Transportation Research Interdisciplinary Perspectives 30 (2025) 101354

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.elsevier.com/locate/trip

Spatio-temporal impacts of unplanned service disruptions on public transit demand

Benjamin Cottreau [a,b,*], Ouassim Manout [a], Louafi Bouzouina [a]

[a] LAET ENTPE-CNRS-University of Lyon, France
[b] Keolis Lyon, France

ARTICLE INFO

Keywords:
Disruption
Public transit
Demand analysis
Smart card
Ridership
Resilience

ABSTRACT

Public Transit (PT) operators often face unplanned subway service disruptions, which are challenging to handle because they need rapid and effective management strategies. Under such conditions, their goal is to limit user dropout by proposing attractive solutions. With the development of Automatic Fare Collection (AFC) systems and the ability to deal with increasing data volumes, research has recently focused on demand-oriented analysis of disruptions and demand-based strategies. However, AFC data can be incomplete and delayed when uploaded in real time and, therefore, are poorly monitored by PT operators. For this reason, this work seeks to show the relevance of using AFC data to enhance the implementation of disruption management strategies. It aims to understand when, where, and why service disruptions occur. Findings show that 39% of the service disruptions observed between 2021 and 2023 have a negative impact on demand. Thanks to the Gaussian Mixture Model (GMM) algorithm, we distinguish 3 levels of negative impacts: high, medium and low intensity. Results highlight that the PT system is more vulnerable to long-lasting disruptions, which occur during peak hours. The connection with the national railway system also increases the vulnerability of PT systems to disruptions. Using a Multinomial Logistic Regression model, this work highlights the causes that mostly harm the demand and calls for line-specific management strategies. In addition, the contextual analysis introduced in this study reveals the clues left by disruptions in demand levels and argues for the development of online disruption detection coupled with flow redistribution models to enhance decision-making.

1. Introduction

Public Transit (PT) is one of the main levers to achieve a sustainable urban mobility transition. Consequently, PT operators and public authorities increasingly encourage the population to use PT instead of individual and polluting travel modes. The increasing interest in PT raises the question of its efficiency and reliability, especially in the presence of disruptions that can reduce or even hinder PT uptake. *Disruption* is defined as an event that causes a substantial deviation (positive or negative) from the usual performance levels (Bešinović, 2020; Bešinović et al., 2022). Serious PT disruptions can deter the reliability and safety of the PT system and undermine users' satisfaction. Consequently, PT operators and mobility planners need to understand PT disruptions, prevent their advent, and manage their consequences when they occur. A disruption is notably different from an *incident*, which is an unexpected event that does not necessarily affect the performance of PT systems. In the literature, the notion of *disturbance* can also be found and defined as a disruption with a minor impact (Malucelli and Tresoldi, 2019). These differences are relevant as they may call for different analysis and management strategies.

PT operators often store data on *incidents*, which contain information on *disruptions* (Yap and Cats, 2021). We characterize a disruption by two attributes: its **source** (endogenous or exogenous) and its **predictability** (planned or unplanned). The source of a disruption refers to where the disruption occurs. Endogenous disruptions refer to disruptions that occur inside the PT system (e.g. maintenance work, vehicle failure), while exogenous disruptions occur outside the PT system but still have an impact on it (e.g. concerts, strikes, sports events) (Noursalehi et al., 2018). Predictability is the ability to plan a disruption: planned disruptions are known in advance (e.g. concerts, sports events, maintenance work), while unplanned disruptions are unexpected (e.g. failures, unattended baggage, pandemics). The source and predictability of disruptions are essential for optimal disruption management strategies. Exogenous disruptions rely mainly on external sources of data, and their management requires prior cooperation between several organizations. Conversely, endogenous disruptions are exclusively managed by the operator. In addition, planned disruptions

* Corresponding author.
E-mail address: benjamin.cottreau@proton.me (B. Cottreau).

https://doi.org/10.1016/j.trip.2025.101354
Received 24 September 2024; Received in revised form 29 December 2024; Accepted 1 February 2025

can easily be handled in advance, while unplanned disruptions are unpredictable by nature. In the event of planned disruptions, the operator can rely on the available data early enough to prepare relevant substitution plans. However, such informative data are hard to retrieve when unplanned disruptions occur. Unplanned disruption management strategies require rapid implementation to be effective. The measurement of system performance during a disruption is crucial in the implementation of these strategies.

Performance, and therefore disruptions, can be analyzed through the lens of supply or demand Mo et al. (2022). On the supply side, disruptions are assessed through their impact on the PT network or service performance. Network performance assesses the resilience of PT networks and provides information on their optimal design (Bešinović, 2020). This methodology is inspired by graph theory, notably relying on notions such as robustness (Cats, 2016) and vulnerability (Rodríguez-Núñez and García-Palomares, 2014). Service performance focuses on the deviation of the actual supply from the theoretical (or planned) one. On the demand side, disruptions are assessed by characterizing passenger flow redistribution under disruptions. The redistribution process is driven by the willingness of passengers to fulfill their mobility needs while spending a reasonable amount of resources (e.g. money and time). From the passenger's perspective, this process emerges spontaneously when an incident occurs: after a certain amount of time considered unreasonable by customers, the attractiveness of alternatives becomes greater than the solution chosen in the first place. Flow redistribution can also be enhanced with the use of passenger information.

Disruption management strategies can also be grouped into two categories: supply-oriented and demand-oriented. On the supply side, an operator can initiate a service adaptation to restore adherence to the schedule or minimize delays at the vehicle or passenger level. Supply-oriented strategies have always been preferred by PT operators, bridging buses being by far the most used strategy in case of unplanned disruption (Pender et al., 2013; Zhang et al., 2021). On the demand side, management strategies focus on providing relevant information to customers affected by a disruption. The purpose of this strategy is to better allocate the demand in space (*Which alternative stop/line is better for each customer's trip?*) and time (*How long are customers willing to wait before choosing an alternative?*). With the increasing availability of highly granular data, demand-oriented strategies interest researchers and PT stakeholders. The emergence of Real-Time Information (RTI) especially helps better reallocate demand and contributes to a decrease in travel time (Brakewood and Watkins, 2019).

Thus, unplanned disruptions are specifically challenging because they require a quick response from PT operators, with limited ability to monitor the real-time performance of PT systems. The service data are available in real time and already participate in the global management strategy of the PT operators. However, demand data are often incomplete and delayed when uploaded in real-time (Tao and Tang, 2019) and therefore are poorly monitored by PT operators. Hence, when a disruption shows up, PT operators struggle to have a good representation of the evolution of ridership levels from the ground.

In this research, we investigate unplanned service disruptions in the subway system of Lyon, through the lens of demand. This study aims to answer the following questions: *What are the impacts of service disruptions on subway demand levels? How can this information help the implementation of demand-oriented management strategies for subway disruptions? Is there a spatio-temporal variability with respect to these impacts?* The investigation of these questions is a prerequisite for choosing the right strategy to manage disruptions and mitigate their impacts.

The paper is organized as follows: a general review of the literature in Section 2 (Background), methodology and data description in Section 3 (Methods & Data) with results shown and interpreted in Section 4 (Results & Discussions), followed by a conclusion in Section 5 (Conclusions).

2. Background

This work aims to contribute to demand analysis, focusing on service disruptions. To this end, two data types are used: service disruption logs (SD logs) and Automatic Fare Collection (AFC) data.

Service disruptions provoke either a temporary reduction in supply or a temporary cancellation of supply. For these reasons, most studies can retrieve disruptions thanks to service data (Section 2.1). The AFC data are suitable for evaluating demand levels (Pelletier et al., 2011; Fadeev et al., 2018; Pasini, 2021). Nevertheless, surveys can also be used to achieve the same purpose (Rahimi et al., 2020). **Origin–Destination (OD) flows** are useful to retrieve trips' characteristics, such as travel times, distances, or passenger delays. They rely on data for the entry (tap-in) and the egress (tap-out) of the PT system (Section 2.2). **Time series** are valuable for spatio-temporal analysis, as they highlight the amount of ridership at a given stop for a given time period. They only rely on data for entry in the PT system, which makes them suitable for real-time operations (Noursalehi and Koutsopoulos, 2016) (Section 2.3). In this section, we rely on these 3 types of data to identify the main advantages and usage that arise from each of them.

2.1. Service disruption

According to Pender et al. (2013), "any disruption can be categorized according to duration, cause, time, and location". However, the data provided by the PT operators often lack one or several of these information. Therefore, spatio-temporal boundaries of unplanned disruptions can be hard to estimate. Yap and Cats (2021) explained this property of incident data by the fact that most of the time, data related to incidents are filled manually by the operator, in real-time. This results in difficulties in properly interpreting the data, where duplicates and inconsistencies can often be found. In such a case, spatio-temporal boundaries of disruptions can be identified by deriving delays from Automatic Vehicle Location (AVL) and General Transit Feed Specification (GTFS). Yap and Cats (2021) considered disruptions as incidents that provoke a delay superior to 2 min on the Washington subway network, while Marra and Corman (2019) set the delay threshold to 6 min for the Zürich public transit network. Other works mention different delay thresholds, although it is not clear if the operator deliberately does not record the information for small disruptions: Louie et al. (2017) considered delays superior to 2 min using data from Toronto PT network while (Weng et al., 2014) set a threshold at 5 min with Singapore's disruption dataset and 8 min with Hong Kong's disruption dataset. Liu et al. (2024) used GTFS and real-time GTFS data to assess the change in accessibility and service reliability related to short-term (football games) and long-term (COVID) disruptions. These indicators are proven to be useful in detecting PT disturbances.

Hence, taken individually, SD logs provide low-quality information, and several works use AVL or GTFS data to enhance the quality of SD logs. In the following, we investigate the studies using demand data for the same purpose.

2.2. Origin–destination flows

The main benefit of OD flows is the possibility of retrieving information at the passenger level. Some studies investigate the impact of disruptions on passenger delays using OD matrices. Sun et al. (2016) retrieved passenger delays and defined 3 types of customers: customers who give up using PT, customers who take detours, and delayed customers. Considering a disruption that occurred on the Beijing subway network, they estimated that only 26% of the customers were on time during the disruption period, while 12.5% gave up using PT. Paulsen et al. (2021) assessed passenger delays when providing different levels of RTI. Using data from Metropolitan Copenhagen, they have implemented an agent-based path choice model and shown that RTI rationally improves passengers' decisions and thus improves punctuality. These types of frameworks have also been applied to disruption management and the results show that RTI reduces passenger delays (Siegrist et al., 2022; Rahimi et al., 2021), travel times (Mo et al., 2023) and increases user satisfaction (Leng and Corman, 2020).

Other works investigate the impact of disruptions in terms of travel times and variations in the levels of demand. Liu et al. (2021b) used AFC data from urban railway systems to develop demand-based metrics such as average journey time or passenger accumulation. They have also been able to assess spatial and temporal changes in passenger behavior and assess the efficiency of alternative buses. The results highlight that even if the average journey time and passenger accumulation increases in case of disruptions, users mostly prefer to wait and take the subway rather than use bridging buses. These alternative services are less efficient in terms of travel time, being more than 1.5 times higher than travel time using the subway. Malandri et al. (2018) used capacity-based metrics to assess PT disruptions through an agent-based simulation model. Capacity-based metrics, in contrast to demand-only metrics, can help match the demand and supply levels, and point whenever there is a mismatch between both. The authors show that the impacts of disruption can be retrieved up to 15 km from the place where it occurred and last up to 6 times the closure time.

OD flows provide exhaustive information about demand. However, the methodology to retrieve OD flows for real-time operations is computationally expensive, and it requires several sources of data to be appraised (AVL, counting data) depending on the type of fare collection system. For instance, Toque et al. (2016) have proposed a dynamic OD forecasting deep learning model, using the data of the PT system of Rennes, France. Their methodology outperforms traditional approaches, but the authors emphasize the need to reduce computation time for the training phase. The use of raw times series is less restrictive in terms of pre-processing and computational needs. For this reason, we investigate the research that has been done using this data in the context of service disruptions.

