to schedule more train recovery time at the terminal, thus allowing more effective re-spacing of intervals between trains.

No literature on the performance effects of three- and four-track operation were available. Research on the placement of crossovers has been conducted by Abromovici (1986). He presented a method for determining the optimal location of emergency crossovers, based on travel times between the candidate locations and the nearest stations on either side.
Route design

Route design encompasses the decisions on what services to provide on the available track network. This affects the quality provided by the nominal service, as well as the operational difficulty which influences service management. Route types may be single lines, lines with a trunk and branches, merging-and-diverging service, lines with dual central area routes, corridors with express and local service (or alternating skip-stop service), or lines with turnbacks. Models for optimal route design have been proposed by Furth (1981) for bus systems as well as Ceder and Israeli (1997) for general transit systems. These involved formalization of waiting times, travel times and operating costs incurred by crew and fleet size. No literature on route design techniques used in practice was located, although agencies with complex track networks such as NYCT have planning departments specifically for the investigation of route design changes.

Stations and platforms

Station location is difficult to change except for surface lines, where station relocation, consolidation or infills can be built. These actions can be taken to improve capacity, change travel times, and attract ridership. Additionally, if a track has platforms on either side, the configuration of which side is used may be changed. Some stations have double-berthing of trains on the same track. Lipfert (2001) describes a simulation-based capacity analysis of the MBTA Green Line under single and double berthing in subway stations. Models of stop location have been proposed for bus routes such as Furth and Rahbee (2001), and these may be applicable to surface rail.

Platform attendants may be used to expedite boarding. Their general function is to help the on-board crews close the doors at the appropriate time, and to prevent holding of doors. Platform attendants are especially useful on very congested lines at the stations associated with bottlenecking,
in order to improve efficiency and reduce annoyance to passengers. In some systems they are used only on special event days where large surges of passengers are expected. Early research by Hankin and Wright (1965) studied the relationships between passageway dimensions and passenger flow rates to influence the design of London Underground stations.

New York City Transit (NYCTA, 1997) conducted an experiment using platform attendants on the southbound express platform at Grand Central Station on the Lexington Avenue corridor. They began with one attendant per door (30 attendants total) and achieved a 10 percent reduction in average dwell time between 8 and 9 AM. The platform coverage was reduced to one attendant per two doors, which achieved a 6.5 percent reduction over no attendants. Under a further reduction to one attendant per three doors, a 5.8 percent reduction in average dwell time was observed. The report states that a rush hour delay appeared to cause between 5 and 7.5 percent fewer late trains, and that throughput had increased by 2.1 trains in the peak hour. Unfortunately, nothing was said in the NYCT report about the change in distribution of dwell times, rather than simply the average. The CTA (Chicago Transit Authority) conducted a similar experiment in reducing dwell times of trains going out of service at Howard station through the use of additional platform attendants.

The locations of staircases affect the distribution of passengers on the platform. Several researchers have investigated the distribution of passengers along the train’s length at the time of boarding. Mori (1988) and Wirasinghe and Szplett (1984) found that passenger distribution along the train’s length is strongly influenced by the position of platform entrances and exits. Mori’s data was re-analyzed by Krstanoski (1996), who remarked that both Mori (1988) and Wirasinghe and Szplett (1984) found that for platforms they studied with single entrance locations, there was a negative exponential spatial distribution with high passenger concentration at the train doors nearest the platform entrance.
Signal system type

Changing from one signal system type to another may have both short-term and long-term effects, involving service quality and cost-efficiency. Newer signal systems are aimed at reducing the maintenance costs of wayside equipment, as well as providing increased operating flexibility and capacity. These improvements however, cannot be generalized and depend on design on a case-by-case basis. The installation of a new signal system on existing lines, termed a cutover, has in some cases had negative effects on level of service (OTA, 1976 p. 93). A more significant body of literature exists discussing the assessment of signal system types and their effect on capacity, cost and safety. Genain (2001) discusses developments in communications based systems in operation internationally. Gill (1998) remarks that moving block ATC systems represent an improvement in line capacity, exhibit improved recovery from disturbances, improve travel times due to coasting and more efficient speed regimes, and provide room to increase capacity over the nominally scheduled capacity. Gill remarks that an improvement in delay recovery may result in congestion at the downstream terminal (p. 238). Gill and Goodman (1992) comment that moving block systems should enable the effects of disruptions to be minimized (p. 273). Hubbs and Mortlock (2000) describe a communications-based train control standard being implemented on the NYCTA L line. Miller (2000) describes Advanced Automatic Train Control (AATC) being implemented at BART.

The Ontario Ministry of Transportation and Communications (Abrahamson et al., 1977), Office for Technology Assessment (1976) and the Transit Capacity Manual (1999) compare the theoretical train throughput capacities and show that ATC systems have slightly higher capacity. The Transit Capacity Manual (1999) discussion is slightly misleading since the operating speeds at the minimum headway are not compared. Further, the capacities of fixed-block signal designs will depend on block layout and choice of speed commands available, which is a matter of design rather than technology. These and other
track layout factors will affect the actual capacities, as pointed out by Abrahamson et al (1977). The Office for Technology Assessment (1976) evaluates Automatic Train Control system in terms of potential cost, efficiency and safety benefits.

Signal design

In the area of signal design, engineers work to set various parameters of the designs based on a set of objectives and constraints. A poorly designed signal system may cause unnecessary congestion delays and ultimately may lead to schedules that work around the signal system’s limitations. Poor signal design may also cause an uncomfortable ride. One area of analysis would be whether the existing design meets current objectives. Setting the objectives is another area of analysis.

For fixed block systems, the track circuit layout and choice of speed commands available is a matter of engineering design. The typical inputs to the design process are a pre-specified minimum headway, track geometry including gradient and curvature, and safe braking distance model parameters. The safe braking distance parameters include the worst-case braking rate of the fleet used, and may also include the acceleration rate.

The objective of the signal engineer is to minimize the amount of wayside equipment associated with the signal system, while achieving the desired headway and not violating safe braking distances. Some researchers have studied the service quality implications of operating trains close to the minimum headway. Throughout Krstanoski’s (1996) analysis of a moving-block system, the existence of a minimum headway bottleneck due to the signal system has a big influence on service quality variables, and has important implications on the speed of operation at a minimum headway (p. 198). He argues that practical capacity is not well defined, because operations at the minimum headway can be severely degraded. Bergmann (1970) studied the signal
designs prevalent in 1970 in terms of minimum headway and travel times. However, he did not study the effect of station spacing—although close station spacing can significantly decrease performance. Gill and Goodman (1992) note that in many systems, restrictive signal codes or aspects slow peak period service due to increasing service levels. They note that this also increases energy costs due to braking. They note that resignaling is a viable option for increasing line capacity. Gordon and Lehrer (1998) note that a train following closely behind another will teeter totter between speed commands as the trains go over hills, or as they accelerate and brake between stations. They note that this is uncomfortable for passengers and also wastes energy, but through advanced train control technology it can be avoided without significantly impacting overall travel time. They suggested that this feature be integrated into advanced control systems. Janelle and Polis (1980) describe software for interactive signal design of fixed-block systems. Systems integration of a new signal system for the TTC (Toronto Transit Commission) has been discussed by Allen (1995). The modernization of SEPTA’s Broad Street subway signal system has been described by Childs (1995). Signaling in Japanese rail lines has been described by Takashige (1999).

A common design feature at interlockings is the prevention of trains from entering an interlocking until the preceding train has cleared a prespecified interlocking limit location on the other side. This prevents a train from becoming stuck over an interlocking. This design criteria makes sense in the case of flat junctions, where a train stuck over the interlocking will block crossing traffic. It also makes sense at stub terminals, where an exit route must remain for trains in the station. Unfortunately, this feature has also been implemented even on normal crossovers, where this type of protection for operational flexibility does not make sense given its impact on headways. Another abnormal case is when two interlockings overlap in such a way that a train may be held outside both interlockings until both are clear.
Researchers have developed models to try to optimize signal designs. These include Gill and Goodman (1992), Gill, Goodman, McCormick and Aylward (1987), and Chang and Du (1999). Gill and Goodman (1992) propose a nonlinear optimization model. The objective is to achieve a specified headway while minimizing the complexity of the signaling layout, without infringing on safety constraints. The model output is a signal block layout and available speed commands. The impact of achieving a specified minimum headway is not the over-riding concern of operations, however. The performance of the system depends on the intermediate speed commands, and their effect on running speeds not just at the minimum headway but also approaching it. The layout of non-headway critical blocks affects the speed of travel while approaching, but not yet at, the minimum headway.

Another performance measure that depends on signal design is frequent acceleration and deceleration. The objective of achieving a minimum headway will not also minimize acceleration and deceleration. An operating phenomenon that occurs is the teeter tottering between speed commands while closely following a moving train. The layout of blocks and choice of speed commands can affect this operational property. In many cases, transit authorities do not operate at nearly the minimum headway due to the poor level of service (including frequent changes of speed commands, interstation stopping and slow speeds) that would result.

Gill, Goodman, McCormick and Aylward (1987) discuss an optimization technique for setting the levels of intermediate speed commands available in an ATC system (for the London Underground’s Central Line). The objective was the minimization of travel times, thus resulting in a smaller fleet requirement. An alternative objective was the minimization of headway. Most research has implied that the objective of signal design is to minimize headway rather than optimize running times. The authors note that an optimal speed command setting for a dense urban network may not be well suited to suburban headway and speed characteristics, but also remark that the low
frequency of suburban services mean that the urban characteristics control the decision process. Chang and Du (1999) propose a model to optimize signal block layout based on the criterion of headway minimization, but this does not optimize running times. Lastly, Gill (1998) remarks that moving block control system design must be fine-tuned to reach acceptable ride comfort characteristics such as frequent adjustments in acceleration, implying that ride comfort may be worse in moving block than in fixed block systems.

2.4.2 Control system settings

Some rail transit lines have train regulation systems that operate as an overlay to the signal system. Train regulation systems may influence running speeds and dwell times by using various algorithms. Vuchic, Bruun and Krstanoski (1996) conducted a study evaluating whether programmed dwell times in the BART system at a single station could be improved. Similar research was undertaken by Barker (2001) for San Juan's Tren Urbano system under construction at this writing.

2.4.3 Energy management

Energy costs, although not a major portion of operating costs, are large enough that transit agencies are often concerned about whether their system could be more energy efficient. The question of whether energy costs may be reduced is an area for analysis. Transit agencies typically purchase power from other providers, under contracts that specify the cost structure. The transit agency can negotiate a new contract with a modified cost structure, or can modify its energy use to reduce costs subject to other constraints (Uher and Sathi 1983). These techniques have been the subject of some research.

A report by Uher and Sathi (1983) discusses the potential reduction of peak power demand in rail transit systems. The authors summarize power rate structures, characteristics of power demand and factors affecting it, and the
monitoring of power demand and billing. Several strategies of energy reduction were discussed in the report, including coasting, top speed reduction, running speed optimization, shorter train lengths, and regeneration of braking energy. A more recent discussion reviewing potential strategies for reduction in energy costs was by Levin (1999, in German).

Some researchers have studied the potential benefits of train control features that attempt to reduce energy consumption or avoid energy demand spikes. Gordon and Lehrer (1998) noted that voltage levels could be reduced, thus saving costs, if the control system could minimize the spikes in voltage demand by avoiding unnecessary stops and starts. Under ATC, civil speed restrictions may be higher than the highest allowable ATC command, in order to restrict large currents that may lead to increased failure rates and associated maintenance problems (Gordon and Lehrer, 1998 p. 174). The authors note that coasting is a potential enhanced ATC feature that would decrease travel time while saving energy. They suggest a power use profile during acceleration that limits power use to 5MW, while having only a marginal impact on trip time.

Sanso and Girard (1995, 1997) studied the potential for reducing power peaks through controlling the departure times of trains. The power peak reduction could be achieved by avoiding cases where two trains drawing power from the same substation accelerate from a station simultaneously. The objectives of the model that the authors designed was to minimize power peaks while respecting constraints on how significantly operation could be affected through the holding.

A similar scheme was suggested by Gordon and Lehrer (1998). They noted that if train acceleration could be coordinated with braking of another train within the same area, regenerated power could be transferred rather than using the substation as a source. They suggested that the first train to stop in a station be held there by the algorithm to wait for the arrival of the opposing
train; when the train is allowed to depart, it will use the energy available from the braking train rather than using power from a substation. Alternatively, a train approaching a station can be commanded to begin braking early to allow the departing train to leave on time. Gordon and Lehrer remark that these algorithms are probably not worth pursuing due to travel time impacts, but they suggest that a schedule that encourages these forms of energy transfer would be justifiable.

Murata, Nagata, Kawabata, Tashiro and Takaoka (1995) presented a model and algorithm for optimizing energy consumption by selecting an interstation speed profile for a train. The suggested optimal speed profile is one that recommends speeds slightly slower than the fastest running profile, while not affecting running times too greatly. Sujitjorn, Mellitt and Rambukwella (1987) describe a method of regulation with an objective of minimizing energy consumption. This objective is achieved through coast control and dwell time regulation. The authors state that energy consumption can be reduced by 30% with only a 5% increase in interstation running time (p. 303).

2.4.4 Route management

Route management systems provide either automatic or manual mechanisms by which the movement of trains through crossovers is controlled. Route management systems mainly affect service management and thus, service quality, but also may affect operating speed and capacity. The real-time decisions in the management of priority and route choice of train movements at junctions are a matter of service management. The setup of the infrastructure by which these decisions are carried out is a potential area of analysis.
2.4.5 Vehicles

Fleet size is a constraint on the maximum scheduled vehicle requirement, train length and number of spare trains available for run-as-directed trips. Transit agencies are continually decided on the rebuilding or retirement of older cars and purchase of new cars, all of which have a direct affect on fleet availability. Dersin and Durand (1995) created a model that associates service quality with several factors including the number of trains in reserve. The authors argue that this model may be used to set fleet size and fleet reliability standards necessary to achieve a desired level of service quality, given many contextual factors affecting service quality. Pierce (1995) surveyed rail transit properties to determine their spare vehicle ratios and discussed ways of sizing a fleet. Vandebona and Richardson (1985) describe differences in dwell times due to fare collection strategies (i.e. on-board payment, proof-of-payment, etc.). The NYCTA conducted dwell time tests using two different fleet types (NYCTA, 1991).

