

Public Transp (2013) 5:149–176
DOI 10.1007/s12469-013-0065-9

ORIGINAL PAPER

A framework for evaluating operations control on a metro line: integrating multiple perspectives and automatically collected train and passenger movement data

Andre Carrel · Rabi G. Mishalani ·
Nigel H.M. Wilson · John P. Attanucci

Published online: 5 July 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Transit operations control, the task of implementing the operations plan in daily operations on a metro line, plays a key role in service delivery because it determines the quality of the service experienced by passengers. Yet, it is one of the most poorly understood aspects of rail transit operations. Faced with a disruption or infeasibility, dispatchers typically choose between several response strategies. However, to date, it has been very difficult to evaluate the positive and negative effects of individual control strategies with respect to operations and passenger travel times under real-world conditions. This paper proposes a framework for the study of rail operations control decisions that integrates automatically collected service and passenger demand data, which are increasingly available and accessible to transit agencies. The framework supports a multiperspective analysis methodology that can inform operational policies and plans, and help operations control decision-makers choose the most appropriate strategies to manage service. By using automatically collected data, taking into consideration the operations control decision environment, and acknowledging that the reliability of the resulting service depends on many factors endogenous to it, this paper takes a distinctly different approach from previous studies, which have relied heavily on modeling, assumed simple operating contexts, and did not consider the full range of available data. Two real-world applications of

The research presented in this paper was carried out when the corresponding author was a student at the Massachusetts Institute of Technology.

A. Carrel (✉)

University of California at Berkeley, 109 McLaughlin Hall, Berkeley, CA 94720-1720, USA
e-mail: acarrel@alum.mit.edu

R.G. Mishalani

The Ohio State University, 2070 Neil Avenue, Columbus, OH 43210-1226, USA

N.H.M. Wilson · J.P. Attanucci

Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4301, USA

the framework, where control decisions are evaluated in terms of their operational and passenger impacts, are presented. The methodology is found to be versatile and valuable in providing insights that could not have been gained otherwise. Although the framework is applied to the London Underground, its logic, structure, and procedures are applicable and transferable to other metro systems recognizing that certain specifics would need to be tailored to the available data.

Keywords Rail transit · Operations control · Disruption recovery · Transit reliability · Automatic data collection systems · Total passenger travel time

1 Introduction and motivation

On any rail transit line, the basis for the service provided to customers is the service plan, which defines routes, service frequencies, and the span of service. The service plan provides the input to the development of the operations plan, which describes the manner in which the agency's resources—rolling stock, personnel and infrastructure—are to be utilized in order to fulfill the service plan. The operations plan generally consists of a working timetable (including scheduled non-revenue train movements), a crew schedule, a vehicle assignment and a crew roster collectively representing a plan aimed at achieving the best possible level of service for passengers subject to constraints on agency resources.

The implementation of the operations plan in everyday operations is the task of the transit agency's operations controllers and involves all front-line staff (e.g., train operators and station supervisors) as well as vehicle and infrastructure maintenance divisions, engineers and operational support personnel. It is often impossible to operate the service exactly according to the operations plan due to unforeseen events (e.g., disruptions), infeasibilities (e.g., arising from high demand leading to congestion), and unavailability of resources (e.g., absent crew and unavailable equipment). An important part of the operations controllers' tasks is to modify the operations plan, including the train schedule reflected in the timetable, in real-time to account for such situations. This process is commonly referred to as operations control, or synonymously as service control or operations management. Therefore, operations control is the central, real-time control of scheduled operations that directly results in the daily operations provided to passengers. Its centralized nature, in which controllers are provided with information on the state of the entire system, sets it apart from local dispatching techniques, where a supervisor is positioned at a terminal or a station along the line to monitor local conditions and issue local instructions.

Operations control directly governs the interaction between the supply of transportation capacity and the demand for it on a daily basis. The activities of the operations control center of a metro line or network capture the integration of the network's central communication, decision-making, and coordination functions. The operations control process, as described in Froloff et al. (1989), consists of a transit line (or network) in a state of optimal operations at the beginning of the service. Due to disruptions or delays, it then deteriorates to a disrupted state. If nothing were done, operations would further deteriorate. To prevent this deterioration and to restore good

operations, operations controllers take corrective actions, known as “interventions”, toward a target state. Example interventions include holding of trains in stations, canceling individual train runs, diverting trains or unplanned crew changes. Further train and crew related interventions are discussed in Sect. 2 and definitions including detailed descriptions are provided in Carrel (2009). These interventions extend over time and the system only gradually moves toward its target state. During this process, further unforeseen events may occur, prompting controllers to intervene again. This process is continuous over time, taking place from the beginning of the service day until the end. The set of interventions taken to restore service generally follows a control strategy, which is either explicitly defined by agency policies or is based on the operations controllers’ experience, knowledge of the system, and rules of thumb.

During the evolution of this process, operations controllers work with three streams of input. First, controllers receive realtime information on the current operating state of the line (or network) including the positions of trains from the operations control system (OCS). Depending on the design of the OCS, additional information may include the scheduled train run a train is covering, the train’s destination, the train’s lateness with respect to the schedule, the train’s loading, and the driver’s ID. Secondly, a controller has information on the target state of the system. With the exception of operations under severely disrupted conditions, the target state is commonly the state the operations plan aims to achieve at a given time-of-day. Thirdly, operations controllers are in communication with drivers, rolling stock and crew managers, and other field personnel and could receive field information directly from them.

While operations planners of a metro line provide essential inputs to operations controllers, in turn they need to have a good understanding of the operations on the line in order to be able to make informed decisions on service policies, validate assumptions made during the operations planning process, and identify possible changes to the operations plan that could have a positive effect on the functioning of the line. Moreover, as Rahbee (2001) points out, operations control techniques are often passed down by word-of-mouth to new controllers; therefore, a good understanding of operations is essential for improving and formalizing operations controller training. However, operations control remains heavily reliant on human judgment and often informal and seldom documented practices mostly driven by anecdotal evidence without data to back them up. Feedback from passengers is sometimes consulted. In addition, mathematical models have played some role, albeit limited, in supporting the development of such an understanding.

In recent years the availability of train and passenger data from automated data collection systems has markedly increased, making a wealth of information on the functioning of a transit line available. Such data could support empirical evidence-driven modifications or improvements to the understanding of the complexities of line operations. In this paper a framework is proposed to empirically evaluate line operations based on such data for the purpose of either identifying limitations and suggesting alternative approaches or validating the effectiveness of certain operations control strategies.

Section 2 of this paper summarizes pertinent literature and discusses the focus and value of the study with respect to the literature. The proposed framework and

the description of its elements are presented in Sect. 3. Section 4 presents two real-world applications of the framework. The final section provides a summary and a concluding discussion.

2 Literature and focus of study

There is a wide range of literature on operations control, much of which is concerned with developing optimal control strategies. This paper does not focus on prescribing or improving specific control strategies, but rather seeks to provide practitioners and researchers alike with a framework and corresponding tools that enable the evaluation of field implemented strategies. This study focuses specifically on operations control interventions that utilize the resources (equipment, crew, and infrastructure) of the metro line under consideration and that are therefore under the control of the respective operations control center. The literature presented in this section concentrates on studies that investigate these types of interventions. Of course, future research on the proposed framework could be extended to include control strategies that make use of complimentary systems such as buses to bridge inoperable metro line segments, as presented, for example, in Kepaptsoglou and Karlaftis (2009) and Jin et al. (2013).

Although research on operations control strategies and dispatching dates back to at least the early seventies, studies on operations control considering all the interventions operations controllers have at their disposal and commonly resort to are limited. Example train related interventions include changing of schedule, holding, early dispatching, short-turning, withdrawing, diverting, extending, changing track assignments, expressing, and adding service. Example crew related interventions include substituting drivers, dropping back and jumping up in case of breaks during layovers, and switching drivers. As mentioned above, Carrel (2009) provides definitions including detailed descriptions of train and crew related interventions. The exact set of interventions available to an operations controller at a given time and location is a function of the layout of the network, the nature of the particular service disruption, the design of the signaling and train control systems, and the availability of spare rolling stock and crew. Furthermore, depending on the degree of system automation, the operations controller may implement changes remotely through the signaling system, the automated train control system, or by directly communicating with train drivers. Many previous research efforts dealt with simple line layouts and a restricted set of control interventions. This limitation may not be surprising given the complexity of the problem and the large variety of service patterns and intervention strategies that a controller typically has to consider. Moreover, many studies where mathematical formulations intended to produce optimal control strategies on a transit route or line are developed focus on passenger in-vehicle travel time or waiting time and headway regularity, which may limit their adoption in real-world situations.