2.3. Raw time series

Time series are also suitable for studying ridership levels. The main advantage of working with raw time series is that they require less computing time to retrieve compared to OD flows, making them suitable for real-time operations (Noursalehi and Koutsopoulos, 2016). De Nailly et al. (2022) have investigated the opportunity to use structural decomposition to analyze time series resulting from aggregated flows at the station levels. Their results highlight the potential of this method to quantify the effect of planned (maintenance work) and unplanned (COVID) disruptions with the interpretation of regression coefficients. Individual approaches can also help to understand behaviors: Eltved et al. (2021) have analyzed long-term (i.e. 13 weeks closure) planned disruptions using travel behaviors clusters. This presents the advantage of combining both longitudinal and fare-type data. They have shown substantial changes before and after the disruption, notably bringing out a dropout effect for commuters. However, this approach is less suitable for short-term and unplanned disruption, as longitudinal effects may not be observable for several days or weeks.

Several works have been done to forecast and detect the impacts of disruptions using time series. On the forecasting side, models can be used to assess passenger flow redistribution and provide relevant advice to organize additional or existing supply alternatives (Gao et al., 2020) or provide relevant information to passengers in real-time (Noursalehi and Koutsopoulos, 2016; Noursalehi et al., 2018). Other real-time demand forecasting models could be transposed to disruption scenarios (Zhang et al., 2017; Liyanage et al., 2022).

On the detection side, models are not widely used in the PT literature, although they offer an interesting opportunity to find information on disruptions. Tonnelier et al. (2018) investigated the possibility of retrieving disruptions by detecting anomalies in AFC data. They introduced a range of models capable of detecting disturbances, as well as major breakdowns or PT system failure. Pasini et al. (2022) used short-term prediction models to perform the detection task and implemented a contextual anomaly score. Such detection tasks are relevant for the real-time operational management of disruptions, which can help automate the triggering of disruption management strategies.

Our study distinguishes itself from the existing literature in two main aspects. First, it connects service disruptions with the demand levels by distinguishing several levels of impact of disruptions on the demand. Second, it uses a demand-oriented framework that aims to provide relevant insights for the choice of disruption management strategies.

3. Methods & data

The methodology introduced in this paper is twofold. First, we set the conceptual frame of the time-series analysis that is used to decompose ridership levels into reference and disrupted levels (Section 3.1). Then, we explain the notion of **intensity** of the disruptions' impact, its calculation, and the explanatory model that is used to evaluate the influence of disruptions' attributes on the level of intensity observed (Section 3.2). Lastly, the data provided by the PT operator of Lyon is introduced (Section 3.3).

3.1. Determination of ridership levels

3.1.1. Definition and decomposition of ridership levels

Ridership levels obtained with AFC data are time series (Y_t, t) that can be decomposed using the following form:

$$Y_t = T_t + S_t + R_t, \quad t = [0, 1, \dots, T] \tag{1}$$

T_t is the trend component. S_t is the seasonal component. In the case of PT ridership, 3 seasonal components are classically identified, namely day, week, and annual. Most ridership levels' information is enclosed in seasonality and trend. Hence, we call the estimated term $\hat{Y}_t = T_t + S_t$ the **reference ridership levels**.

R_t is a residual component due to deviations from the reference demand. Let us denote X_t^+ the set of explanatory variables that leads to lower ridership levels and X_t^- the set of explanatory variables that leads to increases in ridership levels. These are included in the residual components such as:

$$R_t = X_t^+ - X_t^- + \omega_t, \quad \omega \sim \mathcal{N}(0, \sigma_\omega^2), \quad X_t^+, X_t^- \geq 0 \tag{2}$$

In other words, while the reference time series is solely explained by seasonality and trend, X_t^+ captures positive outliers, while X_t^- captures negative outliers. We assume that the impact of exogenous shocks on ridership levels can be captured in the residual component. More precisely, we would expect service disruptions to provoke a drop in ridership levels and to be captured by X_t^-. As we do not have comprehensive data about all the events that could yield positive or negative deviations in ridership levels, we are not able to specifically distinguish X_t^- from X_t^+. Consequently, we extract all R_t for each time step t and look at its evolution during service disruptions. To this end, R_t can be estimated by $R_t = Y_t - \hat{Y}_t$.

In some contexts, other events (e.g. football games or concerts) might provoke an increase in demand at the same time as a disruption, such as $X_t^+ > X_t^-$. Focusing on R_t only could lead to incorrect conclusions on the effect of service disruptions on ridership levels. Hence, we need to add contextual information that helps to grasp this decomposition's potentially counter-intuitive results.

3.1.2. Assessment of reference ridership levels

As mentioned above, reference ridership levels are defined by trend T_t and seasonality S_t. The aim of this section is to explain how both terms are calculated.

Trend - To retrieve the trend, we use a simple linear regression fitted with the least-squares method, such as the estimated trend can be written as follows:

$$T_t = \beta t + \alpha \tag{3}$$

Seasonality - Let us assume that the seasonal component comprises s different periods (e.g. monthly seasonality can be decomposed into 12 different periods). Let us denote α_i as the estimated mean of Y_t for the period i, and d_{it} as a dummy variable that takes the value 1 if t belongs to period i, 0 otherwise. Therefore, the seasonal component can be expressed as (Plosser, 1978):

$$S_t = \sum_{i=1}^{s} \alpha_i d_{it} \tag{4}$$

The number of periods s depends on how seasonality is understood. In this work, we analyze the day component on one hand, and the annual and week components on the other hand. The day component captures changes that take place during the day (e.g. peak hours/non-peak hours). According to the level of disaggregation of the data, we have chosen to capture day seasonality at the minute level. Working with such small time steps is only possible with subway data, as the station's inflow is continuous. Tonnelier et al. (2018) has notably worked with the same kind of time series.

The week (e.g. weekdays/weekends) and annual (e.g. working days/school holidays) seasonalities are harder to grasp in the context of recent PT data. COVID has strongly disturbed usual seasonal patterns, because of the impact of lockdowns and curfews, but also because the pandemics caused long-term changes in users' mobility behaviors (Jenelius and Cebecauer, 2020; Kellermann et al., 2022; Cottreau et al., 2023). Consequently, we have defined both week and annual components without considering calendar attributes, but by identifying the days having similar ridership levels using clustering methods. In doing so, we expect to retrieve seasonality patterns more flexibly. For instance, a Monday with an exceptionally low ridership level can be assigned to a cluster that gathers Saturdays because their ridership patterns are similar. The K-means algorithm with Dynamic Time Wrapping (DTW) has been chosen for the study. DTW allows to calculate distances between a set of time steps that are close to each other, instead of only pairing time series values at time t. The resulting minimum distance is chosen and therefore, differences observed in the time series are smoothed. For instance, if a peak hour is observed on a Monday at 08:00 and next Monday at 08:15, these two days have more chances to end up in the same cluster with DTW rather than classic Euclidean distance. This methodology has been selected over other calendar-based, regression (Zargari et al., 2022) or structural decomposition (De Nailly et al., 2022; Deschaintres et al., 2022) methods.

3.2. Determination of disruptions' intensity levels

3.2.1. Definition and calculation of intensity

Intensity is a measurement of the magnitude of the impact of a disruption. It measures the difference in the level of ridership between disruption and the corresponding reference period. A negative difference means a lower level of ridership during disruption while a positive difference means a higher level of ridership. Eq. (5) uses Min-Max scaling at the station level: for each observed station, the greater impact observed sets the value to 1 or −1, and the lower impact (i.e. the ridership level under disruption is very close to reference level) sets the value to 0. Let us denote δ a disruption, for each time step t and each station s, the intensity $M_{s,t,\delta}$ can be measured as:

$$M_{s,t,\delta} = \frac{Y_{s,t,\delta} - \hat{Y}_{s,t}}{max_{t,\delta}(|Y_{s,t,\delta} - \hat{Y}_{s,t}|)} \tag{5}$$

Taking discrete observations of $M_{s,t,\delta}$ for every minute of observed disruption. For a disruption δ lasting n minutes, all $M_{s,t,\delta}$ are stored in a vector $M_{s,\delta}$ of length n. This work aims to cluster disruptions according to their intensity levels. However, the length of vector $M_{s,\delta}$ depends on the corresponding disruption's duration, which is unsuitable for clustering purposes. To overcome this problem, we define vector $I_{s,\delta}$ as the distribution of $M_{s,t,\delta}$ over the disruption period. As $M_{s,t,\delta} \in [-1;1]$, we choose a distribution step of 0.1 which gives 21 intervals within $[-1;1]$ (last interval only contains value 1). For clarity purposes, Fig. 1 illustrates the different steps in defining intensity vectors. Let us consider two disruptions δ_1 and δ_2 occurring at station s, with different durations. Disruption δ_1 lasts 4 min ($n_1 = 4$) and disruption δ_2 lasts 6 min ($n_2 = 6$). For each time step, $M_{s,t,\delta}$ measures the difference between reference ($\hat{Y}$) and disrupted (Y) demand levels, i.e. the **intensity** of the disruption. However, vectors M_{s,δ_1} and M_{s,δ_2} have different lengths and are unsuitable as inputs for the clustering algorithm. $I_{s,\delta}$ represents the distribution of $M_{s,\delta}$ over the disruption period. For each step $p \in [0;20]$, the proportion of $M_{s,t,\delta}$ between $0.1p$ and $0.1(p+1)$ is calculated, resulting in $I_{s,\delta}$ having a fixed length of 21.

A **Gaussian Mixture Model (GMM)** based clustering is then performed on all intensity vectors to identify distinctive disruption profiles according to the intensity of their impact on ridership levels. The advantages of GMM are two-fold. First, GMM assigns a probability of belonging to a given cluster for each disruption. This property is useful for understanding the relationship of different clusters with each other. Second, GMM is particularly useful when working with data including outliers and for anomaly detection purposes (He et al., 2019).

3.2.2. An explanatory approach to intensity levels

Once intensity clusters have been defined thanks to the GMM algorithm, we perform a multinomial logistic regression to know which disruption attribute has a significant influence on the intensity observed. We use the traditional four attributes of disruptions, namely duration, time, location, and cause (Pender et al., 2013), that we adapt according to the data available in this study. We also add a fifth attribute to this framework: consequence. Here is the description of each attribute:

- **Location** is defined as the set of stations impacted by the disruption. If the disruption impacts the whole line, the variable is set to 0; if it impacts a subset of stations within the line, then the variable is set to 1. The line is also introduced as another variable. At the station level, we distinguish stations that have transfers with national train stations, from those that have transfers with other subway stations.
- **Time** is defined using hour intervals. An interval is a 4-hour window, and a total of six intervals are used to capture the whole day. More specifically, peak hours are included in intervals $[\![6;9]\!]$ and $[\![14;17]\!]$, while off-peak hours correspond to intervals $[\![10;13]\!]$, $[\![18;21[\![$ and $[\![22;23]\!] \cup [\![0;1]\!]$. The last interval $[\![2;5]\!]$, mostly corresponds to periods without service (service starts at 5 a.m. in Lyon).
- **Duration** is given in minutes. Four categories are identified: less than 5 min, between 5 and 15 min, between 15 and 60 min, and more than 60 min.
- **Causes** are aggregated groups of the most commonly observed causes of disruptions on the PT network. Table 1 summarizes these groups and their definitions.
- **Consequence** is a boolean variable. If the disruption leads to a total blockage on the set of impacted stations, the value is set to 0; if the disruption leads to a service frequency reduction, then the value is set to 1.