2.4.6 Work rules and labor contract terms

Work rules and labor contract terms have a side-effect on both service quality and cost efficiency performance. An example of work rules coming into play as a cause of service unreliability would be in a transit system where a minimum dwell time policy exists, but only some of the crews comply. Labor contract terms are a major factor in the generation of crew duties, which is described later in this chapter.

2.4.7 Maintenance

Decisions in the area of transit vehicle and right-of-way maintenance are largely undocumented in the literature. Nevertheless, they are activities that have a significant impact on service quality and cost efficiency performance of a transit line. Deferred maintenance, a tactic used by US transit agencies
in times of fiscal crisis, results in poor service quality while achieving its
goal of reducing short-term costs. Henderson and Darapaneni (1994) calibrated
a model associating on-time performance with vehicle failure rates, among
other variables. They argued based on the application of their model to
historical data that much of the improvements in on-time performance
achievable through upgrading the fleet had already been achieved.

2.4.8 Incidents

Incidents may be of many types, including passenger problems, police
action, track fires, persons on the track, signal system problems, vehicle
problems, or crew problems. These problems and the delays they cause may be
influenced by the agency. Vehicle failure rates, measured as the mean miles
between failures (abbreviated "MMBF") of vehicles are typically monitored at
transit agencies as a performance indicator. A common activity among watchdog
parties is to look for a worsening of the vehicle failure rate. Agencies may
sometimes propose target levels of improvement for mean miles between
failures. Some speculate that the results of these targets may be to shift the
service quality impacts in a way in which the MMBF measure is not structured
to detect. As mentioned previously, Henderson and Darapaneni (1994) estimated
a model relating vehicle failure rates to overall performance. They
interpreted the results as indicating the level of improvement in service
quality achievable by improving maintenance. Peer comparison of metro systems
could highlight whether it is different practices or different contexts that
account for different levels of incident types and the delays they cause.

2.4.9 Passenger flows

Passenger flows greatly influence service frequency and consist length,
both of which are discussed later. They also greatly affect service management
strategies. Passenger arrival rates are not directly under the control of the
transit agency. In circumstances of unusual demand, transit agencies
communicate to the general public to help spread demand spikes. Fare structures which vary fare by time of day may also influence passenger arrival rates. The RATP once had gates to control the flow of passengers onto the platform. The service frequencies and train lengths of other lines influence arrival rates at stations where bulk transfer are made.

The capacity and length of passageways between the lines influence instantaneous passenger arrival rates. Route choice behavior may be influenced by route design and schedule design and especially with complex route networks, passenger volumes may be influenced. Analysis may investigate whether passenger arrival rates are too variable or too highly spiked; they may also ask whether passengers' origin-destination route choice should be influenced to redistribute passenger volumes away from overcrowded service. Wiener and Lidor (1978) argued that the NYCTA should encourage more passengers to use the less crowded local trains than the overcrowded express trains that passengers were using because they felt they were achieving a higher travel time savings than they actually were. Analysis could determine whether the transit agency should try to allow for more even passenger loading across train length. According to Komaya and Asuka (1997), passenger guidance is a technique researched in Japan in which passengers waiting to board are encouraged to move along the platform to align with cars that are less crowded. They remark that such a system has been implemented on a very congested line with crowded trains. Research into identifying systematic variation in ridership data has been conducted by McCord and Cheng (1986).

2.4.10 Supervision

Analysis can attempt to determine whether the number of field supervisors is adequate, although such attempts reviewed here have not succeeded. Duerr and Wilson (1991) studied this issue by doing a peer comparison of four agencies. They compared the ratio of supervisors to operators in the control center versus the field. Their results show wide differences, and they argued
that using these ratios as performance measures to compare the four agencies cannot help form conclusions about appropriate levels of supervision, due to the differences in context between them. Aside from levels of supervision, there are open questions as to how to distribute command between central control, line control and field control. Equally important, there are open questions as to how to present real-time information to supervision in these locations to maximize effectiveness. Vuchic (1990) evaluated New York City Transit Authority’s control center. He used a survey to study attitudes towards improving relations between the control center and other personnel, and to study attitudes toward the use of new technology.

2.4.11 Miscellaneous

LaMarca and King (1980) studied winterization technology. They made recommendations on the deployment and effectiveness of various winterization hardware and on winter operation plans for transit agencies.

2.5 Operating Plan

The second area of research is the operating plan. The operating plan is the nominal schedule of the line, which involves decisions in setting a host of parameters and resolving complex logistical problems. Given the vehicle storage locations, schedule design (a) sets timetable parameters, (b) schedules vehicles to trips, and (c) assigns crews (see Ceder 1991). The timetable parameters include many values that change by time of day, including service frequencies, train lengths, running times, train recovery times, and crew layover times. The scheduling of vehicles to trips results in an initial vehicle schedule. Crew duties are developed to cover the vehicle schedule, given work rules and other constraints.

One type of analysis could focus on how the transit agency intends to set the timetable parameters and what the objectives are in the vehicle scheduling
process. A second type of analysis could determine how the transit agency should set these parameters and what the scheduling objectives should be. A third type would investigate how existing operating plans perform in terms of the agency’s own objectives. A fourth type of analysis would investigate whether the frequency needs of the passengers are being met, or whether the cost efficiency needs of the agency are being met. Many of these research areas have been addressed in the literature.

Transit agencies operating under budget crises may take a variety of actions to reduce vehicle and crew requirements while trying to maintain adequate service. Wheeler (1988) discusses the minimization of fleet size, while Shriver (1981) discusses a wider variety of techniques used under crisis conditions. The use of computers in the schedule design process has highlighted the necessity for research on setting the software parameters to achieve the desired result. Chang (1995) studied the sensitivity of the HASTUS scheduling software parameters and the acceptability of the resulting schedules for MBTA bus routes.

2.5.1 Vehicle scheduling

Train length

Train length decisions are linked to capacity and manpower requirements over the day, to fleet availability, and to constraints on when and where train length can be modified during the day. Antonisse (1985) studied the service quality and cost impacts of schedules that differed in terms of train lengths for the MBTA Red Line. Among the various alternatives tested were those that had uneven train lengths for trains on two branches. He discussed the impacts of different train lengths on capacity, crowding levels, dwell times and reliability. The assumption of even headways between the branches was made, although he suggested relaxing this assumption for future research.
Krstanoski (1996) used a simulation to investigate the theoretical effects of having alternating trains of different lengths.

Setting frequency

Literature on setting frequencies has mainly been focused on bus systems, including works by Furth (1980, 1981). At around the same time, Cury, Gomide and Mendes (1980) described a simple model for setting frequencies on a rail line. The objectives and constraints in setting frequencies are complex, and are briefly discussed here. A main objective of peak-period service is to carry the passenger loads without exceeding load factor standards, or at the very least without leaving passengers behind, given fleet size constraints. Service levels during other time periods may follow policy (minimum) headways. Because operator platform hours during the off-peak are less expensive than during the peak, in systems where passenger volumes are highly peaked across the day, proportionally higher levels of service may be provided during off-peak periods. In fully automated systems, high frequencies may be provided all day with varying train lengths. Fleet size or operator efficiency constraints may lead to having scheduled turnbacks, which are discussed below. A common constraint in setting frequencies is that there must be a minimum frequency in order to meet the needs of the passengers, and a minimum cost efficiency standard in order to meet the needs of the transit agency.

The existence of scheduled turnbacks affects the service frequencies that can be provided. Peak volumes may be imbalanced by direction. If there is sufficient storage capacity at both ends of the line, there may be an imbalance of service serving either direction within a peak period. In this case, the excess vehicles build up at one terminal at the end of the AM peak and are reserved there until the PM peak where they provide extra service in the opposite direction. This technique achieves fleet efficiency but may involve crews deadheading back to the other terminal. On lines with low-frequency service, and in which many passengers are likely to time their
arrivals, the transit agency may have as an objective to keep clock-face headways for easier memorization. In systems with many feeder route connections, it may be beneficial to match train headways with the feeder connection headways, to minimize transfer delay. Research areas could also investigate whether train lengths should be shortened while providing increased frequency, to avoid passenger queuing in stations passageways. The signal system and operational characteristics of the turnbacks present a limitation on the service frequency provided. Poor headway regularity on a route with branches will also reduce the ability to operate a high frequency service effectively.

Beyond setting frequency, actual departure times must be chosen. These are largely based on the frequency criterion, but actual departure times may deviate from the nominal frequencies. On lines with merges, the nominal schedule may reflect deviations from the nominal frequency in order to mesh with other service on the same track.

Hall (1987) studied the effects of scheduled headway on passenger delay at BART stations. He created models for estimating queuing delays in stations, and then used them to predict whether they would improve or worsen, depending on whether ridership patterns are sensitive to service increases. Vuchic, Bruun and Krstanoski (1996) studied load factor standards on BART, discussed whether the current standards may be problematic, and suggested other ways of meeting the objectives of providing adequate capacity. Antonisse (1985) studied service frequency and train length on the MBTA Red Line and evaluated various alternatives in terms of cost efficiency. In lines providing multiple services, whether branches or short-turns, the scheduled sequence of trains will affect performance. Researchers may investigate whether the scheduled sequence is optimal.
Running times

Running times must be periodically reset in response to changes in passenger flows affecting dwell times, extended periods of track reconstruction, or in response to operations practices such as those that may hold trains at merges and terminals. Running times may vary across time of day, by weekday, Saturday or Sunday, by season or in other systematic ways. Due to the constraints on the production of schedules, only one weekday schedule is typically produced at transit agencies. During the 1950s, the MBTA Red Line had two weekday schedules to account for extra shopping trips made on certain days of the week. If 24-hour service is provided, the production of schedules may require an extra weekday schedule that overlaps with the initial weekend service. Schedules are generally revised four times per year. Running times may be defined at the level of line segments, or simply one end-to-end segment. In lines with branching, the running times of each branch should be properly timed for the desired headways in the trunk section. Analysis areas include whether there are problems with these segment-level running times.

Another form of research area may investigate whether running times are highly variable. High variability in running times is especially problematic on lines with merging service, since trains may miss their slots and cause unbalanced loads, high waiting times and large service gaps.

There has been theoretical research into schedule-based holding for transit service, but the projected benefits have sometimes been based on inappropriate generalizations from bus to rail. Time points slow down vehicles in order to maintain improved headway regularity or schedule punctuality. Fong (1981) studied the selection of time-point locations and schedule design methods, but generalized characteristics of bus service to rail. Wirasinghe and Liu (1995) conducted similar research. Muller (1999) provided methods for determining scheduled running times based on the goal of having eighty-five percent of arrivals at the terminal within the scheduled running time. This
was achieved through mid-route ‘passing moments,’ in which the scheduled running time was extended. These researchers mainly focused on the en-route causes of poor service quality. This is applicable in cases where lines are long and have many stops, and in systems where dwell times are very sensitive to passenger flows. In other cases, other factors are more likely to have a more significant effect on operational service quality.

Train recovery times

Train recovery times must be long enough to absorb incoming delays while minimizing delays to outgoing service. Higher train recovery times result in higher fleet requirements. The terminal track layout and service frequencies are important in determining the maximum train recovery time by time of day. If train recovery times are set too high for the terminal track layout, delays will develop for incoming trains that have nowhere to berth. Long train recovery times in the peak periods would cause lower utilization of vehicles. If there are extra trains that can be inserted into service, as well as extra crews through scheduled cover or crew fallbacks, then the effects of low train recovery time may be tempered. Unless there are scheduled mid-route holding locations, the cycle time along the route is composed of the running time and the terminal recovery time. As the fleet requirements are integer values, the train recovery times are limited to a choice between certain values, unless there are mid-route time points which allow the end-to-end running time to be increased slightly at the scheduler’s discretion. Some researchers have tried to create models that would predict the reliability resulting from a transit schedule (see Carey (1998, 1999) and Carey and Kwicinski (1995)).
2.5.3 Crew scheduling

Crew layover

Rules of thumb exist for crew layover times as a percentage of end-to-end running time, with a typical figure of 20%. Union work rules may specify a minimum value of layover time. Crew layover time must be sufficient to allow the crew to have a rest break and change train ends if necessary. In practice, it may be distributed unevenly throughout the day, with the higher crew cost of a peak platform-hour persuading the transit agency to schedule less crew recovery time during the peak periods than in the off-peak periods. Larger values of crew layover time can provide increased ability to close headway gaps if trains are available. Research may investigate whether more crew slack time would result in improved service quality.

Analysis may also give insight into the station at which crew reliefs should be located. Typically, crew reliefs are located at terminals, although they may also be located mid-route. The relief location affects the generation of the crew duties and can result in differing levels of crew productivity and hence, cost efficiency. A mid-route relief point on lines where there are also terminal relief points is intended to bring worked time (termed 'platform time') closer to the number of paid hours, in cases where the cycle time does not divide evenly into the number of paid hours in the shift.

Crew size

Most rail transit lines operate with either one- or two-person crews. Reducing the crew size to one person obviously reduces the manpower requirements of train operation (Vuchic, 1983) but there are also indications that safety in train operation may improve from the reduction, as boredom is minimized by the increase in responsibilities of the train operator.
Given constraints on crew layover time, the crew schedule is developed for
the vehicle schedule of each pick, usually by season. Union work rules present
many constraints. The objective of the agency is usually to minimize the total
cost of the crew schedule, although if there is to be a reduction in manpower
requirements, the agency may have as an objective to graduate the reduction.
Union work rules vary, but usually specify work time restrictions as well as
the premium pay structure for crew runs having long spread times, swingoffs at
different locations, paid meal break, and overtime.

