

D'autres messages suivront.

Modélisation des
interruptions de service
du métro de Montréal.

Projet de recherche réalisé par
Samuel Mehenni

Supervisé par
Prof. Richard G. Shearmur

Projet présenté en vue de
l'obtention du diplôme de
Master of Urban Planning

School of Urban Planning
McGill University
Août 2022

REMERCIEMENTS

Ce projet non sans détour n'aurait pas été possible sans l'appui de nombreuses personnes. Je tiens d'abord à remercier mon superviseur de recherche Richard Shearmur pour son soutien continu, ses commentaires judicieux et sa confiance manifeste. Richard a plongé dans les données avec moi et sa rigueur a renforcé les résultats de cette analyse.

Je veux aussi remercier Vincent Chapdelaine, mon ami et fidèle partenaire de responsabilisation. Plusieurs jalons de ce projet de recherche faisaient partie de mes objectifs bimensuels de responsabilisation des six derniers mois. Nos rencontres m'ont forcé à mettre de l'ordre dans mon plan de rédaction tout en maintenant une hygiène buccale quotidienne irréprochable.

Je remercie également mes amis et collègues de la maîtrise en urbanisme, avec qui j'ai eu beaucoup de plaisir et de nombreuses discussions passionnantes ces dernières années dans un contexte covidien singulier. Merci aussi à ma collègue Marie-Christine Desharnais d'avoir passé ce document au peigne fin à la recherche de coquilles.

Un merci spécial à ma fiancée, mon amoureuse, Mary-Ann Wiggill, de m'avoir soutenu dans la rédaction de ce projet et qui est devenue malgré elle une experte des pannes de service.

Je souhaite remercier le soutien financier du Conseil de recherches en sciences humaines (CRSH) via la bourse Joseph-Armand-Bombardier, puis la bourse de Norbert Schoenauer et David Farley de l'université McGill.

Finalement, chapeau à la STM pour sa politique de transparence des données. Et coucou à tous les usagers du métro: en cas de panne, merci d'être patients et cordiaux. N'oubliez pas: d'autres messages suivront!

RÉSUMÉ

Des centaines de milliers d'usagers du métro de Montréal sont confrontés chaque année à des interruptions de service non planifiées. Ces perturbations sont la cause d'itinéraires prolongés, de frustrations et d'incertitudes pour la clientèle. La perturbation fréquente du service est en fait la raison principale de l'insatisfaction des usagers du métro (STM, 2022). Une communication efficace et continue est un facteur déterminant de la satisfaction de la clientèle en cas de perturbations.

Ce projet de recherche a un objectif double; d'abord de présenter un portrait détaillé des interruptions de service, puis d'évaluer s'il est possible de prédire rigoureusement la durée des arrêts avec les informations disponibles en début d'interruption. Les jeux de données utilisés pour cette analyse proviennent de la STM par l'entremise du portail de données ouvertes ainsi qu'une demande d'accès à l'information. La période d'étude s'étale du 1er janvier 2019 au 31 octobre 2021.

L'analyse préliminaire des données contextuelles révèle certaines tendances peu étudiées du métro montréalais. On note une diminution du nombre total d'incidents entre 2014 et 2021 malgré une hausse du nombre de kilomètres parcourus. Les arrêts de service causés par la clientèle et le matériel roulant représentent ensemble une grande majorité du nombre total d'incidents ainsi que de la durée cumulée des interruptions pour la période d'étude. Quant à l'emplacement, 84% des 2 528 incidents entre 2019 et 2021 ont affecté les lignes orange et verte. Finalement, les données révèlent aussi que la majorité des interruptions (66%) ont été résolues en moins de 10 minutes, soutenant l'idée que les usagers devraient s'en tenir à leur itinéraire habituel dans la majorité des cas et attendre la fin de la panne.