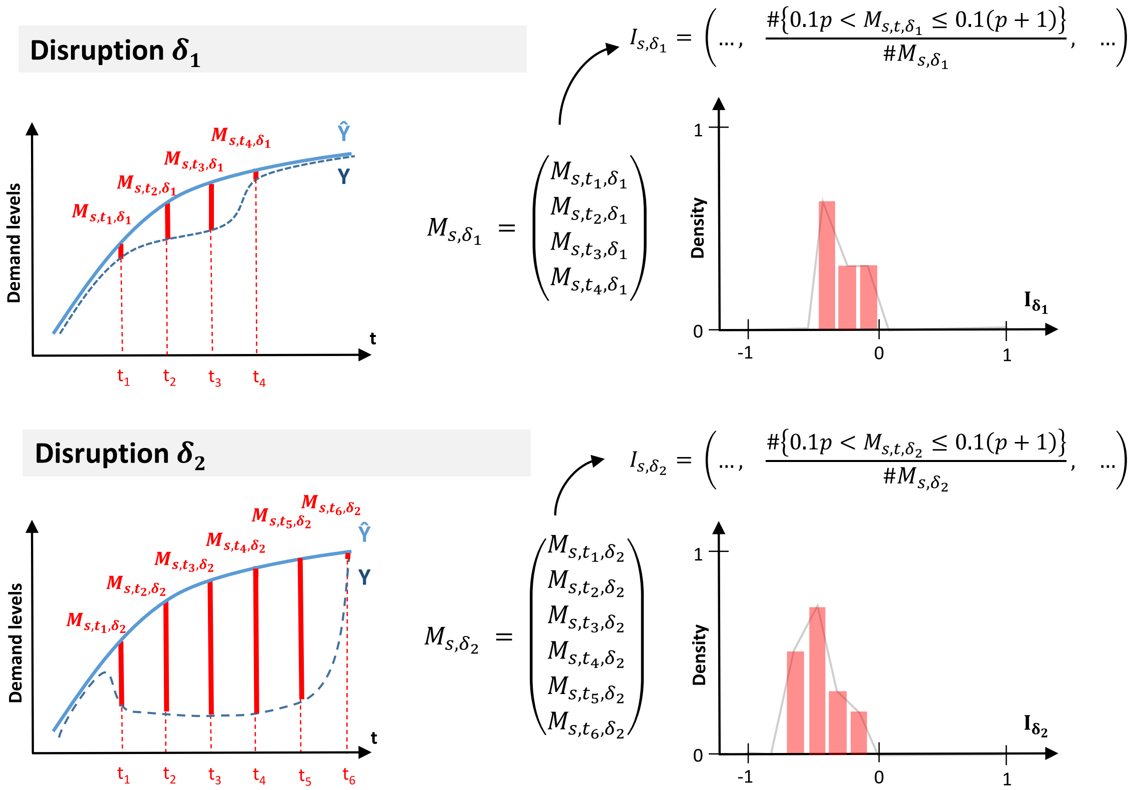


Fig. 1. Definition of intensity vectors - *The two disruptions given as example have different durations and corresponding vectors M have different lengths. Using distribution of M values, intensity vector I have the same length regardless disruptions' duration and is suitable for clustering purpose.*

Table 1
Causes of disruptions - *Aggregated groups of disruptions according to their cause.*

Cause	Definition
MATRO	Rolling stock failure or malfunction
VOYAG	Injured or sick customers, detection of customers on the railways
MAINT	Unplanned or emergency maintenance and operation failure
DETEC	Detection of unattended items on the railways
COLIS	Unattended baggage
FORCE	Force majeure (fire, flood, extreme meteorological events)
PROJET	Incident related to project infrastructure
VENDA	Vandalism
ENERG	Power failure
POSTE	Informatics failure
AUTRE	Other/Unknown cause

In this paper, we investigate the effect of each variable on the intensity to see which of them is more relevant to mitigate service disruptions' impact and improve demand-oriented management strategies. To this respect, we first develop an exploratory analysis to understand the behavior of the considered variable for intensity level and describe the characteristics of the resulting intensity clusters. Then, we design an explanatory approach by building a **Multinomial Logistic Regression** model with the same variables. Variable selection has been performed by first checking the correlation between them, and second using Random Forest (RF) in a 10-fold cross-validation setup. Feature importance scores extracted from RF gave insights into the relevance of implementing the considered features in the model. Friedman H-statistics have been used to assess the effect of feature interactions.

3.3. Case study & data

3.3.1. The public transit network of Lyon, France

The PT network of Lyon comprises 4 subway lines, 2 funicular lines, 7 tramway lines, 26 Bus Rapid Transit (BRT) lines, and more than 130 bus lines are also available. Service disruption data is available for the 4 subway lines only, which explains the focus of this work on the subway system as shown in Fig. 2.

Two major ongoing projects will complete the network with 2 additional tramway lines by 2026. Also, subway line B has been subject to two projects during the same period. First, the line has been extended to its southern terminus and has been put into service on October 20th 2023. Second, this line has been fully automated since June 25th 2022. The network covers the whole Lyon metropolitan area, which represents 746 km^2 where 1.4 million people live. More than 2 million trips are made each day, and approximately 75% of them are made using a smart card. Of all modes, the subway represents almost half of the trips. Similar to many countries COVID had a significant impact on transit ridership in France and Lyon (Cottreau et al., 2023; El Zein et al., 2022; Manout et al., 2023).

3.3.2. Data

We use two data sources in this work: Automatic Fare Collection (AFC) and Service Disruption logs (SD logs).

Automatic fare Collection (AFC) data are a good proxy for PT demand. Access to the subway system in Lyon requires a ticket or a smart card validation. These detailed validations are collected by the operator. In this paper, individual validations are not needed, thus the AFC data was aggregated using 1 min time spans at the subway station level. The resulting time series is called *ridership levels* in this study.

Service Disruptions logs (SD-logs) provided by the operator contains 3 years of data: 2021, 2022, and 2023. During this time, 124,704 incidents were registered and 4593 of them were labeled with a consequence showing an impact on the PT system's performance, and hence were considered service disruptions. When distributing the disruptions across subway stations and splitting them by cases, we end up with 55,937 observations. The existing 4 subway lines of the PT system of Lyon, France are considered in this work. Table 2 shows the distribution of service disruptions according to all the features studied.

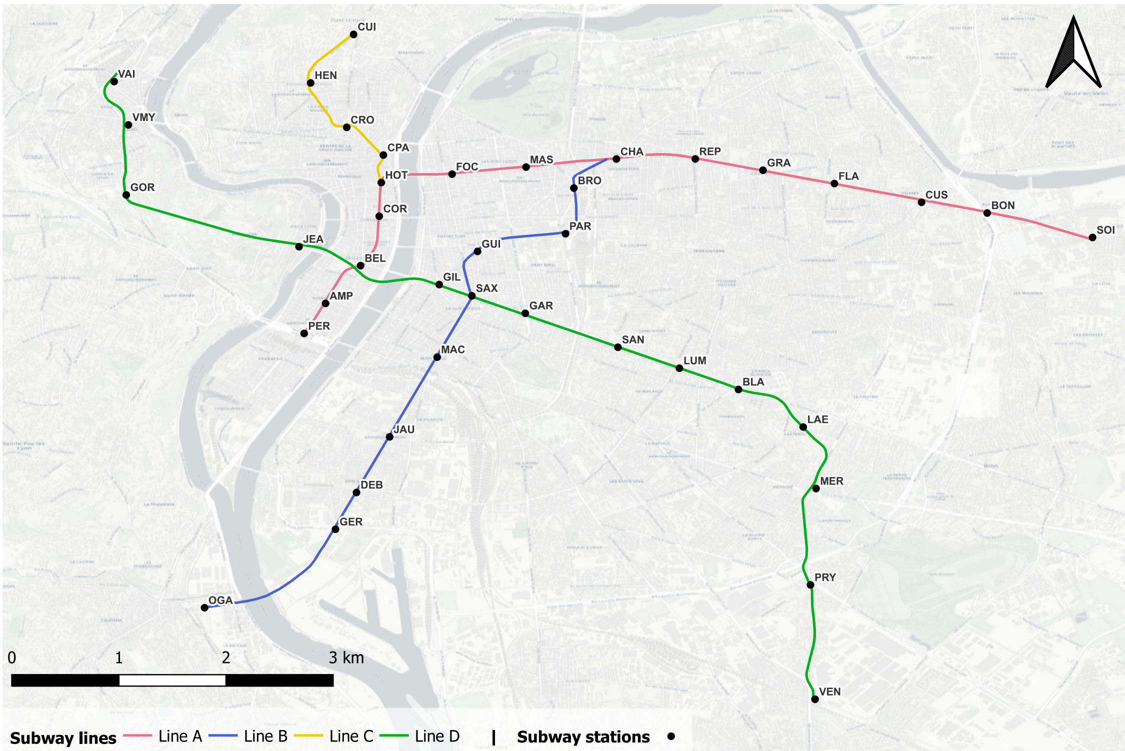


Fig. 2. Subway network of Lyon, France - *The study focuses on subway service disruptions that occurred on line A, B, C and D between 2021 and 2023.* (map tiles: OpenstreetMap).

4. Results & discussions

4.1. Impact intensity and spatio-temporal variability of service disruptions on the ridership levels

First, a general overview of clusters' membership is shown in 4.1.1. The exploratory analysis is then broken down at the temporal 4.1.2 and the spatial levels 4.1.3. Finally, 4.1.4 introduces an explanatory model that assesses the role of disruptions' cause, consequences, and location on the level of intensity. The model also uses the spatio-temporal characteristics studied in the exploratory analysis as control variables.

4.1.1. Resulting intensity clusters

First, the parameters used to retrieve the reference period have been calibrated. For each subway station, the trend of the 1-minute time series has been estimated. Findings show that $\beta < 5.10^{-6}$ and $0 < \alpha < 14$, which is deemed negligible regarding the levels of ridership involved (10 million tap-in validations a station during the studied period on average). Then, the optimal number of clusters is found as described in Appendix B and set to 5 ($R^2 = 0.69$). Estimated level of ridership $\widehat{Y}$ resulting from this procedure is compared to each actual level of ridership during disruptions using Eq. (5). Resulting intensity vectors$I_{\delta,l}$ are used as inputs for GMM clustering for each disruption δ and each subway station l.

Using GMM, five clusters are found as shown in Fig. 3. Three of them gather negative intensity vectors (NEG), one is neutral (NEU) and one is positive (POS).

Negative impact (39%) - Three different clusters correspond to a negative impact on PT demand. NEG+, NEG, and NEG− gather respectively 7%, 11%, and 21% of the intensity vectors. Demand levels are 30% inferior to the reference levels for NEG+, 11% for NEG, and 2% for NEG−. For that reason, we will also refer to them as **high-intensity**, **medium-intensity** and **low-intensity** clusters respectively. High-intensity cluster is characterized by long-lasting disruptions (40 min on average) while medium and low-intensity clusters are composed of disruptions twice shorter on average (between 16 and 21 min).

Neutral impact (53%) - NEU is characterized by values of demand similar to the reference levels and medium-lasting disruptions (19 min on average). The neutral cluster covers a range of different disruptions, as the distribution of intensity vectors spreads out around zero. However, the main distinguishing property of the neutral cluster is that the maximum value observed in the negative spectrum for all intensity vectors is 0.62, while it is 1 for clusters NEG+, NEG, and NEG−.

Positive impact (8%) - POS is characterized by values of demand on average 19% superior to the reference levels. It corresponds to short-lasting disruptions (6 min on average), that mainly occur during peak hours. It seems counter-intuitive that disruptions can lead to an increase in ridership levels and several hypotheses can be made regarding its existence. Data in this cluster could be interpreted as short-lasting disruptions that have a neutral impact on PT ridership but have been captured during a period of time with abnormally high ridership levels, such as planned events (see Section 3.1 for formal explanations). Unfortunately, this hypothesis cannot be formally tested in this work because of the unavailability of an exhaustive dataset including such events. However, we provide contextual information in Section 4.2, which helps to understand in which spatio-temporal context the disruption occurs and the existence of this cluster.

4.1.2. Temporal signature of intensity clusters

As mentioned in the previous section, the duration of disruptions seems to play a role in the intensity of its impact. However, the time of occurrence of disruptions is another crucial temporal factor to take into account. Fig. 4 shows the deviation of each cluster's hourly distribution of $M_{t,\delta}$ compared to the hourly distribution of $M_{t,\delta}$ considering the whole dataset. Consequently, each distribution in Fig. 4 can be interpreted as the temporal signature of the considered cluster. The blue curve corresponds to the distribution of the starting time of disruptions t_{start}, while the red curve describes the distribution of the starting time of disruptions t_{end}.