Much of the early research involving mathematical formulations assumed that the dispatcher had little or no real-time information on the position of vehicles along the line. Those studies generally examine local dispatching or holding strategies at predetermined control points on the line (e.g., terminals or time points), taking as input the timetable and possibly the distance between a vehicle and its immediate

neighboring vehicles, but no real-time information on the state of the line as a whole. The most notable examples of such studies are those by Osuna and Newell (1972), Koffman (1978), and Abkowitz (1986).

The emergence of real-time sensing systems providing data from all vehicles serving a transit line has enabled recent research to take a broader approach to the holding and dispatching problem and to consider, in addition to holding, other types of operations control interventions such as short-turning, expressing, and deadheading. This development creates a strong tendency to use optimization models, calibrated with archived vehicle location data, that seek to find optimal control strategies under disrupted conditions. However, these studies do not simultaneously consider the wide range of interventions and the complexity of the decision environment operations controllers work within.

A series of related studies of this type are those of Eberlein (1995), Eberlein et al. (2001), O'Dell (1997), Shen (2000), and Puong (2001). Eberlein (1995) and Eberlein et al. (2001) considered the holding problem in the presence of real-time position data on multiple trains on a route. O'Dell (1997) considered holding and short-turning as feasible strategies for responding to a disruption in real-time. Shen (2000) extended O'Dell's formulation to include expressing. Puong (2001) built on the studies of O'Dell and Shen, to account for terminal capacity constraints. The objectives reflected in the formulations developed in these studies are to minimize in-vehicle or on-platform delays to passengers assuming centralized operations control. Moreover, in all cases the demand and interstation train travel times are assumed deterministic and known and the duration of the disruption is assumed known. Other example studies include those of Adamski and Turnau (1998) and Walker et al. (2005). Adamski and Turnau (1998) focused on dispatching while Walker et al. (2005) introduced a formulation capturing the simultaneous recovery of a train timetable and of the crew roster after a single disruption on a simple rail line, but without considering passenger delays. More recently, Daganzo (2009) and Daganzo and Pilachowski (2011) developed vehicle holding and speed control strategies, respectively, based on real-time headway information taking into account uncertain demand and vehicle travel times.

Regarding the evaluation of developed strategies, Eberlein (1995), Eberlein et al. (2001), O'Dell (1997), Shen (2000) and Puong (2001) compared the performance of the strategies produced by their formulations with a do-nothing scenario to demonstrate the effectiveness of their proposed strategies. However, it is unrealistic to assume that operations controllers would always "do nothing" (unless there is a good reason to do so). While the passenger travel time savings with regard to a do-nothing scenario may be useful for comparison across formulations, the important question is whether any of the formulations are able to provide an improvement over operations control on the line as it exists in practice. Answering this question obviously requires an analysis of operational and demand data and a reconstruction of the service on a daily basis as a benchmark for performance evaluation. Daganzo (2009) and Daganzo and Pilachowski (2011) analyzed the performance of the resulting control strategies from a reliability perspective, both analytically and numerically. However, as in the other studies referred to above, they did not consider in-service data collected on an operational system.

With the increasing availability of automatically collected operations data, several recent studies focused on the role and use of such data for both operations recon-

struction and analysis purposes. For example, Wile (2003), Dixon (2006), and Ji et al. (2010) specifically focused on the use of automatically collected data for the analysis of operations and presented examples based on data collected off an operational system. Wile (2003) explores how transit agencies can use automatically collected operational data for a variety of applications—including operations control, passenger information and safety—and establishes a range of agency functional needs that can benefit from such data. However, the study does not offer an integrated methodology for operational analysis taking advantage of multiple sources of operations and demand data. Dixon (2006) presents a tool to automatically reconstruct daily operations on heavy rail lines of the Massachusetts Bay Transportation Authority (MBTA) from data produced by the OCS. This tool inspired the more detailed and comprehensive reconstruction methodology developed as part of the framework described in this paper, however, it is limited in scope in that it does not identify operations control interventions and link the reconstructed operations to those decisions. Ji et al. (2010) evaluated the effects of holding and stop-to-stop travel time adjustments on service reliability using a large data set of automatically collected vehicle trajectories on an operational bus route. However, while expected passenger demand patterns are considered in interpreting the results, demand data were not used in the study. In addition to the above studies, the use of operational data is relied upon in a broader framework developed by Rahbee (2001), focused on the analysis of operational problems on rail transit lines.

While the studies by Rahbee (2001), Wile (2003), Dixon (2006) and Ji et al. (2010) present insightful and valuable findings in specific areas of operations reconstruction and analysis using automatically collected data, none of these studies considers the full range of data currently available to transit agencies. Furthermore, although several of these studies mention the potential uses of aggregate demand data from Automatic Passenger Counter (APC) systems, these authors did not systematically incorporate the analysis of passenger movement data in their operational analyses. Doing so is particularly feasible in the presence of disaggregate passenger flow data (e.g., from Automated Fare Collection (AFC) smartcard data) that allow inferences on the time a passenger spent inside the system. These data can be used to estimate demand at a fine level of granularity and infer individual passengers' experiences from the time the service is accessed to the time the traveler leaves the system. Thus, the aforementioned authors do not emphasize the opportunities and challenges related to the integration of highly detailed data from multiple sources relating to both operations and demand for operational analysis.

Finally regarding the literature, Froloff et al. (1989) identify and systematically categorize objectives, constraints and techniques for bus operations control aimed at the design of a simulation system for controller training. Despite the focus on bus services, many of the topics addressed form a good foundation for studying operations control on rail systems in that the study reflects the only attempt to describe the decision environment of operations controllers, i.e., the complete description of variables, objectives and constraints that affect operations controller decisions, thus setting this study apart from virtually all other research on operations control strategies. However, for research to be applicable in a real-world context, it should not only recognize the multi-faceted decision environment of the operations controller

but also specifically recognize the interactions that exist and the possible conflicts that can arise among them. This issue is treated extensively by Carrel (2009) and Carrel et al. (2010), where a detailed description of the rail service controllers' decision environment is provided and the interactions between different objectives and constraints are illustrated. To the extent that they are pertinent in the context of this paper, the findings are summarized in Sect. 3.3.

This study is therefore motivated by the need for a comprehensive framework and accompanying methodologies to analyze operations control questions, especially one that takes into account both the multiplicity of operations control interventions that could be applied in an operational setting and the increasing availability of train and passenger movement data. The key insight is that any effort to analyze and build a good understanding of operations on a line must include operations control, and therefore, the proposed framework would have to allow for considering the full range of interventions available to controllers. In this paper, it is argued that virtually any transit agency with centralized operations control would stand to benefit from taking this perspective. The framework incorporates multiple dimensions relating to operations control, namely control, operations, and demand. This is made possible by the multiple sources of both operations and demand data that are considered and integrated in a consistent manner. Notable regarding the capture of passenger experience is the consideration of the observed total travel time from the moment passengers enter stations to the moment they leave. This is in contrast to the optimization formulations discussed above, in which one or more journey components (typically waiting time or in-vehicle time) were analytically derived.

Another key distinguishing aspect of this study is the application of the proposed framework and accompanying methodologies using data from the London Underground to analyze specific service disruption cases. In doing so, the feasibility of the framework is demonstrated and its value is shown regarding both identifying operations control limitations in one case and validating the effectiveness of an operations control strategy in another. More specifically, the framework has improved the ability of London Underground analysts and engineers to infer, analyze, and evaluate the impact of interventions on the operational level of service and on passengers. Furthermore, it is hoped that it will eventually aid the training of operations controllers in implementable and effective control methods that are not only focused on operational level of service metrics but also on passenger experiences. Although the framework is applied to the London Underground's Central line, and as such some of its specifics are tailored to the nature of the available data on this line, the logic, structure, and procedures of the framework are applicable and transferable to other metro systems recognizing that certain specifics would need to be tailored to the available data. It is imperative to emphasize though that while the framework—once tailored—can be applied in a large range of unique situations, improved or new operations control strategies aimed at remedying problems identified through the proposed evaluation framework will need to be prescribed through subsequent processes and analyses carried out under the auspices of the operating agency. That is, the development of such strategies is beyond the scope of this study.