Les variables étudiées afin de calculer les prédictions de durées incluent la saison, le jour de la semaine, la période, la ligne de métro ainsi que la cause associée. La première régression linéaire multiple par la méthode des moindres carrés ordinaire révèle que les variables sélectionnées expliquent à peine 6% (R^2 ajusté de 0.0609) de la variance des logarithmes des durées des interruptions. Les graphiques de valeurs résiduelles mettent d'ailleurs en évidence la problématique associée aux incidents d'émanation, soit la cause avec les durées observées et prédites les plus longues. On constate ainsi que les informations spatio-

temporelles disponibles en début d'interruption sont de faibles prédicteurs de la durée des incidents, et que ce sont les motifs qui déterminent en grande partie la rapidité de résolution des pannes.

Une régression binomiale utilisant deux sous-ensembles corrobore les premiers résultats. Les deux sous-ensembles constitués — avec les observations de plus ou moins 20 minutes — révèlent encore une fois que certaines causes ont tendance à être beaucoup plus longues que la moyenne. Un incident pour cause d'émanation a par exemple à 29 fois plus de chance de durer plus de 20 minutes qu'un incident associé au matériel roulant.

Bien que ces résultats permettent d'avoir une idée générale des tendances liées aux interruptions de service, il s'avère difficile de prédire avec un intervalle de confiance réduit la durée d'une panne en début d'interruption, élément qui pourrait être bénéfique pour les passagers. Néanmoins, les acteurs du transport en commun — décideurs, opérateurs et équipes de communication — peuvent adapter leurs processus afin d'améliorer l'expérience client. De nouveaux indicateurs de performance relatifs à l'impact sur les usagers pourraient contribuer à une meilleure priorisation des équipes de réponse en fonction des causes déclarées. Une hausse de la fréquence des communications et une diffusion expéditive multiplateforme viendraient aussi aider les usagers qui le peuvent à opter pour des itinéraires alternatifs et ainsi éviter d'éventuels délais prolongés.

Ce projet de recherche souhaite contribuer à la littérature limitée concernant les interruptions de service, ainsi que soutenir la réflexion relative à l'amélioration de l'expérience client afin de favoriser l'adoption du transport collectif.

TABLE DES MATIÈRES

Remerciements	i
Résumé	ii
Table des matières	iv
1. Introduction	2
2. Revue de littérature	5
3. Stratégie de recherche	9
3.1 Objectifs	9
3.2 Données	9
3.3 Méthodologie	11
3.4 Limites	13
4. Mise en contexte	15
4.1 Lignes et achalandage	15
4.2 Performance et probabilité	17
4.3 Portrait et tendances des interruptions de service	20
5. Analyse	25
5.1 Précisions	25
5.2 Statistiques descriptives	27
5.3 Résultats et discussion	29
6. Recommandations	36
7. Conclusion	40
Bibliographie	42
Autres références	44



Panne de service
Photo: STM, 2017

1. INTRODUCTION

Le métro de Montréal est l'un des plus importants réseaux de transport lourds en Amérique du Nord de par son achalandage quotidien. La Société de transport de Montréal (STM), qui exploite le réseau d'autobus et de métro de l'île de Montréal, investit des ressources financières et opérationnelles importantes afin d'améliorer la régularité du service et assurer une expérience client de qualité. Néanmoins, les interruptions de service non planifiées sont inévitables, affectant ainsi le quotidien de centaines de milliers de passagers chaque année.