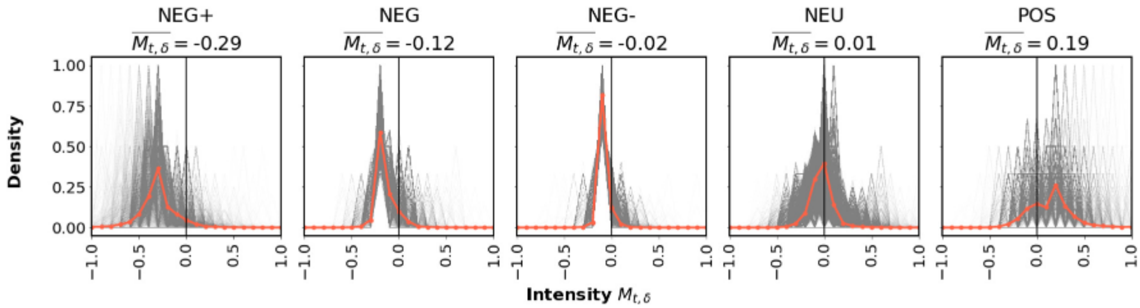


Fig. 3. Intensity vectors per cluster - *Clusters are ranked thanks to the level of intensity of disruptions on the ridership levels, from the most negative (NEG+) to the positive (POS) impact. Red curves indicate the cluster centers, i.e. the mean of all intensity vectors belonging the considered cluster.* (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

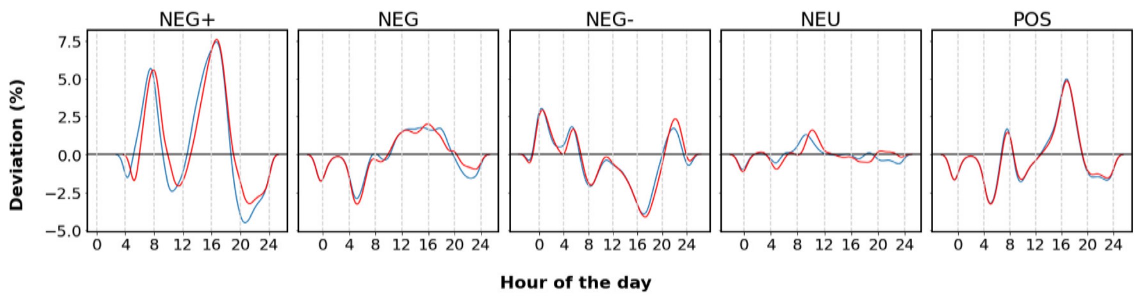


Fig. 4. Temporal signature of clusters - *Each plot shows the deviation between the hourly distribution of $M_{t,\delta}$ for the considered cluster and the hourly distribution of $M_{t,\delta}$ considering the whole dataset. The beginning t_{start} (blue curve) and the end t_{end} (red curve) are displayed for each plot.* (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Each cluster has a signature: NEG+ cluster is predominant in the morning (+5% maximum compared to the sample distribution) and evening peak hours (+7.5%). NEG is predominant during the afternoon (+2.5%) while NEG−is predominant during the evening and nighttime (+2.5%). NEU is close to the sample distribution as it is the dominant cluster. POS is predominant for morning peak hours (+2%) and evening peak hours (+5%). The start and end of disruptions are close to each other, except for NEG+ where a slight lag can be observed. This is because NEG+ disruptions are generally longer than disruptions from other clusters.

This work shows that the PT system is more vulnerable to disruptions during peak hours, i.e. when the highest levels of demand are met. From this perspective, our results are in line with Yap and Cats (2022). The authors have shown stronger demand response during weekdays' peak hours than off-peak hours or weekends when trips can be more easily rescheduled. To a lesser extent, results from our work also highlight the negative impact during off-peak hours with cluster NEG and NEG−. The observed variability in demand levels under disruptions is particularly useful for anomaly detection purposes. For instance, Tonnelier et al. (2018) have noticed that depending on the intensity of disruption on ridership levels, different detection models can be used. The authors notably introduced a "continuous user-based model", which outperforms other models when small impacts are observed. According to our results, this model seems particularly relevant for detecting disruption during off-peak hours.

4.1.3. Spatial characteristics of intensity clusters

Line level analysis - Fig. 5 shows that the intensity of disruptions' impact is consistent with the line where it occurs. A relevant indicator to assess the intensity at the line (or the station) level is the ratio between time spent under high-intensity disruptions and the total time spent under disruptions. Line A endures the strongest impacts of all lines as more than 25% of the time spent under disruptions for this line is related to high-intensity disruptions. Line B is the most impacted line (> 3% of the operating time is spent under disruptions), but the intensity of these disruptions is relatively lower, as 10% of them correspond to the high-intensity cluster. Line C is impacted 2,5% of the time. However, the ratio between the time under high-intensity disruptions and total disruption duration is the best of all lines (i.e. less than 1 out of 35). Finally, line D is the less frequently impacted line, as approximately 1% of the operating time is spent under disruptions. On average, 20% of the disruptions belong to the high-intensity cluster, which makes this line the second most intensely impacted line after line A. The high proportion of disruptions observed for line B can be attributed to its recent automation. Indeed, the debugging of automated lines causes an increase in short-lasting disruptions. Other variables such as the aging of the infrastructure or the vehicles, or the design of the existing infrastructure can be taken into account to explain the strong interline variability observed in Fig. 5. Risk-based management, asset management (Liu et al., 2021a), or restoration plans (Bešinović et al., 2022) are relevant fields of study to improve the disruptions management strategies in the long run. For instance, line C has several portions of infrastructure with only one track, which allows the transit of only one subway at a time. Consequently, other vehicles are being forced to wait at a station, which can cause disturbances on the line and explain the high proportion of time spent under disruption observed for line C. On this topic, Pender et al. (2013) highlighted the benefits of having crossovers for PT operators.

Station level analysis - At the station level, PER and PAR have the highest share of time spent under high-intensity disruptions (respectively 0.61 % and 0.59%). These 2 stations are connected with the national railway system, which makes them vulnerable to specific disruptions such as unattended baggage. Noteworthy, BEL (1.98%), CHA (0.78%), HOT (1.30%) and SAX (0.57%) have respectively their lines' highest shares of the neutral cluster. These are subway-to-subway transfer stations, for which line-specific ridership cannot be measured (e.g. line A and D serving station BEL). Consequently, the impacts of disruptions occurring at transfer stations are underestimated in this study. Transfer stations highly contribute to the resilience of PT systems from the network theory perspective (Weng et al., 2014; Wang et al.,

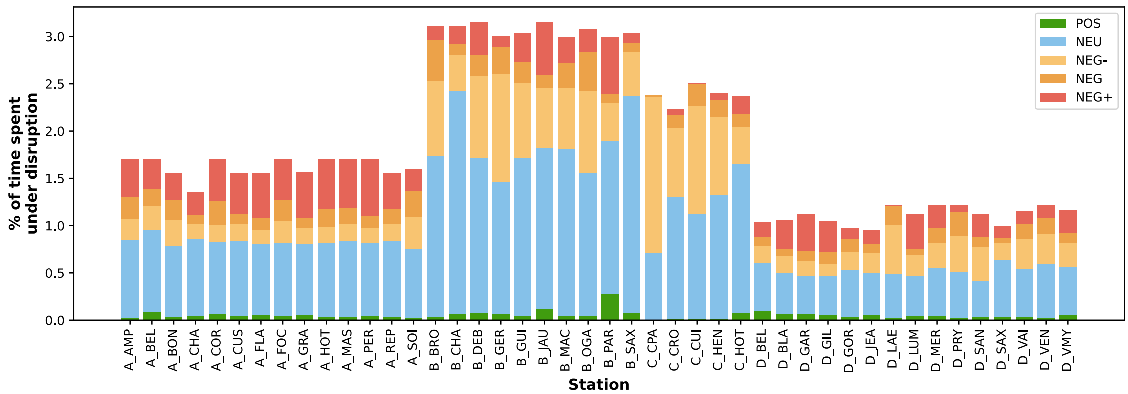


Fig. 5. Spatial distribution of clusters - *For each subway station, the distribution of time spent under disruption is displayed by cluster. The time variable is expressed as the proportion of time spent under disruption compared to the total operating time. A_AMP means line A station AMP.*

Table 2
Distribution of service disruptions according to all the variables considered in the study.

	Frequency Num.	(%)
Cause		
AUTRE	9689	17
COLIS	907	2
DETEC	5290	9
ENERG	266	< 1
FORCE	777	1
MAINT	10,295	18
MATRO	13,590	24
POSTE	174	< 1
PROJE	404	1
VENDA	287	1
VOYAG	14,258	25
Consequence		
Blockage (0)	46,481	83
Frequency reduction (1)	9456	17
Duration		
< 5	29,825	53
≥ 5 min	13,157	24
≥ 15 min	8674	16
≥ 60 min	4281	8
Hour intervals		
$[\![2;5]\!]$	2631	5
$[\![6;9]\!]$	12,184	22
$[\![10;13]\!]$	9354	17
$[\![14;17]\!]$	14,456	26
$[\![18;21]\!]$	12,204	22
$[\![22;23]\!] \cup [\![0;1]\!]$	5108	9
Location		
Full line (0)	52,599	94
Subset of stations (1)	3338	6
Line		
A	16,850	30
B	14,386	26
C	3694	7
D	21,007	38
Transfers		
None	43,337	77
Train station	9936	18
Subway station	2637	5

2020; Yap and Cats, 2021). Some improvements could be made with the consideration of the entry gates related to each validation. Applying ratios based on counting data or working with OD matrices rather than aggregated flows is better for assessing the impact of disruptions on transfer stations, although these approaches are less suitable for real-time management.

4.1.4. Causes and consequences of disruptions: critical factors to improve management strategies

In the following, we investigate the effect of **cause**, **consequence**, and **location** on the intensity of disruptions. To this respect, we have built a Multinomial Logistic Regression model that takes into account the variables described in Section 3.3. Variable selection has been performed by first checking the correlation between them, and second using Random Forest (RF) in a 10-fold cross-validation setup. Feature importance scores extracted from RF gave insights into the relevance of implementing the considered features in the model. Friedman H-statistics have been used to assess the effect of feature interactions, and 3 interaction terms appeared as significant, namely *hour intervals* × *and transfers*, *duration* × *transfers*, and *line* × *cause*. From our perspective, the interaction between line and cause is the most meaningful and was kept in the model. Only significant interactions are displayed in Table 3.

Table 3 indicates the results from the Multinomial Logistic Regression. Results from the exploratory analysis of the temporal (duration, hour intervals) and the spatial (line, transfer stations) aspects are confirmed by the model.

Cause - The category "AUTRE" has been chosen as the reference because this group gathers different kinds of disruptions, that tend to follow the distribution of the whole sample regarding other attributes. Compared with this group, informatics failure (POSTE, $OR = 1.37$), incidents related to project infrastructure (PROJE, $OR = 2.80$), and unattended baggage (COLIS, $OR = 1.49$) have more chances to provoke high-intensity disruption. Unattended baggage has little more chances to provoke medium-intensity disruption (COLIS, $OR = 2.02$), and so do power failures (ENERG, $OR = 3.99$). Power failures also have chances to provoke low-intensity disruptions (ENERG, $OR = 5.10$), just like incidents related to project infrastructure (PROJE, $OR = 3.00$) and unattended baggage (COLIS, $OR = 1.37$). Rolling stock failures (MATRO), unplanned maintenance (MAINT), and customers-related incidents (VOYAG) have more chances to provoke neutral impact (NEU, reference class for target variable) because all significant odd ratios observed are below 1. Vandalism (VENDA) is not significantly representative of any class of intensity. The positive cluster is not positively related to any cause of disruptions. Looking at interactions with lines for the high-intensity cluster, we see that :

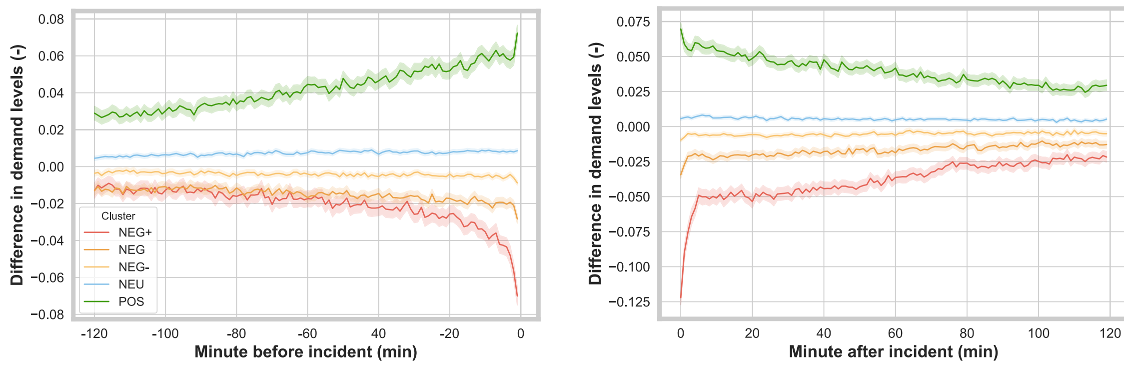
- **Line B** - unplanned maintenance (MAINT*B, OR = 1.39), rolling stock failure (MATRO*B, OR = 1.55) and customer-related incidents (VOYAG, OR = 1.48) on line B have little more chances to provoke high-intensity disruptions, compared to line A. This is due to the ongoing debugging operations consecutive to the recent automation of subway line B.