3 Developed framework

In this section the developed framework and its elements are presented and discussed. Since the two applications of this framework presented in Sect. 4 relate to the Central Line, this line is referred to at appropriate points in this section as well to provide a realistic context to the concepts and methods behind the various elements of the developed framework.

3.1 Overall structure

Before the emergence of automated vehicle and passenger data collection systems, the rail transit sector was by no means a data-rich environment, prompting operations planners and analysts to rely on models for understanding operations and the effect of potential changes. However, recently, the availability, accessibility and quality of operational and passenger data have greatly improved; where possible, the analysis of these data can often provide a considerable amount of information and answers to planning questions and can serve as a sound basis for the development of predictive models. Moreover, the limitations of previous research studies on operations control motivate the need for an approach that integrates multiple perspectives and data sources in order to gain a better understanding of operations control (and therefore of operations).

To achieve this integration, the following elements need to be considered: operations control decisions (i.e., interventions), the reasons for which the decisions are made, the operational level of service on the affected line segments, and passenger travel times. These elements and their relationships are shown in Fig. 1. Collectively, these elements represent the overall structure of the framework. Furthermore, the two types of factors that lead to operations control intervention decisions are also shown, namely disruptions and general operating and demand conditions. If data on these factors are available, they can be included in the analysis.

The interventions represent deviations from the operations plan, and measuring the resulting operational level of service allows analysts to quantify the impact of those strategies on the service. However, this is only an intermediate step since ultimately the concern should be with the impact on passengers. Once the operational level of service has been quantified, the link can be made to passenger travel times by making certain assumptions and applying models linking operating conditions to passenger experiences. However, in the presence of automatically collected passenger data, especially data that allow inferences on the total travel times of passengers between their entry into and their exit from the system, it is now also possible to directly infer the impact of operations control strategies on passengers, as shown in Fig. 1 with a dashed arrow. This direct linkage is one distinguishing aspect of the developed framework, and in the applications presented in Sect. 4 is seen to be of great importance in service evaluations carried out using London Underground data.

Understanding the reasons for the decisions requires a more qualitative approach. As previously mentioned, the decision to perform an operations control intervention is informed by the controllers' knowledge of operating and demand conditions as well as disruptions to the service. It is important to note that a disruption is not a direct

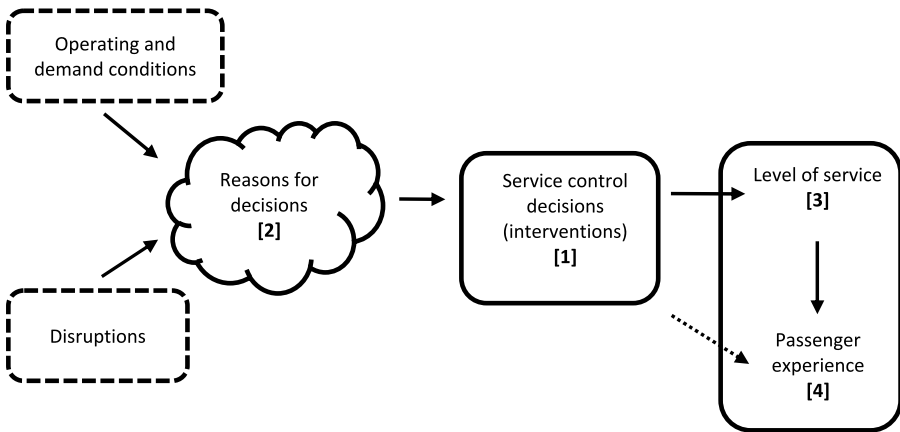


Fig. 1 A framework for studying operations control on a metro line

cause for an intervention, but is rather the root cause of a deviation of the service from the intended service or the cause of a conflict between trains. In either case, the disruption in turn leads a controller to decide whether and how to intervene to restore service quality or to address the conflict. In this sense, since a controller may determine that a small disruption does not warrant an intervention, the decision not to make any changes in the face of a disruption is also a possible outcome of operations control.

This framework focuses on the reasons for intervention decisions, the identification of operations control interventions, and the assessment of their impact. Disruptions and the characterization of operating and demand conditions are only considered as necessary to understand the root causes for intervention. In the remainder of this section the individual elements of the framework are discussed in more detail, and in Sect. 4 two applications are presented showing how the elements are integrated to support the analysis of reconstructed operations control strategies and their impact on operational service and passenger experience.

3.2 Reconstruction of operations and diagnosis of operations control interventions

The first step in the analysis of an operational question is the reconstruction of operations within the time frame of interest and the identification of operations control interventions. In Fig. 1, this element is indicated with [1]. Since the availability and quality of OCS data can vary between lines and across systems, it is not possible to establish a unique methodology for all types of OCS data. A high level of detail is not essential for this task, although the more detailed and the higher quality the data, the easier it is to deliver an in-depth and precise description of what occurred on a line. While the reconstruction methodology presented in this paper is specific to the context of and the level of detail of the data available for the London Underground line applications presented in Sect. 4 (these data are typically only available from modern OCS), the underlying concepts are transferable and the methodology can serve as a

guideline for other metro systems that could be adapted to varying contexts and levels of detail in the available data.

The basic information required as input to the reconstruction methodology is the following:

- Actual times at which a train passed either all stations or a set of predefined points along the line. These locations can be time-points, critical locations such as junctions and reversing facilities or the line terminals.
- Scheduled trajectories train units are operating on, typically in the form of numbers that link the train units to the vehicle blocks they are assigned to.
- Numbers that clearly identify physical train units. Ideally, these would be car numbers if the OCS records them. If trains are assigned to vehicle blocks and retain their block number throughout the day, then these numbers can be used as identifiers. If, however, controllers can reassign trains to different vehicle blocks during service or if trains are withdrawn to depots between the AM and the PM peak and there is no continuity in the block numbers assigned to them, then car numbers are essential.

This information allows the reconstruction of train movements throughout a day at a basic level, for example with a space-time plot. It also allows one to make the link between planned operations and the service that was actually operated by comparing the operations plan with the recorded train movements. The next step is the identification of operations control interventions.

Preferably, there should be a direct way of matching the scheduled with the operated service, as shown in the hypothetical example in Table 1. To facilitate the understanding of Table 1, the station names and associated routings refer to the London Underground Central line, the line this methodology is applied to in Sect. 4 (a line diagram is shown in Fig. 2). The availability of such matches depends on the features of the OCS, and the case where they are not readily available is outside the scope of this paper—see Ji et al. (2009) for a discussion and a possible solution to this matching problem in this case. Each row of Table 1 shows one observed train run, i.e., the movement of a train between two reversals. It features information on the origin and the destination of the train and on the observed departure and arrival times. The right-hand side of Table 1 shows the timetable run that was covered by this train and the scheduled origin, destination, departure and arrival time according to the operations plan.

Any change to a train trajectory is necessarily the decision of an operations controller, and by identifying discrepancies between scheduled and operated services, the analyst can infer and identify the actions of operations control. Even though this table takes into account only the limited amount of information outlined above, it allows the identification of short-turns, diversions, extensions, cancellations and unscheduled trips. Schedule adherence at terminal departures and arrivals can also be inferred through the comparison of the “observed” and “scheduled” parts of Table 1, and under certain conditions (depending on the information reported by the OCS) one can identify when a controller decided to dispatch a train early.

With more detailed information from the OCS, a table of this type can be expanded to include intermediate stations, the vehicle block number a train was assigned to at

Table 1 Juxtaposition of observed and scheduled service

Observed Train unit	Observed			Veh. block	Scheduled			Control actions		
	Origin	Departure time	Destination		Arrival time	Origin	Departure time		Destination	Arrival time
501	Epping	10:02	West Ruislip	11:26	1	Epping	10:00	West Ruislip	11:25	Run as scheduled
229	Epping	10:10	White City	11:12	2	Epping	10:10	West Ruislip	11:35	Short-turned
947	Epping	10:42	West Ruislip	12:08	4	Epping	10:30	West Ruislip	11:55	Canceled
238	Epping	10:49	Ealing Broadway	11:56	5	Epping	10:40	West Ruislip	12:05	Run as scheduled
157	Epping	10:55	White City	11:57	6	Epping	10:50	West Ruislip	12:15	Diverted
										Unscheduled trip

every station, platform numbers, and crew information, leading to a nearly complete reconstruction of the trajectories of each train. This additional information allows the identification of other operations control interventions, namely expressing trains, withdrawals en route, changes in platform and track assignments, the reassignment of trains to different vehicle blocks and unscheduled crew changes (e.g., the use of spare drivers).