Cover crews

In addition to the nominally scheduled train operators, the creation of a
list of additional spare operators, also called a cover list, is a normal
activity of the operating departments of transit agencies. In addition to the
cover list, scheduled crew runs may have blocks of time where they are
scheduled to provide additional cover. The nominal schedule sheets may not
always indicate such cover duties, but the placement of the nominally unused
time is an intentional design decision of schedulers. Research into workforce
planning has proposed models for spareboard optimization of cover duties (cf.

2.6 Service Management

The third area within the framework is service management, also known as
operations control or service control. Service management is the work of
deciding how the nominal schedule should be modified throughout the day in
light of unexpected events. Service management techniques are typically passed
down by word-of-mouth, and the techniques vary across lines. These techniques
greatly affect service quality. The decision-making responsibility is
distributed in some manner depending on the transit agency. The decisions
include vehicular logistics, crew management and other decisions. There are many possible research questions in the area of service management.

Service management is based on a set of objectives, which typically include some or all of the following:

- Picking up passengers at stations without leaving any behind
- Regularity of headways
- Punctuality
- Insuring transfer connections
- Personnel management
- Energy management
- Meeting performance measures

One type of research investigates service management practices for a given context such as a line or group of lines. This would attempt to document the objectives and constraints as they exist within that context. A second question investigates to what degree service management decisions are being made according to the agency’s objectives and constraints. A third question investigates whether the agency’s complex set of objectives and constraints in service management decisions are properly thought out.

The most seminal work on service management located for this review is by RATP (1989) and focuses on bus service, although most of it is applicable to rail. It attempts to inventory and describe the objectives and the techniques of service management in general and then applies these concepts to a number of RATP bus routes. It proposes the conceptual design of a training simulator for service management personnel. Finally, it discusses RATP’s existing computer aids for service management through a description of their features and interviews with service management personnel. New York City Transit has published a guide to service management that lists general strategies (NYCT, 1992).
Some researchers have tried to identify problems in the way service management is carried out. Deckoff (1990), Macchi (1989), Soeldner (1993) and Coor (1997) studied various service management techniques. Deckoff evaluated short-turning decisions on the MBTA Green Line at Boylston station. Soeldner (1993) framed his research as an evaluation of the major elements of service control on the MBTA Green Line. Unfortunately, several important aspects of service management were left out of his evaluation. First, terminal dispatching was not considered an aspect of service control, which it is. Soeldner’s evaluations of short-turning decisions were based on the premise that the objective of short-turning is to close headway gaps. Soeldner mentions in passing that delays may propagate downstream due to limited recovery times, but goes on to assume that this does not influence a service controller’s decision to short-turn a train if it is late. He evaluated control decisions for several days of data as “good” or “bad.” Second, Soeldner does not discuss holding tactics related to the double-berthing of trains at Park Street station in the proper order. Also, at some Green Line stations on the surface as well as the subway, trains may stop at the start of the platform and drop off passengers without picking any passengers up. It is unclear how often this is carried out compared to expressing without dropping off passengers, but it would be worth discussing.

Other researchers have proposed real-time control models. Craven (1989) studied the potential service quality benefits of adding terminal dispatching supervision on a bus route that had no supervision at one end. Eberlein (1995) created general models describing holding, deadheading and expressing. She used some observational data from the MBTA’s Green Line, but only to observe headways leaving a terminal (see also Eberlein, Wilson and Bernstein 2001 and Eberlein, Wilson, Barnhart and Bernstein 1998). Ding and Chein (2001) created a model for holding for the New Jersey Transit’s Newark City Subway. All authors including those previously mentioned failed to describe the objectives and constraints of the control actions they were studying as they exist in the context of the line they were studying. Models pertaining to recovery after
disruptions have been proposed by Furth (1995), O’Dell (1997), O’Dell and Wilson (1997), Shen (2000) and Puong (2001). Models by Furth (1995) and Puong (2001) recognized that recovery strategies affect downstream operations. Their models included these relationships as constraints. In the future, researchers should investigate incorporating downstream effect more realistically into the modeled system’s behavior.

For major disruptions, rail service management involves the decisions related to replacement service—such as what the capacity of segments of the line would be under various blockage scenarios. Research questions may ask what these capacities are, and what the structure of the line’s operation should be. Song (1998) computed the capacities of segments of the MBTA Red Line under various blockage scenarios involving loop, shuttle and single-track operations.

2.7 Discussion

This chapter defined three general types of problems: when the population needs are not being met, when the agency’s objectives are not being carried out in practice, and when the agency’s objectives are formed somewhat in ignorance. For the second and third types of problems, analysis can be categorized into three areas, namely line characteristics, operating plan and service management. These three areas are somewhat hierarchical, since the operating plan is based on the line characteristics, and service management takes as its inputs the line characteristics and the operating plan.

The use of performance measures and targets may lead to several problems. It is unclear that setting targets for performance has been an effective way to achieve improvement in performance. This appears mainly due to performance measure definitions that are not consistent with optimal service management practice. Thus, the targets are prone to cause unintended shifts of the problem into unmeasured areas. If an unintended shifting into an unmeasured
area does occur, it follows that unmeasured aspects of performance tend to be those that are more difficult to measure. One of those aspects may be morale. Another may be a shifting of responsibilities within a reduced workforce. For example, under a reduced schedule a supervisor may step in and do the work of a train operator if there are no cover crews available, but this activity would not be part of their job description. Thus, the problem may seem to disappear, while in fact this shift may not necessarily represent an improvement.

In the case of service management, it may be possible to learn about the objectives and constraints that exist by abstracting from observations of actual decisions. To the extent that patterns exist in the behavior of those responsible for service management (and that those patterns reflect the assumed structure of the objectives and constraints present), one may gainfully develop models of service management that are calibrated to reproduce the patterns observed. This approach is being used by Brezillon, Pomerol and Saker (1998) to create an incident management system that they hold will be consistent with the agency’s current practice.

Areas of decision-making within the three-tiered framework, as well as research questions into those areas have been posed here. The general approach proposed in this thesis is a holistic approach to diagnosing problems on a line. As this chapter demonstrates, there are a great number of areas where problems may exist. While some problems may be strong areas of interest for one transit agency at a particular time, the same problems may not be important to other transit agencies. In a following step, some research questions will be posed in the context of a real line in Chapter 4, as chosen by some combination of what problems seem important and what problems seem likely to be solvable given the constraints at hand. But before this application, it is worthwhile to first discuss the characteristics of available data and the general methods one can use to investigate different types of problems using that data. This is the subject of the next chapter.
Chapter 3
Data and Methods

This chapter describes general types of available data and methods that can be used in the analysis process. In contrast to the study of operational problems, the data types and methods used in the analysis are more similar across lines and systems.

A range of methods is necessary to understand current performance and the potential for improvement under ideal circumstances, and various types of data will be needed as input. This chapter reviews data types that may be available within a transit agency, collected automatically or manually. A second section discusses methods of analyzing this data. Finally, different representations of data are discussed as aids to understanding and describing system behavior.

3.1 Data

A critical part of understanding performance is assessing the type of data available within the agency, as well as data that can be gathered in the future. Data can include both observational data on system entities that may be useful for off-line analysis, and descriptive data on the system. Both types of data are useful for analysis, but particularly when used jointly. They can provide insight into the performance of the operating plan and service management strategies.

3.1.1 System data

Some data simply document the characteristics of the system and the operating plan. One important example of this is the vehicle and crew schedules that comprise the operating plan. Signal designs for fixed-block systems are typically documented on signal control line diagrams. Other diagrams may show civil speed restrictions, platform locations and track
geometry. Diagrams are typically updated only as frequently as changes are made to the signal designs and in agencies with older lines, the data may not be complete or up-to-date. There will also be a safe braking distance model and its associated parameters. Data on the track condition, including the time and location of slow orders on the line, are important. Other sources are vehicle maintenance and diagnostic information.

3.1.2 Observational data

Raw observational data on the operation of a rail line, which is the core of the performance analysis, may be obtained in many ways. The largest quantity of data and variety of data types comes from observational data since most aspects of operation are observable in some way. Observational data allows the identification of patterns, the construction of models, and ultimately a better understanding of a line’s operation.

Observational data collection can be automatic or manual. Automatic data collection may be computer-based or mechanical, and manual data collection may be handwritten or with mechanical or computer aids. There are many data items that are useful for various analyses.

Continuously recorded data is very useful, especially in the case of computer database systems in which the data may be manipulated right away. With mechanical recording systems, data must first be transformed into electronic form before it can be analyzed. There are certain types of events occurring in transit systems that are now automatically recorded at many transit agencies, mostly as byproducts of computer-based systems. New signal and control system technology, as well as new vehicles, may each be installed with computer data recorders, though generally not both at once. These monitor train movements and many other events from devices in the field or user actions at the workstations. Fare collection systems may include databases of individual fare transactions and uses. The level of information provided by
any automatic data collection system depends on the data structure of each particular installation.

Train monitoring systems exist mainly to bring current information on train location to a control center where service controllers or dispatchers have access to it. If the data passes through a computer or is transmitted to a computer via a communications link, then it can be stored in a database. The systems can be categorized by the level of detail on the information they provide:

- Point detection of trains, limited number of locations
- Track circuit occupancy monitoring, limited number of track circuits
- Track circuit occupancy monitoring, all track circuits
- Continuous location monitoring

Aside from continuous location monitoring, track circuit occupancy data is the most useful form of train monitoring system. The operations control system software can make use of algorithms to track trains through the system, based on track circuit occupancy. Track circuits do not read any information from the train, and so train-tracking algorithm must deduce which train is causing the occupancy. Train tracking is most accurate when all track circuits indicate occupancy, rather than a limited number of nonadjacent track circuits.

Automatic vehicle identification (AVI) devices can be used to track trains by identifying them through wayside devices. AVI can be used for a train monitoring system without track circuit occupancy information, or as an overlay to monitoring systems with some or all track circuits indicating.

Monitoring systems are commonly called Automatic Train Supervision (ATS) systems when they involve computerized display of automatically collected information. ATS systems can include a wide range of monitored information,
not just train location. Signal, switch, start light and terminal operation mode indications are commonly monitored and stored by ATS systems. Other information, such as power systems, intrusion alarms, fire alarms, escalators and elevators may also be included. Some ATS systems include train schedules, crew schedules and maintenance schedules. Modifications to the train and crew schedules made in real-time by a service controller may be stored, and provide valuable information for analyzing service control techniques. Service control and incident logs are typically input by control center personnel, and provide information on the times and types of incidents, the ways in which they were resolved, and any abnormal restorative techniques used. Data collection on vehicle maintenance and vehicle failures is typically routine.

Computerized on-board vehicle recording devices may provide further useful information such as speed commands, actual speed, door use and other information as for example the MBTA’s No. 3 Red Line cars. Such systems usually have a rolling time period for recording data after which the data is overwritten. Audio file recording of in-vehicle and on-platform public address announcements would provide a useful way to monitor the effectiveness of announcements during disruptions, but no reference to such systems was found.

Automatic Fare Collection (AFC) systems control the use of farecards, and store information on their use. In some transit systems, cards are used only upon entry to the system, while in other systems cards must be used both on entry to and exit from the system. Data structures vary, with more sophisticated systems tagging each farecard use with farecard ID, time of usage, type of card, location of use, fare type deducted, remaining value (for prepaid cards), and other information.

Passenger counting and tracking systems using infrared and video devices are produced for use in transit systems. One such commercial system is by Point Grey (2000). Applications of image-processing technology in monitoring
platform or in-vehicle crowding are now being introduced (see Lopez 1997, Velastin et al. 1994).

There are also non-computer continuous data collection systems that have existed for some time. These include Angus recorders, which are pen graphs that mark the passage of trains as spikes on rolls of paper (see Fellows 1990, p. 70). Many interlockings include these devices, to provide records in case of a collision. In some cases they are also used to provide locational information for service management. Another example of mechanical data systems is older turnstile equipment that contains mechanical counters tallying passenger entries and exits.

Daily manual data logs are common in rail line operations. Data on crew assignments, train arrivals, departures and passenger loads may be collected routinely by operations personnel such as supervisors in towers, terminals or on platforms. The STATIS pilot project has upgraded New York City Transit Authority’s paper-based sheets with a computer interface directly in the tower room, thereby transmitting manually input information in real-time to the control center, as well as eliminating data entry for off-line purposes. Data from real-time operations are typically used to produce reports on crew overtime usage and cause, as well as reports on train delays.

Non-continuous data sources, whether manual, mechanical or computer, can also be very useful. Not all system entities merit observational data recording at all times. For computer-based systems, the cost of purchasing and maintaining a full monitoring system, and the network associated with it, may outweigh the perceived benefits. Mechanical pen-graph devices, such as voltage meters, are outdated and are being replaced with computerized data feeds. Manual data checkers are costly, and involve scheduling the work, entering and cleaning the data. Thus, spot-checks using computer, mechanical or manual means are quite common in transit agencies.
Transit agencies usually have a staff of checkers for planning-related data collection activities. Traditionally, a checker makes notes on preprinted forms. The checker’s assignments are flexible in that anything a human has the capacity to monitor may be recorded. Typically, these include passenger boardings, alightings, volume, and train arrival and departure times. Manual checkers have started using electronic hand-held devices, eliminating paper-based data forms and the data entry process. NYCTA’s first use of hand-held devices for measuring departure and arrival times was conducted along the Queens Boulevard corridor in Spring 2001. Wong (2000) with this author conducted computer-aided monitoring of passenger flows passing fixed points within station passageways. This author also developed a software tool to computerize train-tracking data through manual input of model board indications.

Self-recording mechanical or computer devices may be temporarily attached to trains, to fare equipment, to power feeds, or placed in passageways or platforms. Data feeds on board trains can monitor speed, ATC speed commands and other information in real-time. Wayside power feeds can monitor voltage and power consumption. Passenger counting devices can be placed temporarily in passageways.