Table 3
Multinomial Logistic Regression results.

	NEG+		NEG		NEG−		POS	
	OR[a]	95% CI[b]	OR[a]	95% CI[b]	OR[a]	95% CI[b]	OR[a]	95% CI[b]
Cause								
AUTRE	–	–	–	–	–	–	–	–
COLIS	**1.49****	1.13, 1.97	**2.02*****	1.56, 2.61	**1.37***	1.06, 1.77	0.90	0.60, 1.34
DETEC	0.78	0.56, 1.10	**0.74****	0.59, 0.92	1.05	0.94, 1.17	1.05	0.91, 1.22
ENERG	1.25	0.53, 2.94	**3.99*****	2.12, 7.48	**5.10*****	2.78, 9.34	1.12	0.34, 3.66
FORCE	0.96	0.68, 1.37	1.17	0.83, 1.64	1.07	0.79, 1.46	1.39	0.92, 2.08
MAINT	1.01	0.81, 1.26	0.83	0.67, 1.03	1.10	0.94, 1.30	**0.79***	0.63, 1.00
MATRO	**0.62*****	0.52, 0.74	**0.81***	0.69, 0.94	**0.85***	0.74, 0.97	0.85	0.72, 1.01
POSTE	**1.37***	1.05, 1.81	1.25	0.98, 1.59	0.96	0.77, 1.19	1.03	0.73, 1.44
PROJE	**2.80****	1.31, 5.99	0.85	0.26, 2.76	**3.00****	1.53, 5.86	0.21	0.01, 4.40
VENDA	0.42	0.12, 1.41	1.15	0.60, 2.20	1.12	0.57, 2.18	0.98	0.45, 2.16
VOYAG	**0.81***	0.68, 0.97	0.97	0.82, 1.13	0.99	0.86, 1.14	1.15	0.96, 1.38
Consequence								
Blockage (0)	–	–	–	–	–	–	–	–
Frequency (1)	**0.13*****	0.11, 0.16	**0.25*****	0.21, 0.28	**0.53*****	0.48, 0.58	**0.79****	0.69, 0.91
Duration								
< 5	–	–	–	–	–	–	–	–
≥ 5 min	**0.69*****	0.62, 0.76	**0.52*****	0.48, 0.56	**0.58*****	0.54, 0.61	**0.45*****	0.41, 0.49
≥ 15 min	**1.59*****	1.40, 1.79	**0.56*****	0.50, 0.63	**0.51*****	0.47, 0.56	**0.18*****	0.16, 0.21
≥ 60 min	**7.06*****	6.19, 8.04	**1.23****	1.07, 1.41	**0.81*****	0.72, 0.90	**0.14*****	0.11, 0.19
Hour intervals								
$[\![2; 5]\!]$	–	–	–	–	–	–	–	–
$[\![6; 9]\!]$	**7.25*****	5.21, 10.1	**8.08*****	5.69, 11.5	**0.35*****	0.32, 0.39	**10.3*****	6.34, 16.6
$[\![10; 13]\!]$	**6.32*****	4.53, 8.81	**11.7*****	6.79, 13.7	**0.30*****	0.28, 0.34	**7.65*****	4.60, 12.1
$[\![14; 17]\!]$	**15.2*****	11.0, 21.1	**13.7*****	8.02, 16.1	**0.22*****	0.20, 0.24	**15.8*****	9.56, 24.9
$[\![18; 21]\!]$	**5.74*****	4.13, 8.00	**11.3*****	6.54, 13.2	**0.41*****	0.37, 0.45	**10.9*****	6.57, 17.2
$[\![22; 23]\!] \cup [\![0; 1]\!]$	**0.41*****	0.25, 0.67	**3.96*****	2.28, 4.82	**1.22*****	1.11, 1.36	**5.08*****	2.96, 8.07
Location								
Full line (0)	–	–	–	–	–	–	–	–
Subset (1)	**5.85*****	5.14, 6.66	**3.67*****	3.23, 4.17	**1.90*****	1.69, 2.13	**1.32***	1.04, 1.67
Line								
A	–	–	–	–	–	–	–	–
B	**0.27*****	0.22, 0.35	**0.57*****	0.48, 0.69	**1.32*****	1.15, 1.52	**0.65*****	0.53, 0.79
C	**0.07*****	0.03, 0.16	**0.54****	0.36, 0.82	**1.95*****	1.54, 2.46	**0.54***	0.31, 0.93
D	**0.49*****	0.39, 0.60	**0.80***	0.67, 0.95	1.13	0.98, 1.31	0.91	0.74, 1.11
Transfers								
None	–	–	–	–	–	–	–	–
Train station	**2.59*****	2.25, 2.99	**1.17***	1.02, 1.35	**0.63*****	0.55, 0.72	**2.09*****	1.84, 2.38
Subway station	**0.63*****	0.57, 0.70	**0.77*****	0.71, 0.83	**0.80*****	0.75, 0.85	**0.77*****	0.70, 0.84
Cause*Line								
DETEC * B	0.52	0.27, 1.01	**0.53****	0.34, 0.82	1.15	0.94, 1.40	**1.35***	1.02, 1.79
FORCE * B	0.46	0.18, 1.14	**0.41***	0.18, 0.93	0.94	0.57, 1.54	**0.24***	0.06, 0.88
MAINT * B	**1.39***	1.01, 1.91	**1.40***	1.06, 1.86	0.89	0.73, 1.09	**1.93*****	1.44, 2.59
MATRO * B	**1.55****	1.15, 2.09	**1.30***	1.03, 1.64	**0.82***	0.68, 0.99	**1.36***	1.05, 1.75
PROJE * B	0.41	0.16, 1.05	1.32	0.37, 4.75	**0.29****	0.14, 0.61	6.79	0.30, 152
VOYAG * B	**1.48***	1.08, 2.04	1.17	0.91, 1.50	0.83	0.68, 1.01	0.91	0.69, 1.20
ENERG * C	**51.6****	3.37, 788	5.10	0.45, 57.2	0.02	0.00, 111	1.01	0.00, 63,342
FORCE * C	**45.3*****	12.3, 167	**4.64****	1.49, 14.4	**4.81*****	2.28, 10.1	0.51	0.02, 16.1
MAINT * C	**2.67***	1.05, 6.76	**2.11****	1.26, 3.54	1.22	0.91, 1.63	1.24	0.61, 2.51
MATRO * C	**3.13****	1.32, 7.42	1.40	0.89, 2.20	0.90	0.69, 1.17	0.99	0.55, 1.79
VOYAG * C	**6.14****	2.03, 18.5	1.82	0.95, 3.50	1.02	0.66, 1.59	1.34	0.57, 3.16
DETEC * D	**1.51***	1.06, 2.16	**1.39****	1.10, 1.77	0.91	0.81, 1.04	**0.78****	0.65, 0.93
ENERG * D	0.59	0.20, 1.79	**0.18*****	0.07, 0.46	**0.11*****	0.05, 0.27	0.41	0.07, 2.28
FORCE * D	1.05	0.55, 2.03	1.06	0.58, 1.93	1.53	0.91, 2.58	**0.29***	0.09, 0.87
MATRO * D	**1.60****	1.20, 2.13	1.20	0.96, 1.51	1.00	0.82, 1.21	1.19	0.92, 1.53

* Significance levels : p-value < 0.05.
** Significance levels : p-value < 0.01.
*** Significance levels : p-value < 0.001.
[a] OR = Odd Ratio.
[b] CI = Confidence Interval.

- **Line C** shows high odd ratios for power failures (ENERG*C, OR = 51.6) and force majeure (FORCE*C, OR = 45.3). Also, the high vulnerability of line C towards power failure and force majeure (fire, flood, extreme meteorological events) can be explained by the fact that its infrastructure is partially above ground. Hence, the line is more exposed to this kind of event than other lines, which are all entirely underground. More generally, the infrastructure and the vehicles used on line C are very different from other lines as it is the only rack-and-pinion subway of the network. This may explain the special results obtained for this line, which call for specific management strategies. Analyzing the effect of other variables such as weather (wind and rain especially) on the occurrence of service disruption is particularly relevant for this line.
- **Line D** shows a high odd ratio for the detection of unattended items (DETEC*C, OR = 1.51) because it is the only automated line to benefit from a detection system throughout the entire studied



(a) Evolution of $M_{t,\delta}$ 120 minutes **before** the beginning of disruptions for each cluster

(b) Evolution of $M_{t,\delta}$ 120 minutes **after** the beginning of disruptions for each cluster

Fig. 6. Temporal consistency of disruptions for each cluster.

period. Rolling stock failure (MATRO, OR = 1.60) also have more chances to provoke high-intensity disruption compared to line A and other causes.

The cause of disruptions is a critical factor to distinguish high-intensity from low-intensity disruptions. Our results are in line with Pender et al. (2013), who concluded that rolling stock failure (MATRO) led to frequency reduction and therefore lower impact on demand. The conclusion drawn for railway systems is closer to the results obtained for line C, especially regarding meteorological events (FORCE). Noteworthy, Yap and Cats (2021) have shown the relevance of implementing cause in their disruption exposure and disruption impact prediction models.

Location and Consequences - For all clusters, disruptions that provoke a frequency reduction tend to have significantly less impact than a blockage. The more negative the intensity, the lower the ratios (NEG+, $OR = 0.13$; NEG, $OR = 0.25$; NEG−, $OR = 0.53$; POS, $OR = 0.79$). Therefore, frequency reduction tends to provoke a neutral impact on demand. For all clusters, disruptions that impact only a subset of stations have a stronger impact on demand. The more negative the intensity, the higher the odd ratios (NEG+, $OR = 5.85$; NEG, $OR = 3.67$; NEG−, $OR = 1.90$; POS, $OR = 1.32$).

This work shows that a diminution in subway frequency provokes minor changes in ridership levels at the subway station levels in Lyon. In these cases, capacity-based metrics are relevant to properly assess the impact of such disruption on the quality of service (Malandri et al., 2018). At the passenger level, these situations often cause overcrowding and denied boarding, which increases waiting times and dissatisfaction. For that purpose, real-time information is proven to improve waiting times due to denied boarding (Drabicki et al., 2021). This work also indicates that a supply shortage provokes a decrease in ridership levels. Stronger discrepancy is observed when supply shortage is limited to a subset of stations rather than a whole line, because of the accuracy brought by more detailed spatial information. In these cases, anomaly detection and demand prediction for reallocation purposes are recommended tracks of research.

4.2. The contribution of contextual analysis to implement relevant demand-oriented management strategies

The deviation of performance levels observed during disruptions is relevant when considering the context in which they occur. Several events (e.g. sports games or concerts) can overlap and SP logs do not capture all the information needed to comprehend the complexity of observed situations. To address this problem with the lack of data about other events that might interfere with our data, we choose to analyze the spatio-temporal context related to the disruptions observed.

For the sake of the analysis, location and consequence have been merged into a single variable called **type** in this section. Type 1 corresponds to disruptions localized on a part of a line (location = 1) and leads to a blockage (consequence = 0). Type 2 corresponds to disruptions localized on the whole line (location = 0) and leads to a blockage (consequence = 0). Type 3 corresponds to disruptions localized on the whole line (location = 0) and leads to a frequency reduction (consequence = 1). However, a frequency reduction on a part of a line is never observed in the data. Indeed, it is operationally challenging to manage two different headways for the same subway line during a disruption.

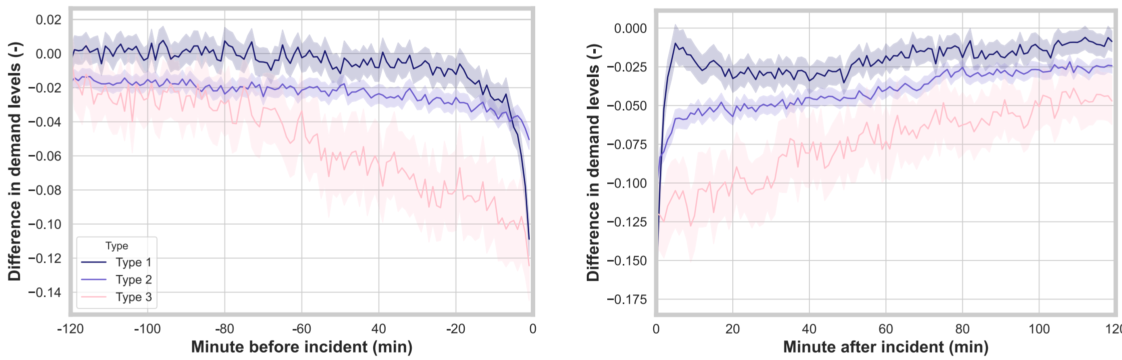
On the temporal aspects (Section 4.2.1), disruptions may be affected by **cascading effect**. For a single disruption, two types can be consecutive to each other due to the cascading effect. This happens when the operator first identifies a disruption that blocks the whole line (Type 2) and then manages to restore the traffic on one part of the line (Type 1).