If the quality of the OCS data is sufficient, both the reconstruction of operations and the detection of operations control interventions can be automated, as described by Carrel (2009). Such automation has several advantages over relying on manual logs completed by operations controllers. The latter will generally be filled out only when controllers have time. During the most stressful moments in control centers, e.g. during major disruptions, controllers are likely to be consumed by their operations management tasks, to the detriment of the reliability of the manual logs. Aside from that, controllers may record poor operations management decisions incompletely or not at all.

3.3 The operations controllers' decision environment

The next step to understanding and evaluating operations on a line, indicated with [2] in Fig. 1 and an important element of the framework, is to identify the reasons for which operations controllers performed the observed interventions. This step is rather unstructured in that it relies on available manual logs and conversations with controllers and other front-line staff. As is shown in Fig. 1, any decision made by controllers needs to be put in the context of the operating and demand conditions and disruptions that led up to it. The availability of such information varies between agencies, but is often available in the form of manual disruption logs, weather logs, and AFC data.

Given the complex interactions among the set of objectives, constraints, and provisions by the operations plan that a controller has to consider when making a decision, a recognition and understanding of these factors and their interactions are necessary to understand operations control interventions. Carrel (2009) and Carrel et al. (2010) present a comprehensive approach to acquiring this knowledge where decision factors and constraints in operations control were observed during extended visits to a control center of the London Underground and the overall decision environment is described as completely as possible. In the remainder of this section, only the aspects pertinent to the subsequent discussion and applications presented in this paper are briefly pointed out.

Probably the most important driver of operations control is the adherence to the operations plan. A significant deviation from the train timetable or the crew schedule can trigger an operations control intervention. Adherence to the train timetable is a well-known factor, but the major importance of crew management has seldom been sufficiently recognized in previous studies. It is one of the most complex aspects of operations control, as it is generally governed by numerous rules and regulations that pose constraints on the possible recovery strategies that controllers can follow under disrupted conditions. Adherence to the operations plan is very important for the efficient management of assets (e.g., trains, infrastructure capacity, and staff), and

the presence of multiple criteria regarding, for instance, levels of service and crew management, can lead to conflicts. In the absence of an official agency policy that states priorities and gives general guidelines, controllers are likely to define their own understanding of what constitutes “good” service in terms of schedule adherence, headway regularity, or traffic patterns, but not always taking into account the impact of those assumptions on passengers.

Furthermore, it is important to recognize that uncertainty is omnipresent in operations control, as controllers are working with a stochastic system where it is difficult if not impossible to predict precisely how the system will evolve, and decisions often have to be made under time pressure. Control center observations revealed that when choosing between recovery strategies, controllers would often favor those where the outcome was subject to the least uncertainty. Another decision factor encountered in the application presented subsequently in this paper is passenger safety. For instance, trains stuck behind a blockage should be held in stations whenever possible since that simplifies the evacuation of the trains in case the duration of the blockage turns out to be long. Overall, control center observations indicated that in dealing with the highest-priority objectives and constraints, operations controllers could sometimes actually cause gaps in the service and delays that would likely require remedies at a later stage. An example of such a situation is discussed in Sect. 4.

3.4 Measures of impact

Following the identification of operations control interventions and the reasons for which they were performed, measures are needed to describe the impact of these interventions on the operational level of service and passengers. Common performance metrics for transit lines often take an aggregate spatial and temporal view of operations, for example by tracking performance indicators for an entire line over time periods ranging from one day to several months. In contrast, the impact of operations control interventions is typically limited to specific segments of the line and to small time windows. In choosing between impact measures, the analyst should be aware of the resolution trade-off in effect. A highly disaggregate representation would support a high-quality analysis of intervention strategies but renders the analysis of operations control over longer periods of time challenging due to the associated high degree of variability from situation to situation. A more aggregate representation, on the other hand, supports a long-term assessment but may hide the effects of operations control interventions. Finding the appropriate balance for the situation at hand is critical and yet is for the most part context specific and, therefore, needs to be addressed in light of the situation at hand. Furthermore, the framework is flexible and can accommodate both simple and complex line layouts. The limiting factor for its application to more complex situations is mostly the analyst’s ability to unambiguously attribute the observed service patterns to the delay and the interventions being investigated.

The two steps of quantifying the impact of operations control decisions (as well as the potential disruptions leading up to them) on operations and on passengers are indicated with [3] and [4] in Fig. 1, respectively. Step [3] takes as input the exact time windows and line segments that are of interest, as identified in the previous steps. Step [4] relies on the same information as well as any insights into the spatial

Table 2 Example intervention matrix with effects on line segments (one per direction)

Time:	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Σ
Short-turns				8	3													1	12
Diversions				2															2
Extensions			1	1						1				1	1	1			6
Canceled trips		1	2							1	2	1	2						9
Unscheduled trips				1															1
Withdrawn trains			2															1	3
Total in this direction																			33
Time:	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Segment 1																			
Reduced service				-3	-2							-1							-1 -1
Additional service				2															1
Segment 2																			
Reduced service		-1	-2	-5	-1						-2	-2							
Additional service			2	1							1		1		1				
<i>etc. for all segments of the line</i>																			

and temporal extent of the operational effects that can be derived from step [3]. In applying the framework to a specific metro network or line, the analyst would choose the metrics most suitable for that situation. In the following subsections, the metrics that were found to be most suited for the applications of Sect. 4 are presented and discussed.

3.4.1 Operational level of service by line segment

All interventions that cause a train to be rerouted or canceled can be characterized in terms of their impact on individual segments of the line. For example, a canceled or short-turned train trip results in all of the line segments that train was scheduled on to be served by one less train. On the other hand, an extension of a trip or an unscheduled trip results in a train added to the service on some segments of the line. A train diversion has both types of impact on different line segments. While the addition or removal of train service at individual stations would be the most precise way of examining the impacts, the data become more manageable without much loss of information if they are aggregated at a segment level. The segments should be defined such that the scheduled level of service is approximately constant throughout the segment and such that within the segment there are preferably few, if any, reversing tracks that are commonly used by operations controllers.

Table 2 shows an example of how operations control interventions and the effects on different segments of the line can be represented in the form of a matrix. The top part of the table shows the number of interventions by types for each hour of the day, and in the lower part, which is separated by segment, the net effects of the interventions are shown either as added service on that segment (e.g., “1” indicates that one additional train traveled over that segment) or a reduction in service (e.g., a “-1”).

The lower part of the table shows the net effects of operations control interventions in units of added (or reduced) number of trains traveling over a segment. These are translated to the ratio of scheduled to operated service kilometers at an hourly and segment level (i.e., measured in units of train-km/h for each segment), as is shown

Table 3 Example matrix of train-km operated (in bold) and % of scheduled train kilometers (in italics) by line segment and time period

Time period		Section number				
		1	2	3	4	5
10:00–11:00	Westbound	33.0	141.2	330.0	235.7	53.2
		<i>111 %</i>	<i>98 %</i>	<i>100 %</i>	<i>94 %</i>	<i>95 %</i>
11:00–12:00		29.7	137.5	300.8	227.9	56.0
		<i>100 %</i>	<i>96 %</i>	<i>91 %</i>	<i>91 %</i>	<i>100 %</i>
12:00–13:00		29.7	141.3	341.6	262.6	58.9
		<i>100 %</i>	<i>100 %</i>	<i>96 %</i>	<i>93 %</i>	<i>100 %</i>
10:00–11:00	Eastbound	33.0	127.6	321.1	227.8	56.0
		<i>111 %</i>	<i>89 %</i>	<i>97 %</i>	<i>90 %</i>	<i>100 %</i>
11:00–12:00		29.7	155.6	320.6	244.3	53.0
		<i>100 %</i>	<i>108 %</i>	<i>97 %</i>	<i>97 %</i>	<i>94 %</i>
12:00–13:00		29.7	135.7	299.8	213.1	57.6
		<i>100 %</i>	<i>93 %</i>	<i>89 %</i>	<i>85 %</i>	<i>106 %</i>

in Table 3. This measure allows for the over- and under-provision of service to be pinpointed to individual segments of the line and time periods, thus capturing the net effect of an intervention strategy. For instance, if one train is shortturned but the one behind it is extended to cover its trip out to the branch terminal, this will show up as better service than if the first train had been short-turned without extending the following one. This distinction is important since only looking at the intervention counts may convey the impression that larger absolute numbers of interventions always result in increased service curtailments. It is worthwhile noting that the adopted service measure is train-km/h for each segment rather than the number of trains per hour (tph)—or alternatively average headway—serving a line segment. The reason for not adopting tph is because, while tph is a good measure for a specific point on the line (typically a station), for a line segment, trains may only serve portions of that segment within a given time window, and tph measures would need to be somehow averaged over multiple stations where they are not necessarily equal. The train-km/h for each segment measure avoids such averaging, can be readily calculated, and offers a clear indication of the operational level of service over space and time.