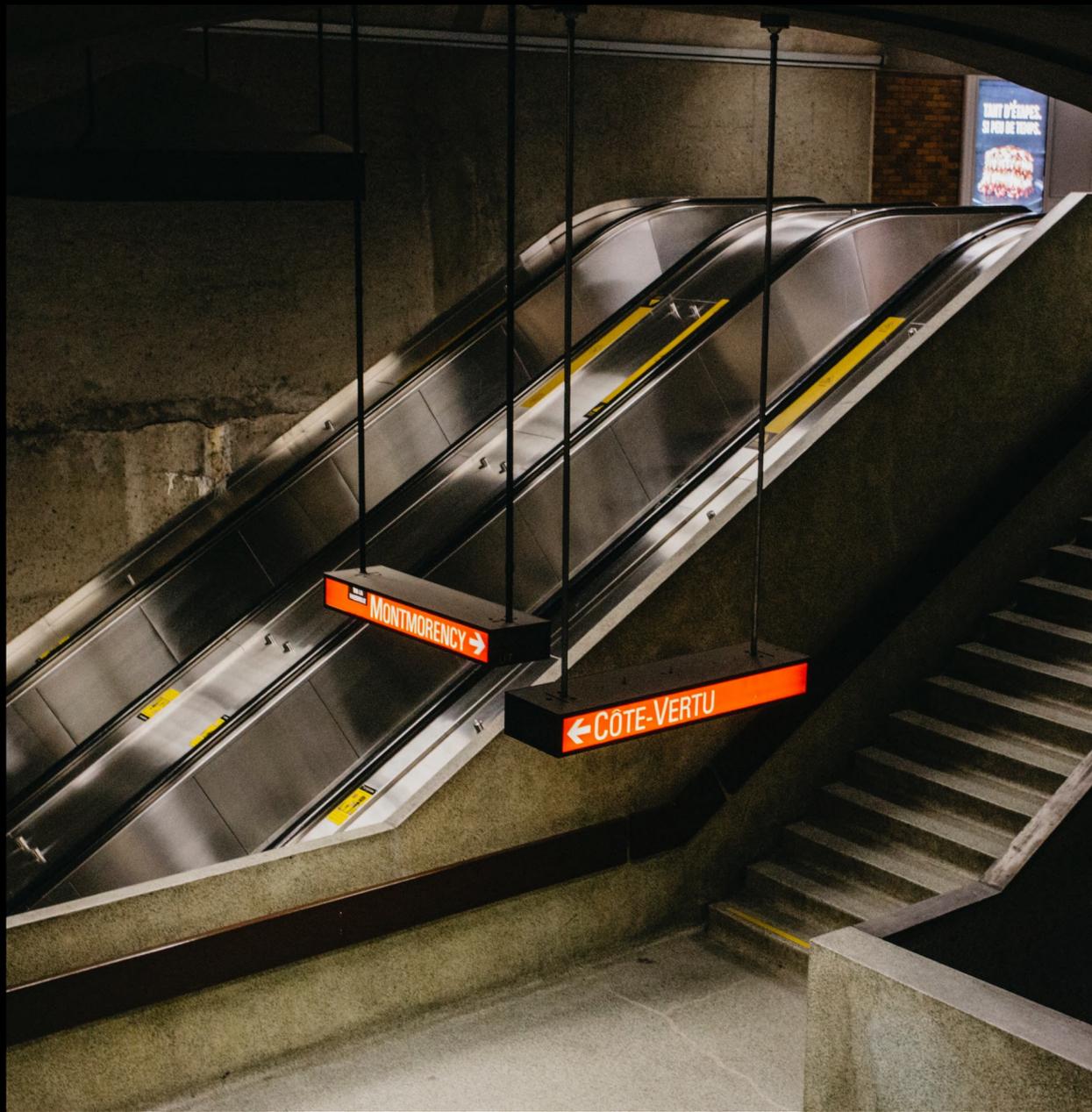
Les habitués du métro ont tous déjà entendu le fameux message lors d'une panne de service, une variante du: "Attention, un incident cause un ralentissement de service sur la ligne orange, d'autres messages suivront.". À peine les premiers mots relayés par les hauts parleurs, tous et toutes à bord se posent les mêmes questions en rafale: quelle heure est-il? Combien de temps avant la reprise du service? Devrais-je attendre quelques minutes ou opter pour un itinéraire alternatif? Quelles sont les autres options pour me rendre à destination le plus rapidement possible? La tension sur le quai ou à bord du train est palpable.