Rizzi and Guichoux (1997) proposed a data object model for the RATP bus system. Their work represents an effort by a transit agency to define important operations-related data to facilitate its use in new and interconnected software systems. The TITAN/1 project, which created the ‘European Reference data model for Public Transport operations (Transmodel)’, was a joint effort of several agencies and had similar goals including facilitating the development of real-time control systems.
3.2 Generated Information and Associated Methods

Because data does not always come in a format that is directly usable, and can be overwhelming in terms of volume, methods are needed to turn it into useful information. Sometimes the most pertinent data cannot be monitored directly but can only be estimated based on what is known. Even when it can be monitored directly, data reduction may be necessary to identify patterns and trends in a large quantity of data. Models and algorithms can be designed to process data into information leading to a better understanding of system performance. Data manipulation tools such as relational databases, spreadsheet applications and programming environments are means by which data can be analyzed, depending on the skill of the programmer. Data estimation, reduction and manipulation are all important parts of the analysis process as discussed below.

3.2.1 Estimation

Since some system characteristics may not be readily observable, estimation methods are useful for producing proxy data based on what is available. The unobserved characteristic may indicate a problem, or may be valuable information from which other problem identification or alternatives analysis tools can be constructed.

Passenger-flow models

A common example of data estimation is the creation of train-level passenger volumes based on passenger arrival rates and train movement data. One such system in routine use is BART's passenger flow model by Buneman (1984). Beginning with data on time and location of exit of each passenger, the model assigns passengers backwards in time to appropriate trains serving the origin-destination pair of each passenger as determined by farecard data, and results in estimates of boardings, alightings and volumes at each station.
for all trains. The resulting estimated peak loads for each trip are compared
with load standards (c.f. BART, 2000). The MBTA’s passenger counts, consisting
of a one-day composite of several different days, used a passenger flow model
assuming no passenger left behind and assuming travel times, and required
balancing to insure that the observed system ons and offs were internally
consistent (CTPS 1998).

Origin-destination models

In systems with older fare collection equipment, fare revenue is often
used to estimate system entries. Such methods can become convoluted,
especially in larger systems with complex route networks and fare structures.
In systems with newer fare collection technology, data estimation has been
used to estimate origin-destination flows. Navick (1997) used a model to
translate boardings data into origin-destination data for bus routes. Furth
(1997) created production software to begin with MBTA rail system boardings
and estimate an updated origin-destination matrix. Barry, Newhouser, Rahbee
and Sayeda (2001) used estimation methods to turn entry-sweep automatic fare
collection data into an origin-destination matrix for the New York City
Transit Authority rail system.

Train tracking algorithms

Another common example of data estimation is the train-tracking algorithm,
which integrates raw point detection data or track circuit occupancies into
train movements where no information on train IDs exists. Operations control
systems typically include train-tracking algorithms. Examples of train
tracking data for the MBTA case where all track circuits are included, will be
shown in Chapter 4.
Dwell time models

The estimation of dwell time models has been the subject of much research, including Puong (2000), Lam, Cheung and Lam (1999), Lin (1990), Krstanoski (1996), Mori (1988), Wirasinghe and Szplett (1984), Szplett and Wirasinghe (1984), Fritz (1983) and others. Lin (1990) estimated multiple linear and nonlinear regression models for two subway stations on the MBTA Green Line with high-floor light rail vehicles operating singly or in two-car lengths. Lin’s model accounted for 70% of the variation in dwell times in his data set. Puong (2000) estimated a single model for data from two MBTA Red Line heavy rail stations. Puong’s model accounted for 90% of the variation in his data set. Krstanoski (1996) modeled dwell time as a gamma process, noting that in his data set at Bloor station in Toronto, larger passenger flows resulted in more variable dwell times. Models may be applied to generate stand-alone dwell time data, such as in research by Barker (2001). Barker estimated necessary boarding and alighting portion of dwell times for comparison with larger preset dwell times in an Automatic Train Regulation system in a rail line not yet built. He used a dwell time model calibrated for similar vehicles running in an existing system.

Running time models

Running time models come in different forms. At their simplest, single equations are used to predict running time between two points. Models with more contextual information include the signal system logic and some representation of vehicle performance characteristics. Kichuchi (1991) presents a simple interstation running time model. Wong (2000) estimated running time models for a component in a real-time control model.
Service control models

Models of service control tactics have been proposed by many researchers. In many cases, the service control models were embedded within simulation models of a line segment.

Simulation models

Some estimation methods are designed to be embedded within larger simulation models. These have been either event-based or time increment-based, with the exception of Kiyotoshi and Masaru (1994) who designed a tool that switches between the two to achieve a combination of performance and detail. Event-based simulations included the following. Woodhead, Shaw and Marlowe (1964) created a simulation model of the then not-yet-built Washington DC Metro, using embedded running time models, a dwell time model based on passenger flows, a simple passenger flow generation model, and an energy consumption model. Welding and Day (1965) published a report on a simulation for the London Underground's Central Line and Victoria Line. Eichler and Turnheim (1978) created a combination continuous and event-based simulation modeling interstation running times. Harsch (1970) used a simulation to test control strategies that were under development at the time for use in BART's control system. Weston and McKenna (1987) created a simulation model of the London Underground Central Line to test various changes to service management and the operating plan. Minciardi, Savio and Sciutto (1994) created a simulation used to size the electrical supply network. Hill and Bond (1995) used a simulation to evaluate theoretical capacities of different signal system types. Krstanoski (1996) used a closed-form simplified running time model, allowing implementation within a spreadsheet, in a line-wide simulation. A similar approach was taken by Malavasi (1998), who likewise used the simulation to test various disturbed inputs. Heimburger et al (1998), in their model of interstation running time, modeled fixed-block ATC speed commands and some representation of crew behavior as well, although the crew
behavior model component was not based on observational data. Heimburger's model used time-increment simulation. Harder (1998) created a simulation of the Braintree terminal with the aim of modeling terminal operations. SYSTRA USA's Railsim has been applied to numerous rapid transit systems, including recently the MBTA Green Line (in the subway section only). Earlier Green Line central subway simulations have been produced by Carson and Atala (1990) and by Humphrey at the Central Transportation Planning Staff. A simulation modeling train regulation at terminals is described (in French) by Doras, Girardot and Heurgon (1978).

Other areas of data estimation include operating costs, which must be allocated by using cost models. Herzenberg (1981) proposed models for operator wage costs and applied them to MBTA bus routes. Other workforce cost models have been proposed in general by Perzli (1992) and for cover work by Hickman (1988) and Shiftan (1991). Antonisse (1985) proposed models of train operating cost with an application to the MBTA Red Line. Energy consumption models have been created, typically as part of simulation models such as for the Washington Metro, then called the NCTA (Woodhead, Shaw and Marlowe, 1964).

3.2.2 Data Manipulation Methods

When the formatting of data represents an obstacle to its use, data manipulation techniques are often applied. Tools for data manipulation are the various relational databases, spreadsheets and programming environments, and their combinations. In this thesis, data manipulation is meant to exclude the estimation models just discussed and focuses only on data transformation. Implementation of data manipulation programs is a matter of design, given that a variety of relational database and spreadsheet tools exist. The following are some examples of data items that can be extracted by manipulation of data from Automatic Train Supervision system (with a database storing track circuit level data).
• Dwell times (as defined in this section as platform circuit occupancy duration)
• Close-in times (defined as headway minus dwell time)
• Headways
• Segment-level running times
• Service control monitoring
• Delay/incident monitoring

It is difficult to make generalizations about when these data items will be necessary. They describe several components of the system's behavior, which enable sub-systems to be described, and then can be used directly in the problem identification process. For example, in order to test whether scheduled running times have been selected appropriately, an intermediate step would be to observe running time performance from a set of data. The intermediate steps of data manipulation such as the running-time example may form a significant amount of work for the researcher.

Extracting dwell times

Some Automatic Train Supervision systems log train tracking records and track circuit occupancy records, but not door opening and closing times. Dwell times can be extracted based upon the time a train occupies the track circuit in which the platform is located. The door opening time can be most accurately estimated from the most recent unoccupancy of the track circuit immediately upstream of the platform edge, as long as a boundary exists immediately upstream of the platform edge.

Data from ATS systems can provide different levels of data for dwell time monitoring. Some conventional rapid transit operations control systems may record the times of door opening and closing (such as WMATA). Additional information is needed in ATS systems if door opening and closing times are not recorded.
Service control monitoring

Service control monitoring would be useful if there are no other logs of expressing or short-turning actions. Expressing can be detected by calculating the dwell times in platform blocks and comparing their duration with threshold values. If short-turning is not recorded as an event, an SQL query may be written that detects short-turning through train-tracking records.

Segment running-time monitoring

SQL queries may be designed to translate train-tracking records into running time of trip IDs between specified track circuits. In a spreadsheet environment, more flexible functions can be written to accomplish the same goal, but on smaller data sets.

3.2.3 Data Reduction Methods

When there are large amounts of repetitive data, data-reduction techniques, also known as aggregation, are useful for summarizing while retaining the meaning of the data. Most data reduction activities carried out at transit agencies are to produce performance indicators. Automatic Train Supervision systems commonly have reporting features for automatic generation of performance indicators. Other measures are collected manually.

Repetitively-collected data can be summarized using various measures of the distribution, including the mean, standard deviation, variance, mode, median and percentiles. Additionally, statistical methods for comparing two populations can be used to infer the statistical significance of apparent differences.
The statistical approach is vulnerable to inappropriate assumptions about the form of distributions. Many data sets describing aspects of transit systems do not fit readily into any particular distributional form. What is often needed is an illustration of the frequency distribution before any measures of the distribution are calculated. This leads into the next topic for discussion: representations.

3.3 Representations

Tabular representations are a common way to convey information. Illustrations are visual representations of data used to convey the meaning of underlying data both to the researcher and to a wider audience. Illustrations often have a role in the decision-making process in the three areas of line characteristics, operating plan and operations control. Almost all reports, papers and theses use illustration techniques to make points supporting their claims. However, not enough use has been made of such illustrations. As discussed in Chapter 1, the methodological approach promulgated in this thesis is that there is a need for a rich description of the rail line's operation. There are many possible ways to illustrate data, but the illustration chosen to argue a point should make that point obvious in terms of its scale and normalization of the variables.

Some common representations are:

- Frequency distributions
- Time-space plot
- Terminal dispatching chart
- Data series (volume profile along line, volume profile at point, etc.)
- Crew duties
- Vehicle requirements
3.3.1 Frequency Distributions

Frequency distributions illustrate the probabilities of any of a range of values occurring. The value of a frequency distribution is equal to the count of data observations that fit into a specified range of values. Many researchers use distribution measures as described above, however they can be misleading, as for example in data sets where the mean is not equal to the mode, then using percentiles as measures of the distribution may be misleading.

3.3.2 Time-space Plot

The time-space plot is a useful representation of vehicle trajectories and events that occur at various times. Figure 3-1 (a) shows an overview of the time-space plot of train movements. Figure 3-1 (b) is a close-up of a single train trajectory. Time-space plots of train movements are a traditional method of railway schedule design, especially in planning when trains will meet and also in prioritizing the use of single-track sections. They have also been used in rail transit analysis. Bruun, Vuchic and Shin (1999) discuss the uses of time-space plots for operations planning and for service management. A thorough description of time-space plots is given in Rizzi, Froloff and Saporito (1989).

3.3.3 Overlaying Illustrations

Illustrations can provide a useful way of comparing two sets of data, typically presenting an observed set of data with another set of idealized, scheduled or different data. For example, Vuchic, Bruun and Krstanoski (1996) compared a one-day estimated load profile against actual passenger capacity along a BART line using a time scale. Antonisse (1985) overlaid volume passing a point by time of day with the passenger capacity by time of day. In Furth
Vehicle recovery

(a) Overview

(b) Zoom

Figure 3-1
Time-space plots
and Rahbee (2001), measures of actual stop spacing were overlaid with results from an optimization model. Krstanoski (1996) plotted measures of the headway distribution along a line, together with measures of the dwell time distribution along the line and the minimum capacity of the signal system along the line. Shriver (1981, p. 145) overlaid cumulative boardings and alightings along a line.

3.3.4 Multimedia

Multimedia illustrations are used only rarely, and for specific uses. Automatic Train Supervision systems typically have a playback feature where all the events observed on the control center workstations may be played back. Control centers typically have audio-tape logs of radio communications as well.

3.4 Discussion

In this chapter, an overview of data and methods used for operational analysis is presented. These methods have a crucial role in the analysis process that will be demonstrated in Chapter 4. Methods will provide answers to engineering questions, and will also be an invaluable tool in helping the researcher build a knowledge base of the system’s behavior.
Chapter 4
MBTA Red Line Application

This chapter applies the previously described methodology to the MBTA Red Line. Line characteristics, the operating plan, and service management should be investigated through intensive study of observational and system data, and of the human and institutional routines and culture by which the system works. This process typically reveals many idiosyncratic features of the transit system that are part of no scheduling manual, part of no guide to service management, are written down nowhere, and may not even be volunteered by personnel. It is only through this labor-intensive, somewhat unstructured process that one can study in proper context problems that may have been apparent from the start, through which new problems will become apparent, and through which attempts at solutions to those problems that will actually improve system performance can be found. In this chapter, examples of this method focus on several problem areas within line characteristics, operating plan, and service management on the MBTA Red Line.

The choice of issues studied was affected by numerous influences: the availability of data to describe subsystems relevant to any issue; the amount of work required to tease out identifiable problems through an iterative process of learning through staff interviews, verifying their comments and describing the system’s behavior; the extent to which the problem being studied had been ameliorated already by coping mechanisms within the framework of line characteristics – operating plan – service management; the likelihood of change given the agency’s institutional characteristics; and, the time and resources available to the researcher.
4.1 MBTA Red Line Description

4.1.1 Line Characteristics

A number of characteristics of the Red Line are important in that they have a strong effect on the performance of the line. These line characteristics include the terminals, crossovers, storage yard locations, signal system, fleet, central control and field supervision system, and other features, as described earlier in Chapter 2 of this thesis.