3.4.2 Passenger impact

The increasing use of smartcard based AFC technology provides transit agencies with easily usable information about passenger flows on their network. In cases where passengers are required to tap their fare cards upon their entry to and exit from the system, such as in the London Underground, AFC data allow the direct measurement of passenger time spent in the system from station entry to station exit, including access time to the platform, waiting time, in-train travel time, transfer time, egress time from the platform and possibly time spent for other activities inside the system, such as waiting for somebody. From the measurement of this total origin-destination

(OD) travel time, it is possible to draw inferences about the impact of disruptions and service changes on passengers. In systems where passengers are required to validate their fare card only when entering the metro system, the alighting locations could be inferred through various procedures (e.g. by considering transfers and the boarding stop of the return trip), as discussed, for example, in Wang et al. (2011). Given the boarding and alighting stops as well as the entry time, the alighting time and thus the travel time can be inferred from vehicle data. Alternatively, in the cases where not even “tap-on” AFC data are available, the analyst can make use of passenger flow models as long as some information about OD flows is available (e.g., from onboard surveys). In the following discussion, it is assumed that passenger data include the entry time and location (origin station), exit time and location (destination station), and that a unique card number allows the entry and exit to be linked. Such data are typically available in transit agencies with distance-based fares, whereas agencies with simpler fare structures may only require passengers to tap their card upon entering the system.

Not all operations control interventions have the same impact on passengers, and different passenger groups will experience the effects of an intervention differently. One challenging task in assessing passenger impact is the quantification of the (dis-)benefits manifested after an intervention for the following two reasons. First, in daily operations, dispatching decisions are made continuously. Some interventions are performed in order to mitigate the negative impacts of earlier interventions, so the impact of a single intervention can often not be assessed without considering the overall strategy. Second, due to the stochastic nature of the environment in which the service operates, the inherent variability of dwell times and running times and unforeseen events (e.g., additional disruptions) could quickly obfuscate the effects of an individual intervention. Therefore, in quantifying passenger impacts, the analysis should be limited to passenger delays (or travel time improvements) that can be as clearly attributed to an intervention as possible.

The impacts of operations control interventions on passengers include:

- Increased or decreased waiting time at the passenger trip origin.
- An additional train transfer and waiting time during the journey in the case of a short-turn or diversion.
- Increased in-vehicle travel time experienced by passengers on board trains that are held at stations, need to travel at reduced speed between stations or make in-tunnel stops.

Although passenger impacts differ across the various components of the travel experience, identifying the impacts on specific components with AFC data is not feasible. Therefore, this analysis is restricted to the study of a sample of passenger total travel times from the time of entry at the origin station to the time of exit at the destination station, aggregated over a certain time period and pertinent OD pairs. The resulting travel time distribution captures all passenger trips that started within that time period and are distributed over all trains that served the respective OD pair. For a more in-depth discussion, see Uniman (2009) and Uniman et al. (2010).

In the specific case of a high-frequency metro line, the temporal aggregation of passenger journeys captures passengers of multiple trains, but often only a few trains

are affected by a delay. For instance, of 12 scheduled trains per hour one or two may be short-turned. For the purpose of illustration only, if one assumes even demand, only 1/12 to 1/6 of the passengers traveling to destinations beyond the short-turning point would be affected, either because they had to alight and wait for a later train or because they experienced a gap in service of one or two additional scheduled headways. Although the average travel time during this time period is an important measure, it is clear that the average is insufficient for capturing the effects of small disruptions and operations control interventions, as it has a limited sensitivity to the (often small) subset of passengers who experience worse service. However, one should expect to see a change in the shape of the travel time distribution, as said subset of passengers experiencing worse service is pushed to contribute to the right side of the travel time distribution.

Before making the trip, a single passenger will not know whether her train is the one which is going to be short-turned or otherwise delayed; i.e., her travel time might fall anywhere within the travel time distribution. The wider the distribution, the more unreliable the service is from the perspective of the passenger. This acknowledgement calls for an additional measure (or measures) capturing the spread in the distribution and, thus, the unreliability in travel time. The Reliability Buffer Time (*RBT*) is such a measure. Various authors have discussed this type of measure, among them Furth and Muller (2006), Chan (2007), Uniman (2009), and Uniman et al. (2010). Uniman defines it as:

$$RBT = P_N - P_M \tag{1}$$

where $P_N, P_M = N^{\text{th}}$ and M^{th} percentiles of travel time, respectively. Uniman proposed $P_M = 50^{\text{th}}$ percentile (i.e., the median) as an indicator of the typical travel time a frequent traveler on the system would anticipate. A passenger wishing to reach her destination by a certain time with a probability $N/100$ must budget the travel time given by the N^{th} percentile of the cumulative travel time distribution. Thus, the *RBT* is the additional time to be budgeted over and above the typical travel time by passengers who wish to arrive at their destination by a certain time with a degree of confidence given by the probability measure $N/100$.

Building on this, the framework uses two measures to summarize the travel time distribution and thus the impact of interventions on passenger travel times: the average travel time, \bar{T} , and the *RBT*, either for every OD pair or across several OD pairs. Both the average travel time \bar{T} and the *RBT* can be expressed as totals denoted by \bar{T}_{total} and RBT_{total} , respectively across the traveling population of interest, which is helpful in putting the numbers into context. These two measures are calculated as follows:

$$\bar{T}_{total} = \sum_{\forall i} V_i \cdot \bar{T}_i \tag{2}$$

$$RBT_{total} = \sum_{\forall i} V_i \cdot RBT_i \tag{3}$$

where V_i = passenger volume for OD pair i , \bar{T}_i = average travel time for OD pair i , and $RBT_i = RBT$ for OD pair i .

The calculation of total average travel time can be performed at a line segment and time period level and represents the total time expected to be spent traveling by all passengers traveling on any given OD pair without accounting for the unreliability of the system. The total *RBT* is most meaningful when compared to the total average travel time, as it represents the total time all passengers would need to budget in addition to their typical travel time in order to arrive at their destination by their desired arrival time with probability $N/100$.

The quality of the numerical determination of the travel time distribution depends on the quality of the AFC data. While there is no reason to believe that the AFC time-stamps are error prone, any action of a passenger which is not directly related to the originally intended trip (e.g., taking the wrong train or spending additional time at a station to make a phone call or meet someone) enters into the travel time distribution despite its irrelevance to the quality of service. Those passengers are likely to be reflected in the higher percentiles of the distribution, giving a false impression of service unreliability. Unfortunately, there is currently no way of filtering out such passengers, so the analyst depends on the assumption that either any inexplicable passenger behavior results in travel times beyond the N^{th} percentile or that this subset of passengers is sufficiently small and does not severely affect the summary statistics. Chan (2007) showed that the impact of such behavior on the spread of the distribution is minimal up to a value of $N = 99$ (i.e., the 99th percentile of trips). However, in the illustrative example discussed in the next section, the minimum sample size used for analysis was 10 trips, so for several OD pairs, a choice of $N = 99$ would have required interpolation to determine the 99th percentile travel time. To increase the number of OD pairs in which the N^{th} percentile corresponded to an observed travel time, a value of $N = 90$ was adopted.