Dépendamment de la destination du voyageur, les options peuvent être limitées, et la variance de durée du déplacement peut être considérable. Les informations disponibles en début d'interruption sont souvent maigres. La STM fournit les premières informations sur l'interruption et l'heure de reprise prévue seulement 10 ou 15 minutes suite à la première annonce. Or, ces précieuses minutes peuvent faire une différence importante pour les clients de la STM, qui ont des obligations et dépendent du réseau de transport en commun pour se rendre à destination.

Mais que connaît-on réellement des interruptions de service du métro? Les utilisateurs du transport en commun sont-ils victimes d'un biais de négativité? Comment le réseau de métro de Montréal se compare-t-il avec les systèmes des grandes villes du monde? Et est-il possible d'estimer plus rapidement l'heure de reprise du service, afin de supporter les usagers dans leur prise de décision?

Cette recherche tentera de répondre à ces questions. La mise en contexte dresse d'abord un portrait de la littérature à propos des interruptions de service et présente les données du réseau de métro de Montréal pour la période à l'étude. Puis, l'analyse tentera de déterminer s'il est possible d'estimer la durée approximative d'une interruption de service à l'aide d'informations connues au moment de l'interruption, telles que l'heure et la journée de l'arrêt de service, la ligne, ainsi que la cause initiale.

L'information passager est une composante centrale de la stratégie d'expérience client des agences de transport. Cette étude permettra notamment d'évaluer si des produits numériques — applications web et mobiles — pourraient utiliser les données ouvertes disponibles afin d'améliorer les suggestions d'itinéraires de transport en temps réel lors d'arrêts de service.



Métro Bonaventure
Photo: JP Valéry

2. REVUE DE LITTÉRATURE

Les agences de transport à travers le monde investissent des ressources substantielles dans l'optimisation de la performance de leur réseau. La STM fait partie du Community of Metros Benchmarking Group (COMET), une organisation qui regroupe 42 systèmes de métros dans 39 villes d'Amérique, d'Europe et d'Asie-Pacifique (COMET, 2022). Les membres de cette organisation partagent notamment une liste de 30 indicateurs de performance (KPIs) à mesurer dans différents domaines tels que la croissance, l'innovation, les finances, la clientèle, les opérations, la sécurité ainsi que l'environnement. Grâce à cette initiative, certaines données standardisées sont récoltées par l'ensemble des membres, puis analysées de manière indépendante par des chercheurs associés au Transport Strategy Centre de l'Imperial College London, en charge des publications du groupe.

Malgré les efforts de standardisation des données, la définition d'une interruption de service peut varier selon l'agence de transport. La STM définit une interruption de service comme un arrêt complet de tous les trains sur une section du réseau, alors qu'un ralentissement de service signifie que les trains sont encore en mouvement, mais à une fréquence moins élevée (STM, 2022). Or, comme nous allons le voir plus tard dans cette étude, la distinction entre une interruption et un ralentissement de service n'est pas toujours aussi évidente. Dans la grande majorité des cas, les arrêts de service ne sont pas planifiés et impliquent plusieurs intervenants — opérateurs, contrôleur de la salle de contrôle du métro, chef d'opération — afin de déterminer la nature de l'incident et mettre en place le protocole prévu.

Une étude de Barron et al. publiée en 2013 propose une approche de gestion axée sur les passagers afin de mesurer l'impact des retards des trains. Seulement deux des 22 agences partenaires du COMET sondées étaient en mesure de fournir des informations détaillées quant aux causes associées aux interruptions de service, un problème qui persiste dans la littérature (Barron et al., 2013). Pour l'agence identifiée comme "métro A" — opérant un des réseaux les plus fiables du COMET selon les auteurs — les causes détaillées incluent le matériel roulant, la signalisation, les équipements, les incidents causés par le personnel, puis ceux causés par la clientèle.