The Red Line is composed of two branches and a trunk portion, as shown in Figure 4-1 (a). The Alewife terminal, shown in Figure 4-1 (b), is located at the northwestern end of the line. It has crossovers on both sides of the platform, but the terminal is operated as a two-track stub end terminal with a center platform. The crossover on the far side of the station is used for put-in and layup moves only. On the southern end of the Red Line there are two branch terminals, Ashmont and Braintree. Ashmont is a relay terminal with dedicated arriving and departing side platforms, as shown in Figure 4-1 (c). Braintree is similar to Alewife with a center platform, crossovers on both ends of the platform, and operates as a stub terminal (see Figure 4-1 (d)). All three terminals can be controlled through Automatic Terminal Dispatching (ATD) with start bells and lights located on the platforms.

Between terminals, the two branches merge between JFK/UMass and Andrew stations using a flyover. Automatic crossovers are located at Alewife, Davis, Harvard, Park Street, JFK/UMass, North Quincy, Quincy Center, Quincy Adams, Braintree and Ashmont. Hand-throw crossovers are located at Kendall/MIT, South Station, Broadway, Andrew, Fields Corner and Shawmut. Storage yards are located beyond Ashmont at Codman Yard, north of Braintree at the Braintree Storage Track, beyond Braintree at Caddigan Yard, beyond Alewife at Alewife Yard, and north of JFK/UMass at Cabot Yard. A pocket track is located at Quincy Adams and is used for short-turns. The Braintree branch stations are
Figure 4-1
MBTA Red Line
spaced farther apart and are located further away from the downtown than the northern part of the line. At Park Street, platforms are used on both sides of each track. Peak volumes are approximately 18,000 passengers per hour, based on counts from 1997 (CTPS, 1998).

The Red Line is equipped with a fixed-block Automatic Train Control system with five speed commands: 0, 10, 25, 40 and 50 mph. The Red Line is part of the MBTA’s Operations Control System (OCS) that became operational in 1997. It is a full Automatic Train Supervision (ATS) system that monitors track circuit occupancy, performs train-tracking functions, and allows control of field devices such as switches, route management, schedule management, communications, and power control. An alarm system to alert dispatchers to headway gaps was developed as part of the original installation, but did not work properly. The MBTA and MIT jointly designed and implemented a terminal dispatching decision support system called the DCS, which is still being tested.

There are three car types currently used on the Red Line, with seated capacity of approximately 50 passengers per car and crush loads of 200 passengers per car. The No. 3 cars, numbered in the 1800 series, are the most recent, and send regenerated braking energy back into the third rail. The No. 3 cars have four doors per side, while the No. 2 and No. 1 cars have three doors per side. All doors face each other across the aisle (i.e. are not staggered). The train attendant, positioned near the middle of the train, controls the doors in two sections.

Vehicle service management is conducted principally through two to three dispatchers assigned to the Red Line in the OCC (Operations Control Center). Train starters located at JFK/UMass station are responsible for crew assignment. Chief Inspectors and/or Inspectors are positioned at terminals as well as at key stations for diagnosing train problems, responding to passenger incidents and checking in crew members. Motorpersons are instructed to operate
the train at the maximum allowable speed as determined by the ATC system. Some motorpersons will slow down a train in advance of an anticipated 0-mph code, in order to provide a more comfortable ride and prevent excessive wear and tear on vehicles. Each motorperson and train attendant operates only on one branch. This scheduling constraint also affects service management. Union work rules tightly limited the amount of built-in overtime, among other constraints. Song (1998) found that there were 323 incidents involving delays of longer than 15 minutes during a recent two-year period. He chose a sample of 57 incidents and found that 42% of the delays were between 15 and 20 minutes, with the remainder longer than 20 minutes.

4.1.2 Operating Plan

The major characteristics of the Red Line operating plan are shown in Table 4-1. Peak period headways in the trunk section are approximately 3.5 minutes, with an off-peak headway of 6 minutes with trains generally alternating between the two routes. The Red Line has two-person operation with motorpersons and train attendants not switching duties. There are only two scheduled running times in the weekday Red Line schedule, with one time only applicable to late night trips, and there are no intermediate time points. Scheduled train recovery times have been set based on past experience, and include some abnormal values which were required in order to compensate for problems in scheduled running times.

Trains are scheduled to operate as 6-car trains during the peak periods and 4-car trains during mid-day and evenings. During cold weather, supervisors may decide not to cut trains, instead running 6-car trains to avoid coupling problems. There are four crew relief points, located at Alewife, Ashmont and Braintree terminals as well as at JFK/UMass. Crews for each branch are scheduled independently and do not work more than one branch. The Red Line’s two branches operate to a common trunk terminal at Alewife. Two extra trips per day depart from the Braintree Storage track and begin service at Quincy.
### Table 4-1
Summary of MBTA Red Line Operating Plan

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Early AM</td>
<td>5-6 min.</td>
<td>49 min.</td>
<td>4 min.</td>
<td>4 min.</td>
</tr>
<tr>
<td>AM Peak</td>
<td>3.5-4 min.</td>
<td>49 min.</td>
<td>3-8 min.</td>
<td>8-21 min.</td>
</tr>
<tr>
<td>Base</td>
<td>6 min.</td>
<td>49 min.</td>
<td>36 min.</td>
<td>18 min.</td>
</tr>
<tr>
<td>PM Peak</td>
<td>3.5-4 min.</td>
<td>49 min.</td>
<td>2-6 min.</td>
<td>4-10 min.</td>
</tr>
<tr>
<td>Evening</td>
<td>6 min.</td>
<td>49-46 min.</td>
<td>4-6 min.</td>
<td>4-21 min.</td>
</tr>
<tr>
<td>Late Evening</td>
<td>6 min.</td>
<td>46 min.</td>
<td>5-4 min.</td>
<td>1-4 min.</td>
</tr>
</tbody>
</table>

(a) Weekdays

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All Day</td>
<td>6.5 min.</td>
<td>7.5 min.</td>
<td>14/15 min.</td>
<td>14/15 min.</td>
</tr>
</tbody>
</table>

(b) Weekends

**Summary of MBTA Red Line Operating Plan**
Adams. There are no scheduled turnbacks or scheduled express trains. The service day spans from approximately 5 AM to 1 AM.

4.1.3 Service Management

The regulation of terminal departure times and sequencing of trains is carried out at the Operations Control Center by dispatchers. Control Center dispatchers may order trains to hold, short-turn, run light (i.e. deadhead), or skip stops by using radio communications. Management of crews is performed at the Red Line Train Starter's Office located above JFK/UMass station. Specific vehicle and crew management techniques will be discussed later.

4.2 Data availability

4.2.1 System data

The MBTA Schedules Department maintains schedule data in a variety of formats. Some of these are in Giro HASTUS software format, and were printable but not exportable at the time of writing. The schedules are extracted from HASTUS and converted to a format recognizable by the Operations Control System (OCS).

Automatic Train Control (ATC) signal control lines are contained on signal diagrams. Based on extra data from the MBTA Signal Engineering Department, the existing Red Line ATC design differs slightly from the drawings for a small number of track circuits. Additional drawings for the relatively new Braintree Storage track exist. A contractor provided the MBTA with software for evaluating the existing MBTA Red Line signal design. It computed safe braking distances starting from the end of each track circuit, based on initial speeds allowed by ATC (i.e. 10, 25, 40 and 50 mph).
4.2.2 Observational data

The MBTA's Operations Control System (OCS) is an Automatic Train Supervision System (ATS) that stores train movement and schedule modification data in a database. A description of some relevant tables in the database and the fields they contain are included in Appendix A. The entry of a train into any of the Red Line's several hundred track circuits is stored with a time stamp to the nearest second, along with a train ID number and other information. Each trip also has a unique record that stores some of its attributes including the scheduled and actual departure and arrival times, allowing easy retrieval of end-to-end running times. This database has not been previously used to conduct operational analysis, and facilitates many types of valuable analysis previously impractical.

Crew assignment logs are recorded manually at the Train Starter's Office located above JFK/UMass station, one of four crew relief points. These logs contain detailed records of train arrivals and departures at JFK/UMass and crew identification numbers by trip. This includes both the crews that were nominally scheduled for each trip as well as the actual crews. Crew swaps, vehicle swaps and any incidents are also noted on the sheets.

Dispatchers' logs are maintained on an older system than the OCS, and are described further in Appendix A.

4.3 Problem Identification and Resolution

A listing and summary of the issues addressed in this case study is shown in Table 4-2. Problems that were investigated included the Red Line's ATC signal design, the intertiming of branch services during weekends and off-peak periods on weekdays, service management practices, the differences between actual and scheduled running times, and train recovery times.
### Table 4-2

**Summary of Issues Studied**

<table>
<thead>
<tr>
<th>Line Characteristics</th>
<th>Issue</th>
<th>Reasons Studied</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Signal design</td>
<td>Historical documents showed that pre-1987 signal design supported 2 minute headways. Queuing delays were known to affect current service quality.</td>
<td>Alternative design recommended for further study. Final design and implementation pending.</td>
</tr>
<tr>
<td></td>
<td>Charles/MGH to Park Street, southbound track.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating Plan</td>
<td>Running time performance.</td>
<td>Crew and supervisor comments of not enough half-cycle time. Schedules department interested in running time recommendations.</td>
<td>Some problems identified, both too much and too little running time.</td>
</tr>
<tr>
<td></td>
<td>Trunk interleaving on weekends.</td>
<td>Problem apparent from data. Resolution would require no significant cost.</td>
<td>Schedules modified slightly, operations staff made aware of problem. Improvement of 10-15% in waiting time at many stations on weekends.</td>
</tr>
<tr>
<td></td>
<td>Terminal dispatching.</td>
<td>Problem apparent from examination of data.</td>
<td>Recommendation to hold trains at branch terminals for even headways at merge. Communicate clearly with personnel the reasons for the holding.</td>
</tr>
<tr>
<td></td>
<td>Terminal congestion.</td>
<td>Problem apparent from examination of data.</td>
<td>Recommendation for study.</td>
</tr>
</tbody>
</table>
The methodology proposed in this thesis starts with analysis and representation of data, and thus an important part of this process has been the generation of illustrations based on train movement data. Several main forms of illustration were used. Time-space plots of train movements were generated for dozens of sample days. Frequency distributions of end-to-end running times were plotted by time of day for all data in the database, by timetable and schedule type. Numerous operational issues were raised through inspection of these time-space plots, although they do not illustrate crew management or terminal dispatching very well. The end-to-end running times were compared to their scheduled values, as well as to the values of vehicle half-cycle times. For a smaller sample of data, other representations were produced, including a comparison of the directed half-cycle times as determined through the start light indications and the actual half-cycle times (based on actual departure times). There was prior knowledge of a capacity problem in one area of the line, which prompted an inquiry into the underlying causes and possible remedies.

In most cases, the following sections do not trace through the intermediate steps that were necessary to identify the problems and in each case, analysis could have been carried further. As Table 4-2 describes, while some cases have produced concrete results, in several cases further study or action is recommended before potential changes are fully defined and implemented.

4.3.1 Signal design, Park Street, southbound

Riders of southbound trains in the rush hours are used to trains stopping frequently approaching Park Street, typically beginning at some point over the Longfellow Bridge between Kendall/MIT and Charles/MGH stations. The bottlenecking is one cause of the increase in end-to-end running times during peak periods (the other major cause is that dwell times increase). The bottlenecking causes increased wear and tear on vehicles, which alternately
brake and accelerate on their way through the area during congested hours of the day. The bottlenecking also wastes energy due to the unnecessary starting and stopping.

Heimburger, Herzenberg and Wilson (1998) showed that removing two peak-period trips from service would result in improved service quality since the bottleneck was causing slow running speeds. However, as shown in Figure 4-2, despite the schedule change resulting from the Heimburger et al. analysis (see description of Alternative II in Heimburger et al.), the travel time from Alewife to South Station still increases significantly during both the AM and PM peaks. Heimburger’s simulation results predicted that the ten-minute increase in travel time over the PM peak would have been reduced to a four-minute increase after the schedule change. As Figure 4-2 shows, travel time on a normal day may still increase by nearly ten minutes, even after the change. Heimburger’s dwell time model is the most likely source of error, as it was not properly verified against observational data at different stations.

Further, as shown in Figure 4-3 the sections of track between Kendall and Park Street southbound have more variable travel times than any other section in that direction. This feature is a result of the capacity of the current signal design being reached. Investigation of the signal design found that a considerable capacity increase is possible with only minor modifications.

The current ATC block layout and speed command logic is shown in Figure 4-4, along with an illustration of some areas where significant improvements are possible. There is an interlocking just northwest of the Park Street platform with two crossovers. Based on examination of the design and input from signal engineers at the MBTA and outside the MBTA, it appears that the design resulted from numerous constraints, only some of which are safety-related. The designs were engineered to meet a specified input headway while both
Figure 4-2
Travel Time, Alewife - South Station
January 22, 2001
Figure 4-3
Interstation Travel Time Distribution along Red Line
11/10/2000
3:30-7:00 PM
Proposed loop in 247T

Figure 4.4
ATC Signal Control Lines
Southbound track, normal direction (through moves only)
minimizing the amount of field equipment and satisfying safe braking distance constraints.

The speed command logic in the area of the interlocking prevents a train from entering the crossover until the preceding train has exited the track circuit between Park Street and Downtown Crossing, even when the safe braking distance would not be violated. This design feature is typical signal design philosophy, which is intended for application on flat junctions to avoid trains preventing movement of other trains. However, the Park Street interlocking is not a flat junction, and preventing trains from entering the crossover does not permit any additional operational flexibility.

The effect of the restrictive lookahead logic is to hold a train far enough behind its predecessor at Park Street that the minimum close-in time of a train entering Park Street is typically greater than it need be. Analysis of a typical PM peak showed that the minimum close-in time at Park Street southbound is no smaller than 1.4 minutes, while at Charles/MGH southbound it is 1.0 minutes. This extra time is most likely not necessary to insure a safe braking distance, especially given that Charles/MGH dwell times are less than the downstream Park Street dwell times. These restrictive lookaheads are currently a major bottleneck on the line.