On the spatial aspects (Section 4.2.2), we rely on the calculation of a spatial consistency index, whose details are presented in Appendix A. The calculation of the spatial consistency index relies on the outputs of GMM clustering. For each disruption, we check whether all impacted stations end up in the same cluster. If yes, the index tends to 1 and the disruption is said to be *spatially consistent*. Otherwise, the index tends to 0 and the disruption is said to be *not spatially consistent*.

4.2.1. Temporal context of disruption: early response and recovery patterns of demand

In order to analyze the temporal context of the set of disruptions observed, we check the ridership levels 120 min before the disruptions' starting time t_{start} and after their ending time t_{end}.

Fig. 6(a) shows the evolution of $M_{t,\delta}$ 120 min before the beginning of the disruption, averaged over all disruptions belonging to the same cluster. The starting points of ridership levels 120 min before t_{start} are different among clusters: POS ridership levels are already 3% higher than reference levels, NEG and NEG+ start −1% lower than reference levels, while NEG−and NEU are between −0.5% and +0.5% compared to reference levels. From there, a linear increase in ridership levels is observed for cluster POS, while NEU and NEG−are quite stable over time. NEG−shows a slight decrease over time (−2%), while NEG+ shows a discrepancy few minutes before the disruption is actually registered (−7%). We call this phenomenon the **early response** of demand regarding disruptions. Symmetrical results are observed 120 min after the disruptions in Fig. 6(b). The **recovery patterns** for NEG+ are particularly interesting. During the first 5 min after the disruptions, ridership levels increase from −12% to −5%. Then from 5 up to 80 min after the disruption, a slow increase is observed from −5% to −2.5%. To understand better the early response and the recovery patterns



(a) Evolution of $M_{t,\delta}$ 30 minutes **before** the beginning of disruptions for cluster **NEG+** and for each type

(b) Evolution of $M_{t,\delta}$ 30 minutes **after** the beginning of disruptions for cluster **NEG+** and for each type

Fig. 7. Temporal consistency of disruptions for cluster NEG+ and each type.

observed for the set of most intense disruptions, we investigate the evolution of $M_{t,\delta}$ for NEG+ according to the type of disruptions on Fig. 7.

Early response - Fig. 7(a) indicates that disruptions of type 1 start with demand levels around 0% difference compared to the reference, while type 2 and 3 start around −2%. Disruption of type 1 initiates a decrease 20 min before t_{start}, and a sharp decrease is observed 5 min before t_{start}, finally reaching −11%. Demand levels for type 2 also decrease, reaching −5% at t_{start}. Finally, demand levels for type 3 show a linear decrease, starting approximately 80 min before the disruptions begin, and associated ridership levels reach −10% at t_{start}. 85% of type 1 disruptions are the results of cascading effect consecutive to type 2 (i.e. 2/1). Consequently, the vast majority of type 1 disruptions are consecutive to a first drop in demand levels, which explains the very early response observed. Demand levels for type 2 also show an early response. However, its magnitude (−3%) is less significant than type 1.

Two hypotheses can be made regarding the early decrease observed for type 3. First, type 3 disruptions really begin more than one hour before the registered beginning of the disruption, which is very unlikely. Second, ridership levels decrease but are not related to the disruption itself, which is more likely. Indeed, type 3 are under-represented in NEG+ cluster as they correspond to only 0.4% of the total number of disruptions. What is more, we have noticed that type 3 disruptions largely belong to NEU cluster. Therefore, we consider these observations as outliers, for which discrepancy in demand does not correspond to the disruptions' impact.

Recovery patterns - Fig. 7(b) show that demand levels for type 1 start from almost −15% and reach −2%, 5 min after the end of disruptions. Recovery patterns for type 2 is slower, starting from −10% and reaching −2,5% after 80 min. Demand levels for type 3 start from −15% and increase linearly until they reach −5%. The slow recovery patterns observed for type 2 is partly related to the cascading effect. After the end of these disruptions, type 1 disruptions might follow, which still have an influence on ridership levels. However, type 1 also shows a first rapid recovery pattern, but the demand levels then slightly decrease until they finally reach a steady state. Therefore, there is a remaining gap in the ridership levels between 60 and 80 min after the disruptions. This topic needs to be further investigated, but we can assume that it is related to capacity issues and misinformation about the state of the traffic after the disruption ended. To a lesser extent, conclusions drawn from early response analysis for the set of high-intensity disruptions are transferable to the set of medium-intensity disruptions. However, recovery patterns for medium-intensity disruption is usually shorter (between 5 and 40 min).

The temporal boundaries that are logged in the incident database are sometimes inaccurate, and AFC data provides relevant information to improve the delimitation of disruptions. There is a range of disruptions for which the impacts on ridership levels can be foreseen few minutes before an alert is registered by the operator. In this paper, we observe a discrepancy starting approximately 5 min before these disruptions. This result implies that for some disruptions, real-time information from the ground could improve the triggering of emergency procedures. In this case, anomaly detection models seem particularly useful to handle real-time management of disruptions (Tonnelier et al., 2018; Pasini et al., 2022). Similarly, in some cases few minutes are needed to recover from a disruption. We also estimate the recovery time (i.e. the time needed to retrieve reference levels) between 5 and 80 min, with a fast recovery stage between 0 and 5 min, and a long recovery stage between 5 and up to 80 min after the end of a disruption. However, these duration may vary at the network level (Malandri et al., 2018).

4.2.2. *Spatial context of disruptions: are all subway stations impacted the same way when a disruption occur?*

Part-line disruption (type 1) - Fig. 8 shows the results of spatial consistency. Type 1 has consistent patterns for sets of relatively close stations. Line A indicates a dichotomous pattern, with one hand stations from PER to MAS and, on the other hand, stations from REP to SOI having very strong spatial consistency. Line B and D indicate more complex patterns, but still some portions can be identified as spatially consistent. For instance, line B shows a strong consistency for stations between CHA and DEB, but also between OGA and JAU. Line D has strong correlation patterns between VAI and BLA, GAR and VEN and BLA and VEN, which sometimes overlap (as GAR to VEN and BLA to VEN for instance). Line C with fewer stations shows that impacts for COR are consistent with HOT and CPA, while HEN and CUI are consistent with each other. These results are due to the fact that most of the time, a closure is implemented by the operator on a relevant part of the line because of the design of the infrastructure and the cause of the disruption. First explanation is related to the track circuit: in case of power failure, only a part of the track can be concerned by the disruption. Second explanation is related to the need for a return loop to be able to operate a subway line. The portion of the track that is still in service needs access to the depot and/or side or tail tracks and/or specific switches or crossovers. Therefore, the decision to close a part of a line is very linked to the design of the infrastructure, which constrains the decision-making process.

Full-line disruption (type 2 & 3) As expected, type 2 and 3 have homogeneous patterns at the line level. Also, type 3 is more consistent than type 2. However, a larger number of type 3 disruptions (71%) belongs to the neutral cluster (NEU), which reinforces the spatial consistency of this type of disruption which mostly has a very low impact on the ridership levels. On the contrary, type 2 shows less spatial

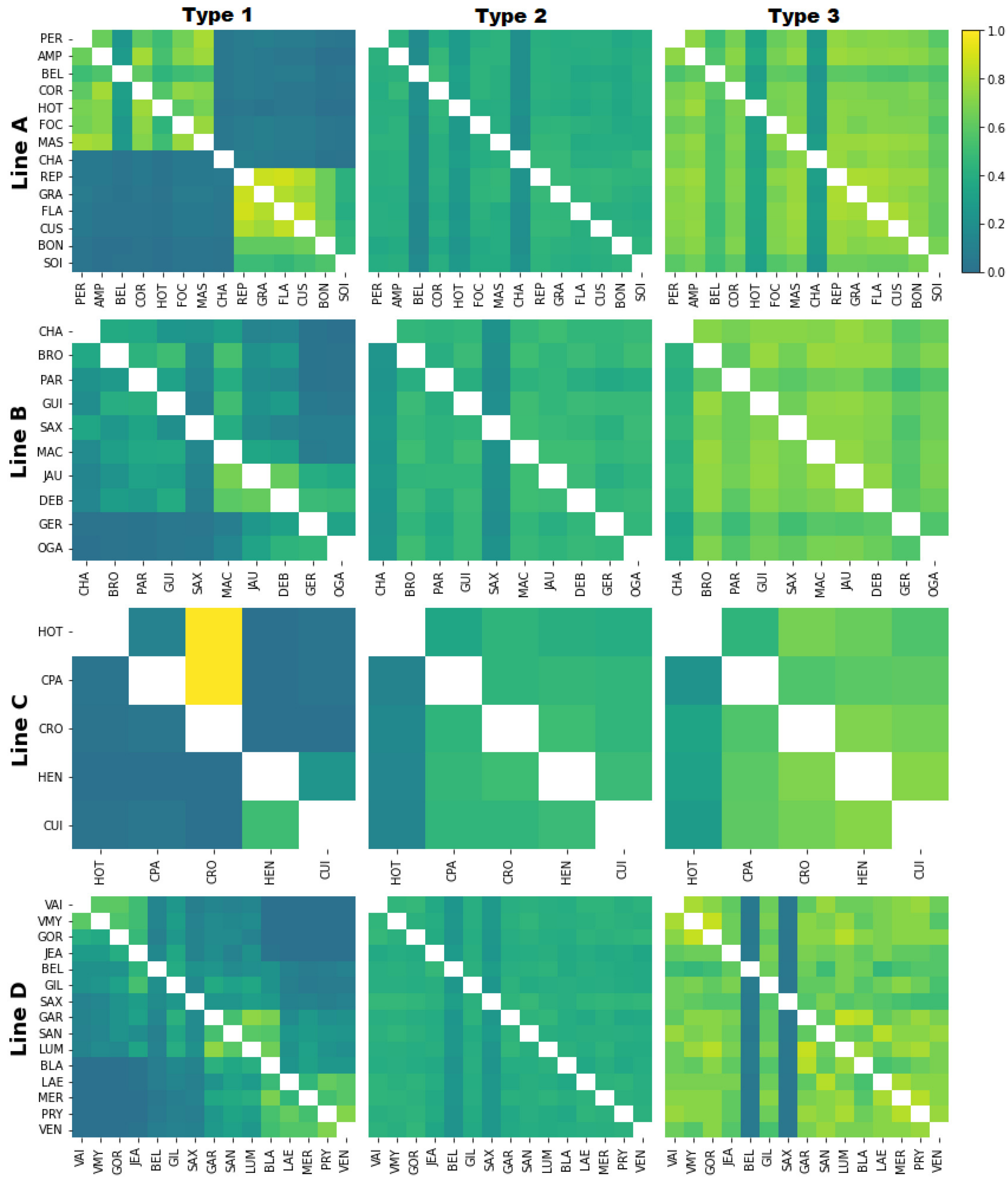


Fig. 8. Spatial consistency of disruptions - *The heatmaps show the correlation of stations in terms of disruptions impact. A value close to 1 (yellow) means that the demand for both considered stations tends to evolve in the same direction when facing a disruption. In that case, we consider that the disruption is spatially consistent, otherwise a value close to 0 (dark green) means that the demand evolves differently. In that case, the disruption is not spatially consistent.* (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

consistency than type 3, but higher intensity. This result suggests that disruption management strategies should be adapted station-wise. For instance, additional supply may be crucial to fill the gap left by the lack of service for some stations, while others can rely on the existing alternatives. Finally, results highlight low spatial consistency for transfer stations, namely, CHA, HOT, BEL and SAX, irrespective of the type of disruptions. By definition, these stations belong to 2 lines at the same time and there is no clear procedure to distinguish ridership levels from both lines. Consequently, the impacts measured for these stations are often different from the rest of the line.