Finally, it is not straightforward to distinguish whether the delays reflected in a travel time distribution are caused by disruptions or by operations control interventions because gaps in the service can be the consequence of either of these two events. One could argue that passengers may not even care about the reason behind the gap. However, to the transit agency, the reason is of great interest. Addressing the problem of determining the reason behind a delay is highly specific to the context of the situation under examination, as can be seen in the next section.

4 Example applications

In this section, two different applications of the framework are presented. They are intended to illustrate how the various elements of the framework are integrated and how the framework can be applied in a real-world setting as well as to exhibit the value of analyzing operations control from multiple perspectives using data from multiple sources. The applications focus on the Central line, one of the longest and most congested lines of the London Underground network. A simplified line diagram is shown in Fig. 2 indicating the stations referred to in the discussion of the applications (in total the Central line serves 49 stations). Even though the Central line does not share tracks with other lines, its layout includes three branches and a loop, and trains operating on the line are generally divided into groups serving different pairs of branches



Fig. 2 Central line diagram

on either side of the trunk section, or the trunk section only. Thus, from the point of view of an operations controller, it is very similar to a set of different lines that share a common trunk section.

In the first application of the proposed framework, a potentially problematic disruption recovery strategy is identified and recommendations for improvement are made. In the second application of the proposed framework, a strategy for dealing with rolling stock shortages is validated and a few recommendations for scheduling are made. Between the two applications, the value of the proposed framework as both an evaluation and improvement identification tool and a validation tool is clearly demonstrated. Two additional applications using this framework can be found in Carrel (2009) and Carrel et al. (2010).

4.1 Analysis of a disruption management strategy

This example shows how a retrospective analysis of the way a disruption was managed by operations controllers through the application of the proposed framework can help identify problematic operations management techniques. The insights from such an analysis can, for example, be used to establish guidelines on priorities to operations controllers and to consequently improve operations controller training. They can also be used to motivate the development of improved control strategies in light of the developed in-depth understanding of the complex situations controllers have to work within and the consequences of the actions they take.

The focus of this analysis is a disruption that occurred on the Central Line on April 1, 2008. The disruption can be summarized as follows: At 08:29, eastbound train #141 experienced an in-tunnel delay between Lancaster Gate and Marble Arch station due to a track circuit failure. (A track circuit is used to communicate operational commands of the automated signaling system to the on-board computer of the train unit. In the event of a track circuit failure, trains cannot operate normally over the affected segment of the line.) The scheduled headway at the location and time of the disruption was 120 seconds. However, due to train congestion, it was not uncommon for individual headways to be observed at values as low as 90 seconds.

The disruption was cleared at 08:35, and the affected train departed Marble Arch station at 08:36. This is the only information given by the Underground's incident log, and to understand what took place in the wake of the incident and how the disruption and control actions taken affected passengers, the developed framework is applied to draw together information from multiple sources and determine the resulting operational level of service and passenger experience metrics. First, the demand and operating conditions are considered during the period of interest. Demand on the London Underground has a large peak-to-base ratio, with demand peaking between 08:00 and 09:00 and between 16:00 and 17:00. Furthermore, April 1 being a regular weekday, demand is highest on the central London segment of the line, roughly between Notting Hill Gate and Stratford. Hence, this incident occurred at one of the most congested times and locations on the line.

The goal of the analysis is to determine the amount of passenger delay caused by recovery actions and operations control interventions, and to separate those from the initial delay caused by the disruption. The first step involves the identification of the operations control interventions that occurred on the Central line eastbound at or shortly after 08:29, the start time of the disruption. The result of the analysis based on the methodology described in Sect. 3.2 is that operations controllers held all trains upstream of the disruption between Lancaster Gate and White City in stations. White City is an important crew relief location. To the east of White City, the Central line runs below ground through central London, while to the west of White City, the line is above ground. On April 1, no holding could be detected to the west of White City, as trains continued passing those stations with average-length dwell times. Furthermore, two cancellations of consecutive eastbound trains out of White City occurred, one during the disruption, at 08:33, and one shortly after the disruption was cleared, at 08:36.

The next step focuses on the reasons for the interventions. First, the holding of trains in stations in the tunnel segments is performed for safety reasons (to facilitate train evacuation if found to be desirable or necessary). Second, based on the explanations given by controllers and conversations with Transport for London engineers, the strategy of canceling trains during disruptions is common. There are several specific reasons for which these cancellations were likely made. One is to prevent a buildup of congestion between White City and North Acton, which would have persisted even after the disruption had cleared due to dwell times at White City. Another is to reduce the pool of trains and drivers that would be delayed by the incident (thus improving the manageability and reducing the workload of controllers). And, a third is to create two "spare" trains at White City which could be used later for service restoration, for example to fill the gap in front of the incident train (in the westbound service).

The two remaining elements of the framework to be discussed are the assessment of the impacts of these decisions on the operational level of service and on passenger travel times. According to the operational data, the result of the two train cancellations was that there was a gap of 7:20 min (equivalent to 3.7 scheduled headways) in the service after the disruption had cleared. Applying the operated train-kilometers to scheduled train-kilometers ratio measure proposed in Sect. 3.4, only 81 % of scheduled service (measured in train-kilometers) was operated between 08:00 and 09:00 on the eastbound trunk segment of the line, between North Acton and Leytonstone.

Since the Central line OCS continuously records train loading as a percentage of crush load, ranging from 0 % to 100 %, it could be confirmed from the data that train capacity did become a binding constraint. Specifically, train #56, the first train after that gap, departed White City at crush load (i.e., 100 % loading) since it had absorbed all passengers from the canceled trains, and it remained at crush load until Holborn, at the eastern end of the central business district. The following train, #31, had also reached crush load by Marble Arch and remained at that loading until Holborn. The removal of two trains from service during the peak hour reduced throughput capacity by approximately 2100 spaces in that short period of time, leaving very little additional capacity for downstream passengers. Presumably many passengers were unable to board the crush loaded trains, causing them to be left behind and experience longer waiting times than could have been attributed to the gap alone.

To verify this finding and assess the passenger impact of the disruption and intervention, individual passenger travel times from AFC data are analyzed following the methodology described in Sect. 3.4. A high variability in travel times is observed for different passengers who had entered the system shortly before the time when the crush loaded trains passed their station, presumably reflecting the fact that only a fraction of passengers at stations where the trains were at or near crush load were able to board, while others had to wait for the arrival of less loaded trains. The specifics of the corresponding analysis are presented below.

Only eastbound passengers on the trunk segment are considered as they were the ones unambiguously identified as affected by the disruption and the following gap. To distinguish between the impacts of the disruption itself and the intervention (canceling two trains), two time bands are defined for the analysis. Time band 1 includes the 30 minutes before the end of the disruption, including the disruption itself, while time band 2 captures the 30 minutes immediately after the disruption had cleared. Both time bands are defined to relate to passengers depending on the stations where they boarded. Passengers accessing the service at a station downstream of the disruption are attributed to the first group if they tapped in before the passage of train #141 (the originally disrupted train) at that station and to the second group if they entered the station afterwards. However, for passengers accessing the service at a station upstream of the disruption, stationary time bands are considered whereby bands 1 and 2 are delineated by the time the disruption is cleared at 08:35. Stationary bands are adopted for upstream stations because it is assumed that as soon as the disruption is cleared, the queue behind the disrupted train started dissipating.

To get a sense of the magnitude of passenger travel time changes due to the disruption and intervention, passenger travel times were also sampled during the same time periods on November 21, which had very good, generally undisrupted service and therefore served as a benchmark. Although in many systems there may rarely be large time windows or areas of the network that are completely undisrupted, to evaluate the effects of a specific operations control strategy, the analyst only needs to find a benchmark timeframe during which the line section(s) under consideration is (are) generally undisrupted and free of the effects of previous delays on the line. By comparing travel times on the disrupted and the benchmark day, conclusions could be drawn about the passenger impact of the disruption and of the recovery strategy. The travel time distributions were weighted using the same OD flow matrix for both

Table 4 Comparison of passenger impact of the disruption in the morning of April 1, 2008 with November 21, 2008

	Time band 1: Passenger travel times observed before and during time of disruption		Time band 2: Passenger travel times observed during the recovery effort	
	1-Apr (disrupted)	21-Nov (benchmark)	1-Apr (disrupted)	21-Nov (benchmark)
Average travel time per passenger [min]	19.04	16.55	19.69	17.21
Average <i>RBT</i> per passenger [min]	7.35	3.96	9.22	3.65

days to remove the effect of differences in the flow matrices between the two days. The results for both time bands and both days are shown in Table 4. The comparisons clearly show the impact of the disruption management strategy in comparison to the delays caused by the disruption itself. Firstly, the average travel times per passenger calculated based on the total average travel time of Eq. (2) (normalization by total passenger volumes is applied to render the figures in Table 4 comparable) hardly differ between the two time bands, and both are higher than the corresponding benchmark values for November 21. As a matter of fact, passengers after the disruption effectively experienced the same service degradation with respect to November 21 as during the disruption, which is evident from the travel time differences per passenger.