A second problem is immediately downstream of the restrictive lookaheads. A train entering Park Street will receive a 0-mph speed command if its predecessor is still occupying part of the Downtown Crossing platform. This is necessitated by the current configuration of track circuit boundaries and by the safe braking distance. However, the safe braking distance (including a 12 foot vehicle overhang) from the end of the Park Street track circuit 274T is 586 feet at 10 mph, which overlaps into the Downtown Crossing track circuit 257T by only 21 feet. These safe braking distances were obtained from the MBTA and had been calculated for the Red Line’s Number 3 cars.
This situation can be remedied in the following way. A wire loop device, known as a b-point, may be installed at the 6-car marker of the Park Street track circuit 274T. This location is at least 20 feet from the track circuit boundary, which is located west of the Park Street end-of-platform point. The b-point would, if a train were occupying part of the Downtown Crossing track circuit 257T, bring the speed command in the track circuit 274T from 10 mph down to 0 mph. This would occur only if the train operator failed to stop at the Park Street platform. Trains would thus be able to enter Park Street at 10 mph and berth, even while a train is occupying Downtown Crossing. Boundaries for audio-frequency track circuits are typically placed 30 to 50 feet away from the station platform edge, to avoid pre- and post-shunting of the adjacent track circuits.

The two corrective suggestions mentioned here would work best if implemented together, otherwise only a portion of the benefit would be achievable. According to MBTA operations staff, it is relatively common to have a train operator request to use the ATC bypass to enter Park Street southbound at 10 mph.

If both signal design changes were implemented, the design would reduce minimum peak-period headways by an amount equivalent to, or slightly greater than, the dwell time of a train at Downtown Crossing in the peak, plus the unnecessary portion of Park Street close-in time. Since peak-period dwell times can reach 60 seconds at Downtown Crossing (see Figure B-2 in Appendix B), these changes could reduce minimum headways by about a minute. The potential investment in the signal redesign would help clear radios, improve travel times, improve delay recovery, reduce vehicle wear and tear, and reduce peak energy demand. Since the MBTA’s power cost structure is heavily determined by peak-hour demand charges, this design change would also reduce energy costs.
4.3.2 Running Times and Half-Cycle Times

The current Red Line schedule is the result of many small modifications made over an extended period, rather than being designed from scratch. With the data available within the MBTA’s Operations Control System database, it is possible to compare actual end-to-end running time performance versus the scheduled running times and half-cycle times. Using data manipulation, illustrations such as Figures 4-5 and 4-6 were produced to illustrate the frequency distribution of running times by pick, by schedule type and by route. An application was developed to produce these charts for four picks totaling one year, yielding sixteen illustrations using the entire population of running time data. Using data reduction, measures of the actual running time distributions including the mode (the typical value), average and percentiles were plotted along with the scheduled values. The current Red Line schedule uses a single running time for most of the day, with varying half-cycle times. It is clear that there are many discrepancies between actual and scheduled running times.

There is a question of what measure should be used for running time data reduction and statistical summary. Figures 4-5 and 4-6 show several measures of the running time distribution for the fall 2000 pick (see Appendix B for additional running time charts). These include the mode, also referred to as the ‘typical’ value, the median, the average, and various percentiles. As Figures 4-5 and 4-6 show, the average and mode are typically not identical. Because the data set includes observations in which trains were delayed due to incidents, the average will be affected by the incident data. By using the mode, the most typically occurring running time value, the influence of problematic observations is minimized. The average and the median values differ from the mode by up to several minutes. The typical running time shown on Figures 4-5 and 4-6 is equal to the mode value within a 15-minute bin of scheduled departure time at the originating terminal, with one-minute bins for the running time.
Running Time Alewife-Braintree
Weekday Schedule
Fall 2000

Figure 4-5 (a)
Running Time Braintree-Alewife
Weekday Schedule
Fall 2000

Figure 4-5 (b)
Running Time Alewife-Ashmont
Weekday Schedule
Fall 2000

Figure 4-6 (a)
Running Time Ashmont-Alewife
Weekday Schedule
Fall 2000

![Graph showing travel time vs. actual departure time at Ashmont, with various lines representing different running times and percentiles.](image)

Figure 4-6 (b)
In the case of the current Red Line schedule, scheduled running times are less important than the scheduled half-cycle times. This is due to the use of a single running time period for most of the day, which has rendered the nominal end-to-end running time an almost useless number, except for generating the end-of-shift times for swing-offs occurring at the terminals. Thus, if nominal running times were modified based on new information, but half-cycle times were left as is, then some efficiency could be obtained from removing somewhere between 5 and 8 minutes of paid time from the end of each shift (for shifts ending in non-peak periods).

Distributions of actual recovery time spent at each terminal showed that it is typical for trains to spend as little as 3 to 4 minutes on Braintree or Alewife platforms in peak periods before returning to service. This may be an appropriate amount for vehicle recovery time, but for crews it seems too little. This may be contrasted to prior Red Line schedules, which were more generous in their allotted crew recovery times.

Alewife-Braintree running times during the AM peak show that the typical actual running time is extremely close to the scheduled half-cycle time. In this case it may be necessary to redesign the half-cycle times such that resources would be redistributed or added to allow the schedule to be feasibly maintained on a day-to-day basis. Alewife-Braintree charts also show that at approximately 10 PM there is a momentary sharp reduction in crew recovery time, which seems surprising upon first glance.

The scheduled running times need to be reset and the half-cycle times carefully examined, considering terminal congestion, crew recovery, modifications to train length, and time spent clearing trains to go into storage. Reconfiguration of the ATC logic as suggested in Section 4.3.1 would require adjustments to running times that could not be based directly on the running times shown here.
4.3.3 Trunk intertiming

On a branching line such as the Red Line, it is difficult to maintain even intervals between trains after a merge. Schedules must be accurate, trains must be able to leave when planned, and respacing of intervals at the terminals must be done with the trunk headway intervals in mind. The intertiming of branch services in the northbound direction is a critical factor in insuring even loading of trains and the minimization of waiting times in the trunk section. Figure 4-7 is a time-space plot showing a Sunday, in which trunk intertiming causes very bunched service and thus high passenger waiting times. A prior suggestion to MBTA scheduling personnel was to take recovery time away from Ashmont branch trains at Alewife and add it to the Ashmont end, such that Ashmont trains would leave later. Analysis showed that the discrepancy in the intertiming would have required an adjustment of three minutes for the Saturday schedule and two minutes for the Sunday schedule. This change could be made across the day on Saturdays and Sundays because the headway is constant for the entire day, unlike in the weekday schedule.

In response to the time-space illustration, MBTA operations personnel made direct field observations and produced a small redesign of the schedule that was different from that recommended by this author. Braintree departures were set to be one minute earlier for all departures in all schedules, including the weekday schedule, with the exception of the first few AM trips because crews did not want to report to work any earlier. It may have been inappropriate to change weekday departure times by one minute across the entire day, as previous examinations had indicated that there was no consistent intertiming problem on weekday except for the earliest and latest trips. Thus, the schedule change had only a small impact on the poor intertiming of weekend service, and did not necessarily improve weekday service.
Average Daily Previous Headway (min.)

<table>
<thead>
<tr>
<th>Date</th>
<th>Headway (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/4/00</td>
<td></td>
</tr>
<tr>
<td>3/18/00</td>
<td></td>
</tr>
<tr>
<td>4/1/00</td>
<td></td>
</tr>
<tr>
<td>4/15/00</td>
<td></td>
</tr>
<tr>
<td>4/29/00</td>
<td></td>
</tr>
<tr>
<td>5/13/00</td>
<td></td>
</tr>
<tr>
<td>5/27/00</td>
<td></td>
</tr>
<tr>
<td>6/10/00</td>
<td></td>
</tr>
<tr>
<td>6/24/00</td>
<td></td>
</tr>
<tr>
<td>7/8/00</td>
<td></td>
</tr>
<tr>
<td>7/22/00</td>
<td></td>
</tr>
<tr>
<td>8/5/00</td>
<td></td>
</tr>
<tr>
<td>8/19/00</td>
<td></td>
</tr>
<tr>
<td>9/2/00</td>
<td></td>
</tr>
<tr>
<td>9/16/00</td>
<td></td>
</tr>
<tr>
<td>9/30/00</td>
<td></td>
</tr>
<tr>
<td>10/14/00</td>
<td></td>
</tr>
<tr>
<td>10/28/00</td>
<td></td>
</tr>
<tr>
<td>11/11/00</td>
<td></td>
</tr>
<tr>
<td>11/25/00</td>
<td></td>
</tr>
<tr>
<td>12/9/00</td>
<td></td>
</tr>
<tr>
<td>12/23/00</td>
<td></td>
</tr>
<tr>
<td>1/6/01</td>
<td></td>
</tr>
<tr>
<td>1/20/01</td>
<td></td>
</tr>
<tr>
<td>2/3/01</td>
<td></td>
</tr>
<tr>
<td>2/17/01</td>
<td></td>
</tr>
<tr>
<td>3/3/01</td>
<td></td>
</tr>
<tr>
<td>3/17/01</td>
<td></td>
</tr>
<tr>
<td>3/31/01</td>
<td></td>
</tr>
<tr>
<td>4/14/01</td>
<td></td>
</tr>
<tr>
<td>4/28/01</td>
<td></td>
</tr>
<tr>
<td>5/12/01</td>
<td></td>
</tr>
<tr>
<td>5/26/01</td>
<td></td>
</tr>
<tr>
<td>6/9/01</td>
<td></td>
</tr>
<tr>
<td>6/23/01</td>
<td></td>
</tr>
<tr>
<td>7/7/01</td>
<td></td>
</tr>
<tr>
<td>7/21/01</td>
<td></td>
</tr>
</tbody>
</table>

Passenger Wait Time Reduction
6 mo. Before and 6 mo. After Schedule Change
Northbound Trains

Figure 4-8

Daily Average Previous Headway
JFK Junction, Northbound Trains
Sundays Only

Figure 4-9
However, passenger waiting times were reduced at trunk stations by between 5 and 10 percent on Sundays, and between 2 and 5 percent on Saturdays, as shown in Figure 4-8. Figure 4-9 shows that the schedule change resulted in a minor improvement in a time-series comparison of all Sundays 6 months before and after the change. In Figure 4-9, the month of April 2000 is also included, as April 2000 data had been used in the original recommendation to operations personnel. It can be seen that Braintree trains still arrive after a much longer headway gap than do Ashmont trains—suggesting that further work to improve the intertiming of weekend service would result in further reductions in passenger wait times at northbound trunk stations.

4.3.4 Balancing of branch loads

Comments from MBTA scheduling personnel indicated that Braintree trains were more heavily loaded than Ashmont trains in the peak periods. The current schedule has approximately equal capacity supplied to the two branches, with the exception of two additional peak-direction Braintree branch trains. Re-analysis of passenger counts published in 1997 verified that passenger loads were heavily imbalanced with a ratio of almost 2 to 1. Passenger load profiles for AM peak northbound trains are shown in Figure 4-10 (a), and for PM peak southbound trains in Figure 4-10 (b). Due to capacity constraints, as shown in Heimburger et al (1998), it would not be productive to increase net service. One option may be to reduce Ashmont service while increasing Braintree service. If signal design changes proposed earlier were implemented, an increase in Braintree service without cutting Ashmont service would be more feasible.

4.3.5 Service management

It would be a useful research project to create an inventory and description of current Red Line service control techniques. As with all rail lines, the topic of service control is familiar to personnel within the
Figure 4-10
transit agency. What is being suggested is simply creating a written
description of the techniques and a summary of the constraints that went into
each decision. A useful way to build such an inventory as a researcher is to
first identify patterns from time-space plots based on train tracking data,
dispatcher's train sheet data, crew assignment logs of dispatchers' logs. By
learning first from the actual data, the researcher can prompt operations
personnel with questions which are detailed enough that they can respond on a
suitably detailed level. Without the researcher's initial work of identifying
patterns in the data, the questions he or she may pose to a dispatcher or
train starter may not be specific enough for them to respond in any detail.

It was not possible in the time available to create a comprehensive
overview of service control on the Red Line, nor to compare the frequencies of
use of each technique. However, this would be useful information both for
researchers and for the training of dispatchers.

4.3.6 Major disruption example

An unusual strategy using elementary service control actions used on
Monday, March 19, 2001 demonstrates the nature of the constraints involved in
service management decisions. Time-space plots are shown for this day in
Figure 4-11 (a) and (b). Due to a four-hour power outage at South Station
resulting from a track jumper, the line was operated as three loops: one from
Alewife to Park Street, another from JFK/UMass to Ashmont and a third from
JFK/UMass to Braintree with buses operating between Park Street and JFK/UMass.
There was a period of over one hour in which there was no service between Park
Street and Central and the loop established from Alewife to Park presents a
unique application of service management. All southbound trains were held in
all stations for approximately 5 minutes during the loop operation, thus
increasing typical Alewife to Park Street travel time from 18 minutes to 45
minutes. The northbound direction was not held, except at Davis, which is one
station away from the Alewife terminal.
The capacity of turnbacks at Park Street station is limited, as only the southbound track is signaled for reverse moves. To be able to use both Park Street tracks as a temporary terminal, for trains entering the northbound track from the southbound direction, the dispatcher must set up the route into the northbound track as if there were a train in the northbound track attempting to go to the southbound track (i.e. in reverse). The southbound train must then proceed through the interlocking under the 10-mph Automatic Train Control bypass feature, through a double-red signal and use line-of-sight to avoid any potential collision. This technique would only be used by the most experienced of dispatchers.

Because the incident occurred during the peak, there were many trains isolated between Park and Alewife, with capacity to store only three trains at Alewife. This left too many trains in operation with no obvious opportunity to remove more of them from service. Thus, in order to keep all trains moving, the peak direction trains (northbound) were not held, but off-peak direction trains (southbound) were held to increase the cycle time, thus providing stable service with the greater number of passengers aboard the faster trains and the lesser number of passengers aboard the held trains.