Usually, disruptions are logged at the line level (Weng et al., 2014) or the station level (Louie et al., 2017; Yap and Cats, 2021). This result provides a relevant feedback on how the spatial boundaries should be considered. For analytic purpose, the location of the disruptions should be distinguished from the set of impacted stations. By default, this set is adjusted to the whole line, but when available, detailed spatial information is required. The occurrence location is useful for maintenance purpose: if disruptions frequently occurs in a specific location, there is a high probability that action is needed to avoid upcoming incidents. However, the set of impacted stations provides information on where the supply and/or demand is expected to change and consequently, where PT services need to be strengthened.

5. Conclusions

Using a sample of 3 years of subway Automatic Fare Collection (AFC) and Service Disruptions (SD) logs provided by the Public Transit (PT) operator of Lyon, this study aims to answer the following questions: *What are the impacts of service disruptions on subway demand levels? How can this information help the implementation of demand-oriented* *management strategies for subway disruptions? Is there a spatio-temporal variability regarding these impacts?*

This work uses ridership levels in the form of time series to assess the intensity of disruptions. As a result, 39% of the service disruptions observed on the PT network of Lyon between 2021 and 2023 have a negative impact on demand. We distinguish 3 levels of negative impacts: high-intensity (7%), medium-intensity (11%), and low-intensity (21%). On the temporal aspect, results show that the PT system is more vulnerable to long-lasting disruptions, occurring during peak hours. On the spatial aspect, the connection with the national railway system increases the vulnerability of PT systems to disruptions. The causes that mostly harm the demand are unattended baggage, power, and informatics failures, and incidents related to ongoing infrastructure projects. However, lines are sensitive to different risks depending on their exposure to external factors (e.g. weather) or their infrastructure's singularities (e.g. presence of crossovers), which might emphasize other causes. Also, 53% of disruptions provoke almost no impact on demand. They are characterized by frequency reduction and are mainly caused by rolling stock failure or unplanned maintenance operations.

The remaining 8% belong to a positive intensity cluster, which is typical of the limits of the model. The methodology introduced in this paper cannot fully capture the effect of all events that might happen and affect the PT network at the same time. Working with a comprehensive dataset about all kinds of events, not only service disruptions, is recommended to overcome this problem. Another limitation is the fact that the effect of disruption on PT transfer stations is poorly captured. For these stations, working with OD flows, or targeting the entry gates of the station for each validation might help improve the appraisal of disruptions' effect on demand.

The contextual analysis indicates that disruptions' effects on demand can be detected a few minutes before and after the time boundaries registered in SD logs. It also highlights the spatial variability of disruptions' impact at the line level. These insights can help to prepare management strategies (e.g. by identifying where additional services are the most needed) and to implement real-time management strategies (e.g. by building online disruption detection models).

This work is a necessary contribution to understanding the impacts of service disruptions on ridership levels. Even if a one-to-one relationship does not exist between ridership levels and the occurrence of an incident, being able to appraise how demand reacts when operators face incidents of any kind is the first step towards a better integration of such events in their management strategies. This work notably shows when, where, and why disruptions occur, with critical feedback on the data made available by PT operators.

Further works could study the opportunity to dimension emergency transit services, using the results of our work as inputs. On the supply side, the use of bus bridging strategies or micro-mobility are two interesting tracks of research. Also, other variables related to the design and the aging of PT infrastructure could be used to prevent PT operators from frequent disruptions and anticipate actions that need to be implemented during disruption scenarios. On the demand side, the implementation of disruption detection and flow redistribution models could be relevant to managing high-intensity disruptions better.

CRediT authorship contribution statement

Benjamin Cottreau: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ouassim Manout:** Writing – review & editing, Validation, Supervision. **Louafi Bouzouina:** Validation, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Benjamin Cottreau reports financial support was provided by Keolis Lyon. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This paper is part of the "Mobility in the Big Data Era" chair. The authors are thankful to KEOLIS for providing the data on PT use in Lyon.

Appendix A. Spatial consistency matrix

In order to test the consistency of the retrieved set of impacted stations $\mathbb{S}^\delta$, we build a distance matrix D between stations. The spatial impact of the disruption is estimated according to the type and the line to which it refers. Consequently, we test the coherence for each type and each subway line, leading to the calculation of 12 different distance matrices. Let us denote $c(I_{s,\delta})$ the cluster to which the impact vector $I_{s,\delta}$ belongs, and Δ the set gathering all disruptions δ. As shown in Algorithm A.1, the goal of the procedure is to check whether each disruption δ has the same impact on the resulting set of stations $\mathbb{S}^\delta$, i.e. corresponding impact vectors $I_{s,\delta}$ belong to the same cluster.

Algorithm A.1 Calculation of distance matrix D

```
D ← 0_{n,n}
for δ ∈ Δ do
    for s_i ∈ 𝕊^δ do
        for s_j ∈ 𝕊^δ do
            if c(I_{s_i,δ}) = c(I_{s_j,δ}) then
                D[s_i, s_j] ← D[s_i, s_j] + 1
            else
                D[s_i, s_j] ← D[s_i, s_j]
            end if
        end for
    end for
end for
```

Then, relative percentages are calculated by columns. A value of 1 means that the couple of stations (s_i, s_j) is fully spatially consistent, i.e. disruptions' impacts are similar to the considered couple of stations. Otherwise, a value of 0 means that the couple (s_i, s_j) is not spatially consistent, i.e. disruptions' impacts are totally different for the considered couple of stations.

Appendix B. Number of clusters

The number of clusters resulting from the Gaussian Mixture Model (GMM) has been set according to two parameters, as shown in Fig. B.1. First, R^2 is a quality measurement that represents "a proportion of total variance explained by a clustering partition" (Loperfido and Tarpey, 2018). Second, a minimum size threshold is implemented to constrain the resulting clusters to have a minimum number of observations. This choice is motivated by the fact that clusters on the edges (e.g. corresponding to NEG+ or POS) tend to split when increasing the number of clusters. Therefore, R^2 increases as the cluster centers get closer to the input data, but the new partitioning does not provide relevant information on the intensity level (which remains very negative or positive). The minimum size threshold is set to 5% in this study. Therefore, the number of clusters used in this work is set to 5 and $R^2 = 0.69$.

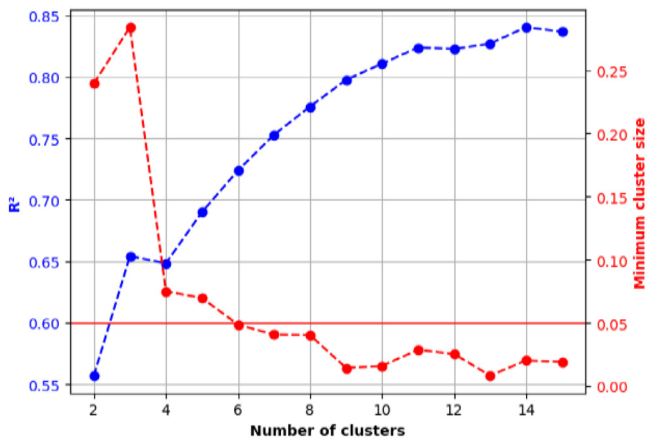


Fig. B.1. Performance (blue curve) and minimum cluster size (red curve) according to the number of clusters in GMM - *The horizontal plain red line indicates the 5% threshold for the minimum cluster size. The chosen number of clusters is 5* . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Data availability

The authors do not have permission to share data.

References

Bešinović, N., 2020. Resilience in railway transport systems: a literature review and research agenda. Transp. Rev. 40 (4), 457–478. http://dx.doi.org/10.1080/01441647.2020.1728419, URL https://www.tandfonline.com/doi/full/10.1080/01441647.2020.1728419.

Bešinović, N., Ferrari Nassar, R., Szymula, C., 2022. Resilience assessment of railway networks: Combining infrastructure restoration and transport management. Reliab. Eng. Syst. Saf. 224, 108538. http://dx.doi.org/10.1016/j.ress.2022.108538, URL https://linkinghub.elsevier.com/retrieve/pii/S0951832022001909.

Brakewood, C., Watkins, K., 2019. A literature review of the passenger benefits of real-time transit information. Transp. Rev. 39 (3), 327–356. http://dx.doi.org/10.1080/01441647.2018.1472147, URL https://www.tandfonline.com/doi/full/10.1080/01441647.2018.1472147.

Cats, O., 2016. The robustness value of public transport development plans. J. Transp. Geogr. 51, 236–246. http://dx.doi.org/10.1016/j.jtrangeo.2016.01.011, URL https://linkinghub.elsevier.com/retrieve/pii/S0966692316000120.

Cottreau, B., Adraoui, A., Manout, O., Bouzouina, L., 2023. Spatio-temporal patterns of the impact of COVID-19 on public transit: An exploratory analysis from Lyon, France. Reg. Sci. Policy & Pr. 15 (8), 1702–1721. http://dx.doi.org/10.1111/rsp3.12718, URL https://rsaiconnect.onlinelibrary.wiley.com/doi/10.1111/rsp3.12718.

De Nailly, P., Côme, E., Samé, A., Oukhellou, L., Ferriere, J., Merad-Boudia, Y., 2022. What can we learn from 9 years of ticketing data at a major transport hub? A structural time series decomposition. Transp. A: Transp. Sci. 18 (3), 1445–1469. http://dx.doi.org/10.1080/23249935.2021.1948626, URL https://www.tandfonline.com/doi/full/10.1080/23249935.2021.1948626.

Deschaintres, E., Morency, C., Trépanier, M., 2022. Assessing the Impacts of the COVID-19 Pandemic on Subway Ridership and on the Interactions with other Transportation Modes. In Review, http://dx.doi.org/10.21203/rs.3.rs-2315989/v1, URL https://www.researchsquare.com/article/rs-2315989/v1.

Drabicki, A., Kucharski, R., Cats, O., Szarata, A., 2021. Modelling the effects of real-time crowding information in urban public transport systems. Transp. A: Transp. Sci. 17 (4), 675–713. http://dx.doi.org/10.1080/23249935.2020.1809547, URL https://www.tandfonline.com/doi/full/10.1080/23249935.2020.1809547.

El Zein, A., Beziat, A., Pochet, P., Klein, O., Vincent, S., 2022. What drives the changes in public transport use in the context of the COVID-19 pandemic? Highlights from Lyon metropolitan area. Reg. Sci. Policy & Pr. 14, 122–142. http://dx.doi.org/10.1111/rsp3.12519.

Eltved, M., Breyer, N., Ingvardson, J.B., Nielsen, O.A., 2021. Impacts of long-term service disruptions on passenger travel behaviour: A smart card analysis from the Greater Copenhagen area. Transp. Res. Part C: Emerg. Technol. 131, 103198. http://dx.doi.org/10.1016/j.trc.2021.103198, URL https://linkinghub.elsevier.com/retrieve/pii/S0968090X21002138.

Fadeev, A., Alhusseini, S., Belova, E., 2018. Monitoring public transport demand using data from automated fare collection system. In: Proceedings of the International Conference "Aviamechanical Engineering and Transport" (AVENT 2018). Atlantis Press, Irkutsk, Russia, http://dx.doi.org/10.2991/avent-18.2018.2, URL http://www.atlantis-press.com/php/paper-details.php?id=25901591.

Gao, H., Xu, J., Li, S., Xu, L., 2020. Forecast of passenger flow under the interruption of urban rail transit operation. In: Qin, Y., Jia, L., Liu, B., Liu, Z., Diao, L., An, M. (Eds.), Proceedings of the 4th International Conference on Electrical and Information Technologies for Rail Transportation (EITRT) 2019. Springer, Singapore, pp. 283–291. http://dx.doi.org/10.1007/978-981-15-2866-8_27.

He, M., Pathak, S., Muaz, U., Zhou, J., Saini, S., Malinchik, S., Sobolevsky, S., 2019. Pattern and anomaly detection in urban temporal networks. http://dx.doi.org/10.48550/ARXIV.1912.01960, arXiv URL https://arxiv.org/abs/1912.01960, Version Number: 1.

Jenelius, E., Cebecauer, M., 2020. Impacts of COVID-19 on public transport ridership in Sweden: Analysis of ticket validations, sales and passenger counts. Transp. Res. Interdiscip. Perspect. 8, 100242. http://dx.doi.org/10.1016/j.trip.2020.100242, URL https://linkinghub.elsevier.com/retrieve/pii/S2590198220301536.