Secondly, in terms of the unreliability of the service as measured by the *RBT*, time band 1 showed a modest increase compared to November 21, amounting to slightly more than 3 minutes per passenger. In other words, passengers during the disruption experienced longer overall travel times than under normal peak-hour conditions, and the service was somewhat less reliable than usual. However, the situation is more severe after the disruption, where the average additional *RBT* amounted to slightly more than 5.5 minutes per passenger or 32 % of the average travel time under generally undisrupted conditions as reflected in the benchmark.

Therefore, not only did the interventions cause an additional average travel time of the same magnitude per passenger as the delays caused by the disruption itself, but the service became considerably less reliable during the recovery than during the disruption. This difference in impact on reliability is presumably due to the two gaps left by the canceled trains as well as to passengers left behind by the crush-loaded trains mentioned earlier, conditions precipitated by the interventions. This result is a clear example of how intervention actions taken by operations controllers, however well intended from their perspective, can lead to increased unreliability experienced by passengers. Aside from the delays to eastbound passengers on the trunk segment that are quantified in this analysis and reported in Table 4, there were additional passenger delays not captured by this analysis that occurred later in the westbound direction since the two canceled trains were missing for an entire round-trip, and therefore, the quantified delays represent lower bounds. A more complete discussion of the analysis, the results and questions that cannot be answered with the passenger travel time data alone is presented by Carrel (2009). Lastly, it is worth mentioning that the present analysis focused only on passenger travel times on the trunk segment

of the line due to sample size limitations. However, the analysis could readily be extended to include passengers traveling from locations on the branches, using the same methodology as presented.

In summary, analyzing this disruption by applying the proposed framework and the corresponding methods associated with each of its elements leads to meaningful insights on operations control. In this particular case, two possible reasons that are found for the controllers canceling the two trains are to manage crew lateness and to avoid train congestion buildup. In the case of the former, it is questionable whether controllers were following the optimal priorities, as the tradeoff between passenger delays and driver delays appears to be strongly out of balance. In the case of the latter, an alternate strategy of holding additional trains upstream of White City would have resulted in lower passenger delays after the disruption. In either case, the disruption management strategy that has been adopted by controllers based on experience and rules of thumb does not appear to be in the best interest of passengers based on the evaluation results derived from the application of the proposed methodology.

To remedy this situation, the transit agency needs to establish common policies and standards which frame the overall priorities of operations control and avoid or diminish negative impacts. However, operations control decisions are also influenced by the design of the operations plan and by the data, information and tools that are available to controllers through the OCS. In this specific case, the design of the OCS makes it difficult for controllers to hold larger numbers of trains since every individual hold must be entered manually in the system and trains cannot be held at all stations. Thus, in a holistic approach for achieving sustained improvements, the aforementioned policies and standards need to be accompanied by changes in vehicle and crew scheduling at the operations planning level and, in the long run, in the OCS design.

4.2 Analysis of a train cancellation policy

Due to various reasons (e.g., shorter than planned rolling stock inspection cycles), the London Underground Central line often suffers from rolling stock shortages, in which case operations controllers must choose which train trips to cancel. A common strategy on the Central Line is to cancel trunk service train trips (referred to as “local runs”), i.e., trains scheduled to run only on the sections between North Acton and Leytonstone, serving central London but not the suburbs on the branch lines. Given that the trunk section also experiences the highest passenger volumes, the question was raised whether this policy causes noticeable delays to trunk passengers whereas passengers from the outer branches experience less significant or no delays.

Once again, the first step involves the reconstruction of interventions of interest. Intervention matrices like the ones shown in Table 2 were used to identify April 1, 2008 as a day on which a steady number of cancellations had been performed repeatedly on the trunk section over a period of three hours in the early afternoon when the scheduled headway was 150 seconds. The next step involved conversations with controllers who provided the reasoning behind the strategy of canceling trunk service in the event of rolling stock shortages. The trunk sections have the highest scheduled frequencies, so in absolute terms, the gaps caused by canceling trunk service trips

were smallest. Furthermore, due to the shorter end-to-end cycle times when compared to branch-service trips, it was easier to re-insert a train into operation once one became available. With regard to the identified cancellations on April 1, 2008, the disruption log and service kilometer plots were used to ensure that the cancellations were in fact due to rolling stock shortages and that no other unobserved major delays occurred during that time. That is, by the early afternoon, the operation had already recovered from the early morning disruption and corresponding interventions of the previous application discussed above.

In quantifying the impact of the cancellations as part of the remaining two elements of the framework following the methodology described in Sect. 3.4, the primary focus was on the trunk section level of service and the travel times of passengers traveling between two stations on the trunk section. In total, there were 11 canceled train trips over the course of the three-hour period on April 1, all of which had been scheduled to serve only the trunk sections of the line and not the outer branches. Service delivery on the trunk sections averaged 93.5 % per hour whereas service delivery on the branch sections averaged 97.3 % per hour. As in the case of the previous application, November 21, 2008 served as a benchmark, where very good service was experienced during the time window of interest—on the branch and trunk sections alike, the average service delivery per hour was 100.7 %. The travel time distributions for two types of passengers were compared between the two days: (a) passengers traveling within the trunk and (b) passengers traveling from the trunk to the branches and vice versa. In general, the interpeak passenger volumes on the trunk are markedly higher than the branch volumes. As in the case of the previous application, the travel time distributions were weighted using the same OD flow matrix for both days.

The results are shown in Table 5. It can be seen that the difference in average travel time between the two days is minimal (in the order of less than 10 seconds per passenger). This is not surprising, as the scheduled frequency on the trunk during the interpeak is high (24 tph) and on average, there were two trains per hour missing. The analysis of the travel time reliability showed only a slightly different picture. The total RBT is between 10 % and 20 % of total travel time, and the differences between November 21 and April 1 do not appear to be large. Nevertheless, it is worth noting that the total RBT is larger on April 1 for the trips within the trunk, whereas it is smaller for the trips between the trunk and the eastern and western branches. Assuming a normal distribution of the RBTs, a standard t-test confirmed the hypothesis that passengers on the trunk experienced less reliable service (at a significance level of 0.01) on April 1 while the experience of passengers traveling between the trunk and the branches is not significantly different from that of November 21. Finally, examining the magnitudes of the differences shows that they are in the order of a few seconds in all cases. Specifically, on April 1, the average RBT on the trunk was 17 seconds larger than on November 21, putting it at 2:33 min. A more in-depth discussion can be found in Carrel (2009).

These results show that, although this strategy does distribute the impacts of rolling stock shortages unevenly and perhaps mostly penalizing the trunk section with its heavy demand, its overall impact is small enough that one can argue that this does not call for immediate attention. However, in the long run, one may think of ways to better align operations control decisions with passenger demand patterns in

Table 5 Comparison of passenger impact of the cancellations in the afternoon of April 1, 2008 with November 21, 2008 (undisrupted benchmark)

	Within trunk		Between trunk and western branches		Between trunk and eastern branches	
	1-Apr	21-Nov	1-Apr	21-Nov	1-Apr	21-Nov
Average travel time per passenger [min]	13.73	13.71	26.66	26.59	21.75	21.85
Average RBT over all OD pairs [min]	2.54	2.27	3.92	4.14	3.86	4.13

the event of rolling stock shortages, for example by incorporating “discardable” trips in the operations plan.