One potential way to have reduced the amount of holding necessary would have been to store one or two trains on the northbound track at Park Street station and in the tunnel between Park Street and Downtown Crossing by setting the route up in reverse, as described earlier. Control Center personnel had indicated in interview that setting up routes in reverse would be acceptable for one or two trains, but not for continuous operation due to the potential safety hazard. By reducing the number of trains in service in the Alewife-Park Street loop, holding times could have been reduced.
4.3.7 Terminal dispatching

There is little formal knowledge on how control center dispatchers carry out the respacing of service intervals of Red Line trains at terminals, despite the model for Red Line terminal dispatching proposed by Song (1998). With this lack of contextual knowledge, and a lack of knowledge on respacing of rail lines with real-time information available, despite Eberlein (1995) and Puong (2001), it was not possible to evaluate terminal dispatching by comparing current dispatching practice with the model’s recommendations. Soeldner (1993) and Deckoff (1990) had taken such an approach to analyzing control actions on the MBTA Green Line, but as discussed in Chapter 2, their assumptions on the objectives of control actions were not well founded.

Figure 4-12 shows a time-space plot for December 7, 2000. This case shows poor coordination of the “flexed” schedule, exhibiting many large northbound service gaps throughout the evening. It appears from this data alone that these gaps could have been reduced with greater attention to the holding times of trains at Ashmont and Braintree. On this day, power was shut off for the northern end of the line due to track fires, and trains were short-turned at Harvard and Park Street.

4.3.8 Expressing, and trains out-of-sequence

Red Line departures are generally scheduled to alternate between Ashmont and Braintree, with the exception of two extra Braintree trips northbound during the AM peak and southbound during the PM peak. When these extra trips are run, some Ashmont passengers may see two Braintree trains before an Ashmont train. Quite frequently, trains traveling northbound on either branch may arrive at the JFK/UMass junction out of order. Usually, it is because one train is significantly late. If the other train is not held, then the trains will enter the trunk out of sequence. If the hold necessary to prevent incorrect sequence would be too long, or in the direction of peak passenger
travel, the dispatcher will typically allow the trains to merge out of sequence.

Upon arrival at Alewife, there may be an opportunity to correct the sequence of trains. It is considered acceptable practice to allow occasional back-to-back Braintree trains in the peak hour, but it would be unacceptable practice to allow two successive Ashmont trains southbound. When trains enter the trunk northbound out of sequence, they inevitably cause an unscheduled occurrence of two Braintree trains followed at some point by two Ashmont trains, or vice versa. An exception to this occurs when one of the trains out of sequence is one of the two scheduled layups at Alewife during the AM peaks, which are the two extra trains from Braintree.

According to operations personnel, it is undesirable to reassign a train or crews from one branch to another except in emergencies. Crews pick to work on a single branch, and the vehicle destination signs cannot always be changed easily. In order to correct the sequence, the dispatcher performs a '1"-in-2" out' procedure. This is possible when the headways between the trains involved are short enough that holding the 1" arriving train to leave 2" will not cause an extremely large gap southbound. The 1"-in-2"-out procedure is common during the peak periods, especially during the PM peak when sequence is critical to avoiding overloading on Braintree-bound trains at South Station.

Expressing is used on the MBTA Red Line, but there are indications from operations personnel that it occurs less frequently than in the past. This is following the views of the MBTA operations management, who had recommended a reduction in the amount of expressing due to the number of passenger complaints received through the MBTA's recently instituted "Write To The Top" campaign. It is possible to identify cases where, given the limited information available for evaluation, expressing probably should have been carried out.
The example shown in Figure 4-13 (a) and (b) is from December 6, 2000. Northbound trains that came into the trunk section out of sequence made all stops to Alewife, arriving at Alewife at 5:40 PM, 8:35 PM and at 9:55 PM and then a 1"-in-2"-out procedure was carried out to correct the sequence. In each case, the first train after the large gap appears to have been overloaded, and following trains were clearly delayed by the first train. Thus, all three trains involved in the 1"-in-2"-out procedure were delayed, which caused a corresponding gap in the southbound direction. It appears that expressing the first train northbound between the typical express stations (i.e. Park or Charles to Harvard and Harvard to Alewife) would have more evenly distributed the loads and potentially provided faster service to the majority of passengers.

The resequencing procedure is more difficult in the off-peaks or on weekends, when headways are long enough that the procedure would cause a very large headway gap. A sequence problem is easier to correct when headways are short because (a) the gap before the out-of-sequence train is likely to be shorter and (b) the time between the first train and the follower is likely to be shorter. Reason (b) more strongly determines the gap length leaving Alewife because it determines the time that both trains must wait in order to comply with a 1"-in-2"-out strategy. This would explain why on Thanksgiving Day 1999, the sequence problem could not be corrected from late morning until the end of the day, as shown in time-space plots in Figure 4-14 (a) and (b). Shorter headways also reduce the amount of passenger annoyance resulting from trains being out of sequence.
The resequencing can potentially be aided by expressing the first late train on its way to Alewife. A train entering the trunk out of sequence should be expressed for some northbound segment if there is (a) enough room in front of the train, (b) a train of the other branch directly behind. If significant travel time savings cannot be achieved through expressing the first train, or if the following train is not close enough behind to be slowed down by the out-of-sequence train, then no control action should be applied along the line (unless the sequence of trains leaving Alewife is not important).

4.3.9 Terminal congestion

Trains carrying passengers can be delayed approaching the end of a line because the platform tracks and/or the interlocking ahead are occupied. Terminal congestion problems can be monitored through time-space plots, or assessed using a larger set of data. One effective way of showing terminal congestion problems is the distribution of travel time through the last interstation section of track approaching the terminal. This method will obscure instances where trains were held at the upstream station, but it will show trains stopped or travelling slowly between stations. Once the severity of the problem has been assessed in terms of passenger service quality, identifying the underlying causes is the next step. There are multiple possible causes of terminal congestion including poorly regulated arriving headways, too much scheduled vehicle half-cycle time, cancelled route requests, too little crew recovery time, or other factors.

In two-track stub-end terminals such as those at Alewife and Braintree, vehicles cannot occupy the platforms for long without delaying incoming trains. In order for trains to occupy the platform for less time, crews in some systems may be scheduled to fall back one or more trains. If crews do not fall back, as is the case on the Red Line, then the trains must occupy the
station until the crew switches ends of the train and is ready to leave. This takes valuable time that may cause delays to approaching trains.

If the track network allows operation of the terminal as a relay, then the delays may be avoided while still giving crews adequate recovery time. Thus the type of terminal operation (e.g. stub versus relay) can be viewed as a ‘cause’ of the delays. It would be useful for an examination of terminal congestion to be focused on the contributing factors that are perceived to be changeable.
5.1 Summary and Conclusions

Most prior research on rail transit operations consists of models describing rail operations based on the researcher's own perceptions of system behavior. Almost no research focused on drawing out a knowledge base from transit agency personnel about operating practice. Especially with regard to service management practice, the expertise exists but is largely undocumented in the literature. The common research approach often lacks effective mechanisms to insure that the incorporated ideas of how the system operates are in fact correct.

There are constructive ways of identifying operational problems in high frequency rail service. A general process was developed during the course of study of the MBTA's Red Line for this thesis. That generalized problem identification process aims at drawing out current practice in operations, and should be based on climbing the following "ladder" to be most effective:

1. Careful analysis of real-time observational and system data. Draw plots of different kinds--time/space plots, time/time plots (cf. RATP, 1989) of interval respacing, as well as other representations such as those in Furth (1981), Herzenberg (1981) and others.

2. Identification of patterns in data. Pay particular attention to respacing of intervals at the terminal, and also to put-ins (pullouts) and lay-ups (pullbacks), and whether they are inserted into service in the scheduled order. Pay attention to how expressing, short-turning and holding are used and what makes their use in that context different from other uses on the same line and on other lines.
(3) Discussion of the observations with the personnel making the decisions. Then they can explain why they took the actions they did—until the illustration shows enough information to tell whether they made the best decision or not. Keep going back and forth between the illustrations and people—until the illustrations contain enough information that someone experienced could tell whether the service management decisions should have been made differently.

(4) At the end of this process two things result: First, it yields a description of the line’s operation. This will be useful for future research. Second, if the illustrations contain enough information (based on what was learned in (3)) some of them may illustrate occurrences when decisions affecting operations could have been carried out more effectively. Perhaps certain trains should have been held longer, or expressed, or berthed at the platform in a different order. In these cases, problems have been identified. Drawing out the rules of thumb by which service management is carried out, or any area within line characteristics and the operating plan, will give an opportunity to evaluate the rules of thumb themselves. Only then can the researcher address whether the objectives are effective and the decisions supportive of them.

Analysis focused on the operating plan seems more likely to result in improved system performance than analyses into changing line characteristics or service management practices—namely, because schedules are typically revised four times per year and require no retraining or capital costs, and ‘buy-in’ is necessary at one time only.

This application has also illuminated the fact that one’s work may instigate actions within the transit agency that are not consistent with one’s intentions. The presentation of the research may become a lightning rod, and precipitate unanticipated reactions. Once, in a presentation, this researcher explained as a side note that running times in a certain track segment could
not be monitored due to a track circuit problem that had existed for a long time. After the presentation, a manager ordered that the circuit be fixed—a result not intended by the presentation. In another presentation, this researcher explained that there were delays between start light indications and actual departures—without stating whether the problem was a significant cause of performance problems. A manager exclaimed that this problem had to be fixed. The fact that there may be many interpretations of the problem must be kept in mind by anyone hoping to improve transit system performance.

In a study of the MBTA Red Line, the four steps listed above proved invaluable in both drawing out practice and identifying problems in several areas. These included the layout of signal blocks, logic of speed commands available, choice of end-to-end running times, the scheduled intertiming of branch services, holding at branch terminals for even headways at the merge location, the use of expressing and the practice of resequencing branch trains at the trunk terminal.

Initial modifications to the schedule resulted in a decrease in passenger waiting time on Sundays between 5% and 10% at certain stations, with no increase in costs. Further changes have the potential to further reduce weekend passenger waiting time significantly. Recommended reconfiguration of the signal design in one area would reduce the peak period southbound minimum headway by the length of typical peak period Park Street dwell time—which may be approximately one minute, barring any unforeseen problems.
5.2 Future Research

Future research could take several general forms, including the following:

(1) Follow-on investigation of the effects of a Park Street ATC design reconfiguration. If implemented, the change would cause a reduction in the southbound running times towards the end of each peak period. The schedules would need modification and further monitoring, both of which could be assisted with the analysis of real-time data from the OCS database.

(2) Following in the footsteps of Rizzi, Froloff and Saporito (1989) one could describe the way service management is done on a rail line. This includes descriptions of resequencing and respacing of intervals at the terminal. Modeling should be left until later, with the first step being a description the way the system works through intensive study of system and observational data.

(3) Analysis into any part of the framework on any rail line, or specifically the MBTA Red Line. Several items discussed in Chapter 4 could warrant further study. Terminal dispatching practice on the MBTA Red Line has not been well documented. Diagrams of the type used in Rizzi, Froloff and Saporito (1989) could be produced, showing the relationship between the nominal schedule and the ‘directed’ schedule that is actually run. These diagrams were not straightforward to produce for the MBTA given the data structure of the OCS train sheet and the database table SCHED_TRIP. Unfortunately, the format of the OCS’s internal vehicle schedule is convoluted and undocumented. Thus, to determine the true order of vehicle runs the data must be input manually. However, the process of creating the diagrams themselves would involve the uncovering of operating philosophies. It would necessitate the monitoring of occurrences in which trains move up or down one slot in the schedule, when trips are dropped, and how often respacing takes
place. These are common service management techniques that need to be better understood.

(4) Challenging of the central arguments of the thesis first posed in Chapter 1, that suggested ways in which independent research can be made more useful for transit agencies. The literature review and experience from transit agencies on which these claims are based was limited. Further literature review across a wider range of transit agencies, independent research institutions and industry consultants could provide further insight into the ways to insure that independent research will be more useful.
Appendix A

Description of OCS Database and Other Data

The MBTA’s Operations Control System (OCS) is an Automatic Train Supervision System (ATS) which monitors trains and other information and control field devices. The OCS covers the three MBTA heavy rail lines (Red, Orange and Blue). A Sybase database stores all indications of device status change from the field as well other information used by the OCS. As of December, 1999 there were 72 tables in the database. The active database contains a rolling one-month period of data, while twelve historical databases contain one month each. Databases over twelve months old are archived to DAT tape.

Two ASCII files used by the OCS (mbta.dev and mbta.gph) contain information on the devices in the database (such as track circuit length) as well as data for generating the three types of user displays (overhead, line view and page view).

On certain days, the time stamps recorded in the OCS database exhibit lag times on the order of seconds to minutes. These lags are most apparent in cases where train tracking records were plotted on time-space plots. Observation of the underlying records indicated that on some days, there were as many as 250 records on the Red Line in which three successive records occurred within the same second. These occurrences are unrealistic and indicate a lag time problem.

TK_ONDEVICE_STATUS and TK_OFFDEVICE_STATUS

When a track circuit occupancy indication is received from the field, it is stored in TK_ONDEVICE_STATUS as well as in CPDEVICE_STATUS. "TK_ON" stands for “track on.” Similarly, a track circuit unoccupancy indication is stored in TK_OFFDEVICE_STATUS as well as in CPDEVICE_STATUS. The fields
CP_NAME and DEVICE_NAME do not work properly and do not contain any values. The field FORMATTED_TEXT contains the values that should appear in CP_NAME and DEVICE_NAME, along with the time from STATUS_DT in formatted as time in seconds and CURRENT_STATE. In order to use these two tables, values from a parsed FORMATTED_TEXT field must be used in place of the blank fields. Otherwise, the table CP_DEVICE_STATUS can be used.

Track circuit names are not unique identifiers by themselves. The MBTA Red Line contains duplicate TRACK_NAMES. Together with CP_NAME, all track circuits have a unique combination in the OCS system (i.e. 21028T CENTRAL SQUARE is different from 21028T ANDREW SQUARE). However, one track circuit may receive indications accompanied by more than one CP_NAME (i.e. 23148T ALEWIFE, 23148T ALEWIFE YARD and 23148T ALEWIFE CROSSOVER are all the same track circuit). All track names correspond to track circuit names on MBTA signal engineering diagrams and predate the OCS.