Kellermann, R., Sivizaca Conde, D., Rößler, D., Kliewer, N., Dienel, H.-L., 2022. Mobility in pandemic times: Exploring changes and long-term effects of COVID-19 on urban mobility behavior. Transp. Res. Interdiscip. Perspect. 15, 100668. http://dx.doi.org/10.1016/j.trip.2022.100668, URL https://linkinghub.elsevier.com/retrieve/pii/S2590198222001282.

Leng, N., Corman, F., 2020. How the issue time of information affects passengers in public transport disruptions: an agent-based simulation approach. Procedia Comput. Sci. 170, 382–389. http://dx.doi.org/10.1016/j.procs.2020.03.068, URL https://linkinghub.elsevier.com/retrieve/pii/S1877050920305056.

Liu, H.J., Love, P.E., Zhao, J., Lemckert, C., Muldoon-Smith, K., 2021a. Transport infrastructure asset resilience: Managing government capabilities. Transp. Res. Part D: Transp. Environ. 100, 103072. http://dx.doi.org/10.1016/j.trd.2021.103072, URL https://linkinghub.elsevier.com/retrieve/pii/S1361920921003692.

Liu, T., Ma, Z., Koutsopoulos, H.N., 2021b. Unplanned disruption analysis in urban railway systems using smart card data. Urban Rail Transit 7 (3), 177–190. http://dx.doi.org/10.1007/s40864-021-00150-x, URL https://link.springer.com/10.1007/s40864-021-00150-x.

Liu, L., Porr, A., Miller, H.J., 2024. Measuring the impacts of disruptions on public transit accessibility and reliability. J. Transp. Geogr. 114, 103769. http://dx.doi.org/10.1016/j.jtrangeo.2023.103769, URL https://linkinghub.elsevier.com/retrieve/pii/S0966692323002417.

Liyanage, S., Abduljabbar, R., Dia, H., Tsai, P.-W., 2022. AI-based neural network models for bus passenger demand forecasting using smart card data. J. Urban Manag. 11 (3), 365–380. http://dx.doi.org/10.1016/j.jum.2022.05.002, URL https://linkinghub.elsevier.com/retrieve/pii/S2226585622000280.

Loperfido, N., Tarpey, T., 2018. Some remarks on the R^2 for clustering. Stat. Anal. Data Mining: ASA Data Sci. J. 11 (3), 135–148. http://dx.doi.org/10.1002/sam.11378, URL https://onlinelibrary.wiley.com/doi/10.1002/sam.11378.

Louie, J., Shalaby, A., Habib, K.N., 2017. Modelling the impact of causal and non-causal factors on disruption duration for Toronto's subway system: An exploratory investigation using hazard modelling. Accid. Anal. Prev. 98, 232–240. http://dx.doi.org/10.1016/j.aap.2016.10.008, URL https://linkinghub.elsevier.com/retrieve/pii/S0001457516303694.

Malandri, C., Fonzone, A., Cats, O., 2018. Recovery time and propagation effects of passenger transport disruptions. Phys. A 505, 7–17. http://dx.doi.org/10.1016/j.physa.2018.03.028, URL https://linkinghub.elsevier.com/retrieve/pii/S0378437118303480.

Malucelli, F., Tresoldi, E., 2019. Delay and disruption management in local public transportation via real-time vehicle and crew re-scheduling: a case study. Public Transp. 11 (1), 1–25. http://dx.doi.org/10.1007/s12469-019-00196-y, URL http://link.springer.com/10.1007/s12469-019-00196-y.

Manout, O., Bouzouina, L., Kourtit, K., Nijkamp, P., 2023. On the bumpy road to recovery: resilience of public transport ridership during COVID-19 in 15 European cities. Lett. Spat. Resour. Sci. 16 (1), 1–15. http://dx.doi.org/10.1007/s12076-023-00338-8.

Marra, A.D., Corman, F., 2019. From delay to disruption: The impact of service degradation on public transport network. http://dx.doi.org/10.3929/ETHZ-B-000368811, p. 8, URL http://hdl.handle.net/20.500.11850/368811, Artwork Size: 8 p. Medium: application/pdf Publisher: [object Object].

Mo, B., Koutsopoulos, H.N., Shen, Z.-J.M., Zhao, J., 2023. Robust path recommendations during public transit disruptions under demand uncertainty. Transp. Res. Part B: Methodol. 169, 82–107. http://dx.doi.org/10.1016/j.trb.2023.02.004, URL https://linkinghub.elsevier.com/retrieve/pii/S0191261523000176.

Mo, B., Von Franque, M.Y., Koutsopoulos, H.N., Attanucci, J.P., Zhao, J., 2022. Impact of unplanned long-term service disruptions on urban public transit systems. IEEE Open J. Intell. Transp. Syst. 3, 551–569. http://dx.doi.org/10.1109/OJITS.2022.3199108, URL https://ieeexplore.ieee.org/document/9858090/.

Noursalehi, P., Koutsopoulos, H.N., 2016. Real-time predictive analytics for improving public transportation systems' resilience. http://dx.doi.org/10.48550/ARXIV.1609.09785, URL https://arxiv.org/abs/1609.09785, Publisher: [object Object] Version Number: 1.

Noursalehi, P., Koutsopoulos, H.N., Zhao, J., 2018. Real time transit demand prediction capturing station interactions and impact of special events. Transp. Res. Part C: Emerg. Technol. 97, 277–300. http://dx.doi.org/10.1016/j.trc.2018.10.023, URL https://linkinghub.elsevier.com/retrieve/pii/S0968090X18301797.

Pasini, K., 2021. Forecast and Anomaly Detection on Time Series with Dynamic Context : Application to the Mining of Transit Ridership Data (These de doctorat). Université Gustave Eiffel, URL https://theses.fr/2021UEFL2016.

Pasini, K., Khouadjia, M., Samé, A., Trépanier, M., Oukhellou, L., 2022. Contextual anomaly detection on time series: a case study of metro ridership analysis. Neural Comput. Appl. 34 (2), 1483–1507. http://dx.doi.org/10.1007/s00521-021-06455-z, URL https://link.springer.com/10.1007/s00521-021-06455-z.

Paulsen, M., Rasmussen, T.K.r., Nielsen, O.A., 2021. Impacts of real-time information levels in public transport: A large-scale case study using an adaptive passenger path choice model. Transp. Res. Part A: Policy Pr. 148, 155–182. http://dx.doi.org/10.1016/j.tra.2021.03.011, URL https://linkinghub.elsevier.com/retrieve/pii/S0965856421000665.

Pelletier, M.-P., Trépanier, M., Morency, C., 2011. Smart card data use in public transit: A literature review. Transp. Res. Part C: Emerg. Technol. 19 (4), 557–568. http://dx.doi.org/10.1016/j.trc.2010.12.003, URL https://linkinghub.elsevier.com/retrieve/pii/S0968090X1000166X.

Pender, B., Currie, G., Delbosc, A., Shiwakoti, N., 2013. Disruption recovery in passenger railways: International survey. Transp. Res. Record: J. Transp. Res. Board 2353 (1), 22–32. http://dx.doi.org/10.3141/2353-03, URL http://journals.sagepub.com/doi/10.3141/2353-03.

Plosser, C.I., 1978. A time series analysis of seasonality in econometric models. In: Seasonal Analysis of Economic Time Series. In: NBER Chapters, National Bureau of Economic Research, Inc, pp. 365–408, URL https://ideas.repec.org/h/nbr/nberch/4331.html.

Rahimi, M., Ghandeharioun, Z., Kouvelas, A., Corman, F., 2021. Multi-modal management actions for public transport disruptions: an agent-based simulation. In: 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). IEEE, Heraklion, Greece, pp. 1–6. http://dx.doi.org/10.1109/MT-ITS49943.2021.9529305, URL https://ieeexplore.ieee.org/document/9529305/.

Rahimi, E., Shamshiripour, A., Shabanpour, R., Mohammadian, A.K., Auld, J., 2020. Analysis of transit users' response behavior in case of unplanned service disruptions. Transp. Res. Record: J. Transp. Res. Board 2674 (3), 258–271. http://dx.doi.org/10.1177/0361198120911921, URL http://journals.sagepub.com/doi/10.1177/0361198120911921.

Rodríguez-Núñez, E., García-Palomares, J.C., 2014. Measuring the vulnerability of public transport networks. J. Transp. Geogr. 35, 50–63. http://dx.doi.org/10.1016/j.jtrangeo.2014.01.008, URL https://linkinghub.elsevier.com/retrieve/pii/S0966692314000180.

Siegrist, M.R., Büchel, B., Corman, F., 2022. Evaluating the effects of disruptions on the behavior of travelers in a multimodal network utilizing agent-based simulation. Transp. Res. Record: J. Transp. Res. Board 2676 (11), 436–445. http://dx.doi.org/10.1177/03611981221094005, URL http://journals.sagepub.com/doi/10.1177/03611981221094005.

Sun, H., Wu, J., Wu, L., Yan, X., Gao, Z., 2016. Estimating the influence of common disruptions on urban rail transit networks. Transp. Res. Part A: Policy Pr. 94, 62–75. http://dx.doi.org/10.1016/j.tra.2016.09.006, URL https://linkinghub.elsevier.com/retrieve/pii/S0965856416303913.

Tao, Z., Tang, J., 2019. Real-time estimation of urban rail transit passenger flow status based on multi-source data. J. Phys.: Conf. Ser. 1187 (5), 052070. http://dx.doi.org/10.1088/1742-6596/1187/5/052070, URL https://iopscience.iop.org/article/10.1088/1742-6596/1187/5/052070.

Tonnelier, E., Baskiotis, N., Guigue, V., Gallinari, P., 2018. Anomaly detection in smart card logs and distant evaluation with Twitter: a robust framework. Neurocomputing 298, 109–121. http://dx.doi.org/10.1016/j.neucom.2017.12.067, URL https://linkinghub.elsevier.com/retrieve/pii/S0925231218302170.

Toque, F., Come, E., El Mahrsi, M.K., Oukhellou, L., 2016. Forecasting dynamic public transport origin-destination matrices with long-short term memoryrecurrent neural networks. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE, Rio de Janeiro, Brazil, pp. 1071–1076. http://dx.doi.org/10.1109/ITSC.2016.7795689, URL http://ieeexplore.ieee.org/document/7795689/.

Wang, Z., Ma, W., Chan, A., 2020. Exploring the relationships between the topological characteristics of subway networks and service disruption impact. Sustainability 12 (10), 3960. http://dx.doi.org/10.3390/su12103960, URL https://www.mdpi.com/2071-1050/12/10/3960.

Weng, J., Zheng, Y., Yan, X., Meng, Q., 2014. Development of a subway operation incident delay model using accelerated failure time approaches. Accid. Anal. Prev. 73, 12–19. http://dx.doi.org/10.1016/j.aap.2014.07.029, URL https://linkinghub.elsevier.com/retrieve/pii/S0001457514002322.

Yap, M., Cats, O., 2021. Predicting disruptions and their passenger delay impacts for public transport stops. Transportation 48 (4), 1703–1731. http://dx.doi.org/10.1007/s11116-020-10109-9, URL https://link.springer.com/10.1007/s11116-020-10109-9.

Yap, M., Cats, O., 2022. Analysis and prediction of ridership impacts during planned public transport disruptions. J. Public Transp. 24, 100036. http://dx.doi.org/10.1016/j.jpubtr.2022.100036, URL https://linkinghub.elsevier.com/retrieve/pii/S1077291X22017362.

Zargari, F., Aminpour, N., Ahmadian, M.A., Samimi, A., Saidi, S., 2022. Impact of mobility on COVID-19 spread – A time series analysis. Transp. Res. Interdiscip. Perspect. 13, 100567. http://dx.doi.org/10.1016/j.trip.2022.100567, URL https://linkinghub.elsevier.com/retrieve/pii/S2590198222000306.

Zhang, S., Lo, H.K., Ng, K.F., Chen, G., 2021. Metro system disruption management and substitute bus service: a systematic review and future directions. Transp. Rev. 41 (2), 230–251. http://dx.doi.org/10.1080/01441647.2020.1834468, URL https://www.tandfonline.com/doi/full/10.1080/01441647.2020.1834468.

Zhang, J., Shen, D., Tu, L., Zhang, F., Xu, C., Wang, Y., Tian, C., Li, X., Huang, B., Li, Z., 2017. A real-time passenger flow estimation and prediction method for urban bus transit systems. IEEE Trans. Intell. Transp. Syst. 18 (11), 3168–3178. http://dx.doi.org/10.1109/TITS.2017.2686877, URL http://ieeexplore.ieee.org/document/7898469/.