5 Summary and discussion

In this paper, a framework is proposed to empirically evaluate metro line operations based on automatically collected data for the purpose of either identifying limitations to certain operations control strategies and suggesting alternative approaches or validating the effectiveness of such strategies. The development of the framework was motivated by several circumstances. First, operations control research to date has relied heavily on mathematical modeling and mostly focused on individual or small groups of interventions applied to simplified systems. In contrast, current control systems mostly rely on controllers’ experience and judgment on the application of a variety of interventions within a very complex decision environment. This disconnect has strongly limited not only researchers’ ability to conclusively determine through empirical field experiments whether the proposed models are able to provide tangible improvement when compared to the state of practice, but also the adoption of these models in real-world systems. Second, within many transit agencies, there is a need to improve the feedback loop from front-line personnel to planning and management levels in order to assess the adequacy and feasibility of operations plans and operations management policies as well as to validate assumptions and models used in the planning process. Third, and probably most importantly, the availability of train and passenger data from automated data collection systems has markedly increased, making a wealth of information on the functioning of a transit line available. Such data can support empirical evidence-driven modifications or improvements to the understanding of the complexities of line operations.

Furthermore, in this paper, it is argued that any effort to analyze and build a good understanding of operations on a line must not only include operations control, but also consider the full decision environment of controllers and the range of interventions at their disposal. The proposed framework accounts for this by incorporating various perspectives relating to operations control and by integrating automatically collected data from multiple sources, thus facilitating the analysis of operational

questions in an operations control centric manner. Although the specific methodologies and measures associated with the proposed framework were developed for application to the London Underground, and therefore are tailored to the type of information available from the respective OCS and operations control logs, the logic behind the framework's design, the structure of its components, and the procedures for deriving its individual elements are applicable to other metro systems as well. Thus, the specific implementation of the framework can be adapted to other systems with their specific input characteristics by following the structure and principles presented in this paper. The presented methodology is primarily designed to make use of fine-grained passenger movement data, including times and locations at which passengers entered and exited the system. However, in systems where such fine-grained data are not available, the analyst may be able to infer or model travel times depending on the type of data or information available as discussed in Sect. 3.4.

To illustrate the use of the framework and demonstrate its value in arriving at clear conclusions and insights regarding the actions and impacts of operations control, two real-world applications are presented in this paper. It is worth emphasizing that the insights gained are possible as a direct result of the integrated structure of the developed framework and the corresponding methodologies relating to train movement reconstruction, inference regarding the reasons driving controllers' actions, and estimation of operational level of service and passenger travel time metrics. A crucial aspect of this framework is the ability to directly link the controllers' actions with passengers' experiences in the form of both average conditions and the resulting uncertainties. While the examination of operational level of service—the predominant perspective typically considered in practice and in some literature—is valuable, this study demonstrates that on its own, it is limited. Only when the passengers' perspective is incorporated, it is possible to arrive at a rich understanding of the impacts of disruptions and, more critically, the resulting controller's actions on service as experienced by the passengers.

The results produced by following the multi-perspective approach incorporated in the developed framework not only provide benefits for short-term operations planning, but can also be useful for long-term investment planning. For example, a detailed analysis of the use of infrastructure elements (e.g., crossovers and reversing tracks) or equipment and human resources (e.g., spare trains, drivers or mechanics) that provide operational flexibility can show where infrastructure upgrades or resource deployments would be most cost effective. In addition, an analysis of the disruption patterns on a line and of controllers' responses to disruptions where no such infrastructure is available can indicate where the construction of additional capacity (e.g., reversing tracks) would have the largest impact on operations performance. Furthermore, when upgrades are made to operations control systems, a careful assessment of the needs of controllers and the conditions they face can provide valuable benefits during operations. Many of the systems are only customizable before installation and are very difficult to upgrade or amend at a later stage.

In closing, while certain aspects of the implementation of the proposed methodology require tailoring to the nature of the available data, the general manner in which the logic and structure of the framework integrate multiple perspectives and multiple sources of data to arrive at meaningful operational insights and conclusions holds

great promise, particularly in light of the framework's demonstrated value through real-world applications. This should encourage the consideration and implementation of this framework in practice, especially given the increasing availability of the automatically collected data it relies on.

Acknowledgements This research was made possible by the generous support of Transport for London (TfL). The data and expertise provided by TfL management and technical staff are greatly appreciated. The views, opinions, findings, and conclusions reflected in this paper are the responsibility of the authors only and do not represent the official policy or position of TfL.

References

- Abkowitz M (1986) Optimal control of headway variation on transit routes. *J Adv Transp* 20:73–88
- Adamski A, Turnau A (1998) Simulation support tool for real-time dispatching control in public transport. *Transp Res, Part A, Policy Pract* 32(6):73–87
- Carrel A (2009) Diagnosis and assessment of operations control interventions: Framework and applications to a high frequency metro line. MS Thesis, Massachusetts Institute of Technology
- Carrel A, Mishalani R, Wilson N, Attanucci J, Rahbee A (2010) Decision factors in service control on a high-frequency metro line and their importance in service delivery. *Transp Res Rec* 2146:52–59
- Chan J (2007) Rail OD estimation and journey time reliability metrics using automated fare data. MS Thesis, Massachusetts Institute of Technology
- Daganzo CF (2009) A headway-based approach to eliminate bus bunching. *Transp Res, Part B, Methodol* 43(10):913–921
- Daganzo CF, Pilachowski J (2011) Reducing bunching with bus-to-bus cooperation. *Transp Res, Part B, Methodol* 45(1):913–921
- Dixon M (2006) Analysis of a subway operations control database: The MBTA operations control system. MS Thesis, Northeastern University, Boston
- Eberlein X (1995) Real-time control strategies in transit operations: models and analysis. PhD Thesis, Massachusetts Institute of Technology
- Eberlein XJ, Wilson NHM, Bernstein D (2001) The holding problem with real-time information available. *Transp Sci* 35(1):1–18
- Froloff E, Rizzi M, Saporito A (1989) Bases et pratiques de la régulation. RATP, Direction du Réseau Routier RC/MSE
- Furth P, Muller T (2006) Service reliability and hidden waiting time: insights from automatic vehicle location data. *Transp Res Rec* 1955:79–87
- Ji Y, Mishalani RG, McCord MR (2009) Assignment-based methodology to match automated vehicle location transit bus trip trajectories to schedules. In: Proceedings of the 11th international conference on conference on advanced systems for public transport, Hong Kong
- Ji Y, Mishalani RG, McCord MR (2010) Analytical and empirical investigations of the effect of bus drivers' reactions to schedules on transit operations reliability. In: Proceedings of the 12th world conference on transportation research, Lisbon
- Jin JG, Kwong MT, Sun L (2013) Disruption response planning for an urban mass rapid transit network. Presented at the transportation research board 92nd annual meeting, Washington DC
- Kepaptsoglou K, Karlaftis MG (2009) The bus bridging problem in metro operations: conceptual framework, models and algorithms. *Public Transp* 1(4):275–297
- Koffman D (1978) A simulation study of alternative real-time bus headway control strategies. *Transp Res Rec* 663:41–46
- O'Dell S (1997) Optimal control strategies for a rail transit line. MS thesis, Massachusetts Institute of Technology
- Osuna E, Newell G (1972) Control strategies for an idealized public transportation system. *Transp Sci* 6:57–72
- Puong A (2001) A real-time train holding model for rail transit systems. MS thesis, Massachusetts Institute of Technology
- Rahbee A (2001) Rail transit operations analysis: Framework and applications. MS thesis, Massachusetts Institute of Technology

- Shen S (2000) Integrated real-time disruption recovery strategies: A model for rail transit systems. MS thesis, Massachusetts Institute of Technology
- Uniman D (2009) Service reliability measurement using AFC smart card data—A framework for the London Underground. MS thesis, Massachusetts Institute of Technology
- Uniman D, Attanucci J, Mishalani R, Wilson N (2010) Service reliability measurement using automated fare card data: application to London underground. *Transp Res Rec* 2143:92–99
- Walker C, Snowdon J, Ryan D (2005) Simultaneous disruption recovery of a train timetable and crew roster in real time. *Comput Oper Res* 32:2077–2094
- Wang W, Attanucci J, Wilson N (2011) Bus passenger origin-destination estimation and related analyses. *J Public Transp* 14(4):131–150
- Wile E (2003) Use of automatically collected data to improve transit line performance. MS thesis, Massachusetts Institute of Technology