CP DEVICE STATUS

The table CP DEVICE STATUS contains records for the change in status of any devices in the MBTA OCS. This includes track circuits, signal indications, start bell indications, switch locking and unlocking, CTC matching, power switches, terminal dispatching mode, terminal platforms being removed from service, Absolute Stop Block (ASB) status and more. While the Red Line and most of the Orange Line has an ATC signal system with wayside signals only at interlockings, the Blue Line has wayside signals. Thus, there are many Blue Line signal indications while there are fewer Red and Orange Line ones. All devices communicate through a control point, designated in the field CP_NAME in numerous tables. The values of CP_NAME correspond to station names and yard names, as the signal room or bungalow containing the signaling and communications equipment are named.

TRAIN TRACKING

The MBTA OCS displays train IDs on the graphical views. The train ID is equal to the 4-digit front car number, plus an optional 1-character trip tag.
which can be added by a dispatcher as a reminder of problems on that train (L for light, D for door, K for early layup). The consist of each train is in the OCS system and is updated through manual input by the dispatcher (with communication to the field) and through a number of Automatic Vehicle Identification (AVI) devices along the wayside.

The MBTA’s track circuits themselves are not capable of identifying the train on top of them. This is left to a train tracking software overlay. When a track circuit occupancy occurs, the software tries to determine whether the occupancy was made by a train or not. Based on the current locations of train IDs already assigned to occupied track circuits, the software looks at adjacent track circuits to guess which train it was that moved.

When track circuits have problems and fail to send positive indications to the OCS when trains enter and exit, the train tracking algorithm may either lose the train ID or assign it to the wrong occupancy indication. This occurs relatively infrequently on the MBTA Red Line.

The TRAIN_TRACKING table contains a time stamp TTT_DT that can be extracted in seconds, although its default is to be outputted in minutes. The record contains the track circuit name in the field TRACK_NAME and control point name in the field CP_NAME. The field TTT_DT appears to be subject to a problem. For certain periods on certain days, records for a given train can come time-stamped in frequent bunches (i.e. three train tracking records in a row have the same TTT_DT value). The patterns observed are consistent with the possibility that the problem occurs when the database is busy. The field SCHED_TRIP_ID links a TRAIN_TRACKING record to a unique SCHED_TRIP record, although sometimes the field SCHED_TRIP_ID may be blank (especially for night work trips).

SCHED_TRIP

The table SCHED_TRIP contains a record for each train listed on the dispatcher’s train sheet. Its contents do not correspond to the dispatcher’s
train sheet itself. The train sheet contains a field for the dispatcher’s integer-minute offset for ringoff time. This value is not stored in the database, although it would be useful if it were. There is no field for consist length, although it appears on the train sheet. The field TRIP_CONSIST contains all vehicle numbers in each train, separated by commas. TRIP_CONSIST is subject to human error.

The field TRIP_ID and NEXT_TRIP_ID provide a link between one trip in SCHED_TRIP and the next trip of that train. ROUTE_NAME provides the MBTA standard route code (i.e. S931_ means southbound Ashmont, while RAD means “run-as-directed”). The fields ORIG_STATION, DEST_STATION and TRANSIT_LINE_NAME are self-explanatory. The fields SCHED_DEPART_DT and SCHED_ARRIVE_DT are based on a schedule inputted quarterly.

The fields ACTUAL_DEPART_DT and ACTUAL_ARRIVE_DT are in minutes rather than seconds, and are equal to the time the front of the train passes over the track circuit boundary immediately beyond the platform. When short-turns occur, the dispatchers should reassign the trip to the other direction using the train sheet. If they do not, the values in ACTUAL_DEPART_DT may have problems.

SCHED_DETAIL

The table SCHED_DETAIL contains the time of arrival and departure at each station. The arrival time is defined as the time the front of the train passes into the track circuit in which the front of the train typically stops. It cannot be used as an accurate approximation for dwell time.

SCHED_CREW

The table SCHED_CREW contains the nominally scheduled crew information that is decided at the beginning of each pick. After the pick it is not updated. It does not represent the crews actually on the train. It can be used by dispatchers to determine which trips are scheduled to have either a motorperson or a train attendant swing off and at what location. A crew
schedule interface which was originally designed with the OCS has not been used by the MBTA’s train starters in particular because of it was not adaptable to their needs.

Crew assignment logs

The train starters and terminal officials keep manual records of which crews are aboard all trains. For the MBTA Red Line these sheets are 11”x17” and are entitled “Train Sheet of Crews and Trains.” They contain many columns, including train consist, scheduled arrival and departure times for terminals and JFK / UMass, scheduled crew IDs for motorpersons, train attendants, and IDs for substitute crews where there are swingoffs. There are columns for actual arrival and departure times, actual crew IDs, and for comments. The comments record unscheduled mid-route crew-swapping, swapping of vehicles, cutting and adding of trains and other information such as incident remarks.

Dispatchers’ logs

Dispatchers logs are, at this writing, maintained on an older system than the OCS, but are available on OCS workstations through a terminal emulation program. The logs are stored on an IBM COBOL system and can be extracted to ASCII files. They contain numerous fields, including a field for dispatchers’ comments relating to an incident record.
<table>
<thead>
<tr>
<th>CP DEVICE STATUS</th>
<th>SCHED TRIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP_DEVICE_STATUS_ID numeric(10) identity</td>
<td>SCHED_TRIP_ID numeric(10) identity</td>
</tr>
<tr>
<td>CP_NAME varchar(40) null</td>
<td>TRIP_ID numeric(10) null</td>
</tr>
<tr>
<td>DEVICE_TYPE varchar(20) null</td>
<td>ROUTE_NAME varchar(10) null</td>
</tr>
<tr>
<td>DEVICE_NAME varchar(20) null</td>
<td>TRANSIT_LINE_NAME varchar(20) null</td>
</tr>
<tr>
<td>STATUS_DT datetime null</td>
<td>ORIG_STATION varchar(20) null</td>
</tr>
<tr>
<td>CURRENT_STATE varchar(80) null</td>
<td>DEST_STATION varchar(20) null</td>
</tr>
<tr>
<td>EQ_VALUE varchar(30) null</td>
<td>SCHED_DEPART_DT datetime null</td>
</tr>
<tr>
<td>FORMATTED_TEXT varchar(255) null</td>
<td>ACTUAL_DEPART_DT datetime null</td>
</tr>
<tr>
<td>TRANSIT_DAY smalldatetime null</td>
<td>SCHED_ARRIVE_DT datetime null</td>
</tr>
<tr>
<td>DAY_OF_YEAR smallint null</td>
<td>ACTUAL_ARRIVE_DT datetime null</td>
</tr>
<tr>
<td>CREATE_DT datetime null</td>
<td>EVENT_NUMBER int null</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TK OFF DEVICE STATUS</th>
<th>SCHED_DETAIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK_OFF_DEVICE_STATUS numeric(10) identity</td>
<td>SCHED_DETAIL_ID numeric(10) identity</td>
</tr>
<tr>
<td>CP_NAME varchar(40) null</td>
<td>SCHED_TRIP_ID numeric(10) null</td>
</tr>
<tr>
<td>DEVICE_TYPE varchar(20) null</td>
<td>TRIP_ID numeric(10) not null</td>
</tr>
<tr>
<td>DEVICE_NAME varchar(20) null</td>
<td>STATION_NAME varchar(20) null</td>
</tr>
<tr>
<td>STATUS_DT datetime null</td>
<td>ACTUAL_ArrIVAL_DT datetime null</td>
</tr>
<tr>
<td>CURRENT_STATE varchar(80) null</td>
<td>SCHED_Arrival_DT datetime null</td>
</tr>
<tr>
<td>EQ_VALUE varchar(30) null</td>
<td>ACTUAL_ArrIVAL_DT datetime null</td>
</tr>
<tr>
<td>FORMATTED_TEXT varchar(255) null</td>
<td>VALUE varchar(20) null</td>
</tr>
<tr>
<td>TRANSIT_DAY smalldatetime null</td>
<td>DAY_OF_YEAR smallint null</td>
</tr>
<tr>
<td>DAY_OF_YEAR smallint null</td>
<td>CREATE_DT datetime null</td>
</tr>
<tr>
<td>CREATE_DT datetime null</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TK ON DEVICE STATUS</th>
<th>SCHED_CREW</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK_ON_DEVICE_STATUS numeric(10) identity</td>
<td>SCHED_CREW_ID numeric(10) identity</td>
</tr>
<tr>
<td>CP_NAME varchar(40) null</td>
<td>SCHED_TRIP_ID numeric(10) null</td>
</tr>
<tr>
<td>DEVICE_TYPE varchar(20) null</td>
<td>TRIP_ID numeric(10) not null</td>
</tr>
<tr>
<td>DEVICE_NAME varchar(20) null</td>
<td>ST_MOTOR varchar(5) null</td>
</tr>
<tr>
<td>STATUS_DT datetime null</td>
<td>ST_G1 varchar(5) null</td>
</tr>
<tr>
<td>CURRENT_STATE varchar(80) null</td>
<td>ST_G2 varchar(5) null</td>
</tr>
<tr>
<td>EQ_VALUE varchar(30) null</td>
<td>END_MOTOR varchar(5) null</td>
</tr>
<tr>
<td>FORMATTED_TEXT varchar(255) null</td>
<td>END_G1 varchar(5) null</td>
</tr>
<tr>
<td>TRANSIT_DAY smalldatetime null</td>
<td>RELIEF_MOTOR varchar(20) null</td>
</tr>
<tr>
<td>DAY_OF_YEAR smallint null</td>
<td>RELIEF_G1 varchar(20) null</td>
</tr>
<tr>
<td>CREATE_DT datetime null</td>
<td>RELIEF_G2 varchar(20) null</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TREG STATION HEADWAY</th>
<th>TRAIN TRACKING</th>
</tr>
</thead>
<tbody>
<tr>
<td>TREG_STATION_HEADWAY numeric(10) identity</td>
<td>TRAIN_TRACKING_ID numeric(10) identity</td>
</tr>
<tr>
<td>TRANSIT_LINE_NAME varchar(20) null</td>
<td>SCHED_TRIP_ID numeric(10) null</td>
</tr>
<tr>
<td>STATION_NAME varchar(20) null</td>
<td>TRANSIT_LINE_NAME varchar(20) null</td>
</tr>
<tr>
<td>DIRECTION varchar(10) null</td>
<td>CP_NAME varchar(40) null</td>
</tr>
<tr>
<td>VARIANCE varchar(6) null</td>
<td>VEHICLE_ID varchar(40) null</td>
</tr>
<tr>
<td>TRAIN_ID varchar(6) null</td>
<td>TRACK_NAME varchar(20) null</td>
</tr>
<tr>
<td>ROUTE_NAME varchar(10) null</td>
<td>TTT_DT datetime not null</td>
</tr>
<tr>
<td>TRIP_ID numeric(10) not null</td>
<td>DAY_OF_YEAR smallint null</td>
</tr>
<tr>
<td>ACTUAL_ArrIVAL_DT datetime null</td>
<td>CREATE_DT datetime null</td>
</tr>
<tr>
<td>SCHED_ArrIVAL_DT datetime null</td>
<td></td>
</tr>
<tr>
<td>ACTUAL_ArrIVAL_DT datetime null</td>
<td></td>
</tr>
<tr>
<td>SCHED_ArrIVAL_DT datetime null</td>
<td></td>
</tr>
</tbody>
</table>

Figure A-1
OCS Database Tables Used
Figures B-1 and B-2 show dwell times at Park Street southbound and Downtown Crossing southbound respectively. Dwell times in this case are approximate, but slightly larger than the time spent with the doors actually open. The unoccupancy time of track circuit PARK STREET (R) 7-2T was used as the arrival time at Park Street, while the departure time from Park Street was taken as any train-ID-move into track circuit PARK STREET (R) 225T. Similarly for Downtown Crossing, the unoccupancy time of track circuit DOWNTOWN CROSSING (R) 225T was used as the arrival time, while the departure time was taken as any train-ID-move into track circuit DOWNTOWN CROSSING (R) 2121T.

Figure B-3 shows estimated station close-in times for a typical day. These were derived from train tracking data alone, although a more accurate estimate could be made by using track circuit unoccupancy data.

Figure B-4 shows passenger load profiles corresponding to Figures 4-10 (a) and (b) in the text.
Figure B-1
Dwell Times at Park Street Southbound
January 22, 2001
Figure B-2
Dwell Times at Downtown Crossing Southbound
January 22, 2001
MBTA Red Line Southbound
Volume Leaving Stations
7:00 AM - 9:00 AM
CTPS 1997 Passenger Counts

Max. Load = 10903

MBTA Red Line Northbound
Volume Leaving Stations
3:30 PM - 6:00 PM
CTPS 1997 Passenger Counts

Max. Load = 12433
Figure B-6: Scheduled Running Time, 95th Percentile Running Time, Scheduled Half-Cycle Time, 85th Percentile Running Time, Average Running Time, Typical Running Time (Mode), Median Running Time.
Running Time Alewife-Ashmont
Weekday Schedule
Winter 2001

Figure B-7

Actual Departure Time at Alewife

Travel Time, (min.)

5AM 6AM 7AM 8AM 9AM 10AM 11AM 12PM 1PM 2PM 3PM 4PM 5PM 6PM 7PM 8PM 9PM 10PM 11PM 12AM

- Scheduled Running Time
- 95th Percentile Running Time
- Scheduled Half-Cycle Time
- 85th Percentile Running Time
- Average Running Time
- Typical Running Time (Mode)
- Median Running Time

Figure B-7
Running Time Ashmont-Alewife
Weekday Schedule
Winter 2001

Figure B-8
References


Chicago Transit Authority, (1999). Special Events Rail Ridership, Summer 1999 Summary. CTA Planning and Development, Data Services and Development


De Coen, L., Netter, B.D., and Vingerhoeds, R.A. Central advice system for fleet management and operations. Second World Congress on Safety of Transportation: Imbalance between Growth and Safety?


Fong, D.J. (1981). Evaluation of Schedule-Based Holding for Transit Vehicles. MS Thesis, Department of Civil Engineering, MIT.


134


Weiss, D.M. New York City Transit's New Rail Control Center.