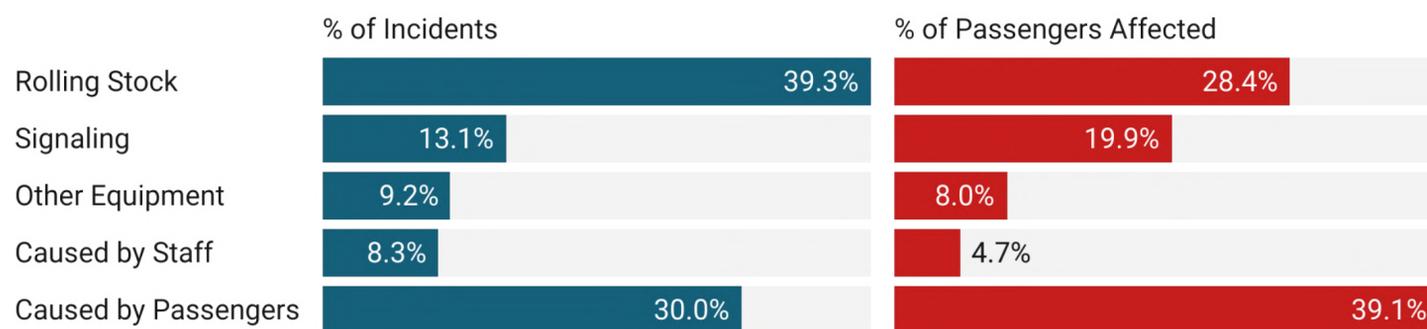


Figure 1. Nombre d'incidents et de passagers incommodés pour le "Métro A"
Source: Barron & al., 2013.

Bien que les résultats de l'analyse montrent un nombre d'incidents supérieur pour le matériel roulant, les incidents causés par la clientèle sont ceux qui incommodent le plus les passagers du métro par année, d'où l'importance d'adopter une stratégie de résolution centrée sur l'utilisateur.

Parmi les facteurs identifiés qui contribuent à réduire le nombre d'incidents, l'automatisation des trains du réseau de métro se place au sommet des recommandations, avec une réduction moyenne de 33% des incidents (Melo et al., 2011). Une augmentation de la capacité pratique de 10% est aussi associée à une réduction des incidents d'environ 9,3% pour les systèmes étudiés. À l'inverse, plusieurs facteurs contribuent à augmenter le nombre d'incidents. Une croissance de 10% du nombre de trajets des passagers est associée à une augmentation de 3% des incidents. Quant à l'ancienneté du réseau, une augmentation de 10% de l'âge d'une ligne entraîne une hausse moyenne de 2,6% du nombre d'incidents (Melo et al., 2011). Ces observations démontrent l'importance des investissements dans la maintenance et la mise à jour des infrastructures des systèmes de métro.

Concernant la modélisation des interruptions de service, une analyse des pannes à Toronto révèle que la cause des incidents est un facteur déterminant de la durée des interruptions. Les incidents qui ont un impact simultané sur la sécurité et les opérations sont généralement plus longs que les causes d'incidents qui ont un impact sur la sécurité ou les opérations uniquement (Louie, 2016). L'analyse révèle d'ailleurs que le lieu de l'incident peut avoir un impact sur la durée des interruptions; les incidents ayant lieu à des stations de correspondances sont résolus plus rapidement qu'aux stations secondaires. En moyenne, les incidents aux heures de pointe ont pour leur part la même durée que les incidents aux heures creuses.

Il est aussi pertinent d'étudier les prises de décision des voyageurs en réaction aux incidents de métro. Des analyses démontrent que des facteurs comme les caractéristiques

sociodémographiques, l'environnement urbain, l'itinéraire en cours et les préférences personnelles sont déterminants du comportement des passagers lors d'interruptions de service non planifiées (Rahimi, 2021). On peut estimer que pour des délais relativement courts de 5 à 10 minutes, une grande majorité des passagers restent à bord du train. Néanmoins, confrontés à des délais prolongés, plusieurs passagers considèrent des trajets alternatifs. Une analyse du métro à Montréal révèle une augmentation de l'intensité d'usage des autres modes de transport lors d'interruptions d'une durée de plus de 20 minutes (Lepage, 2019).

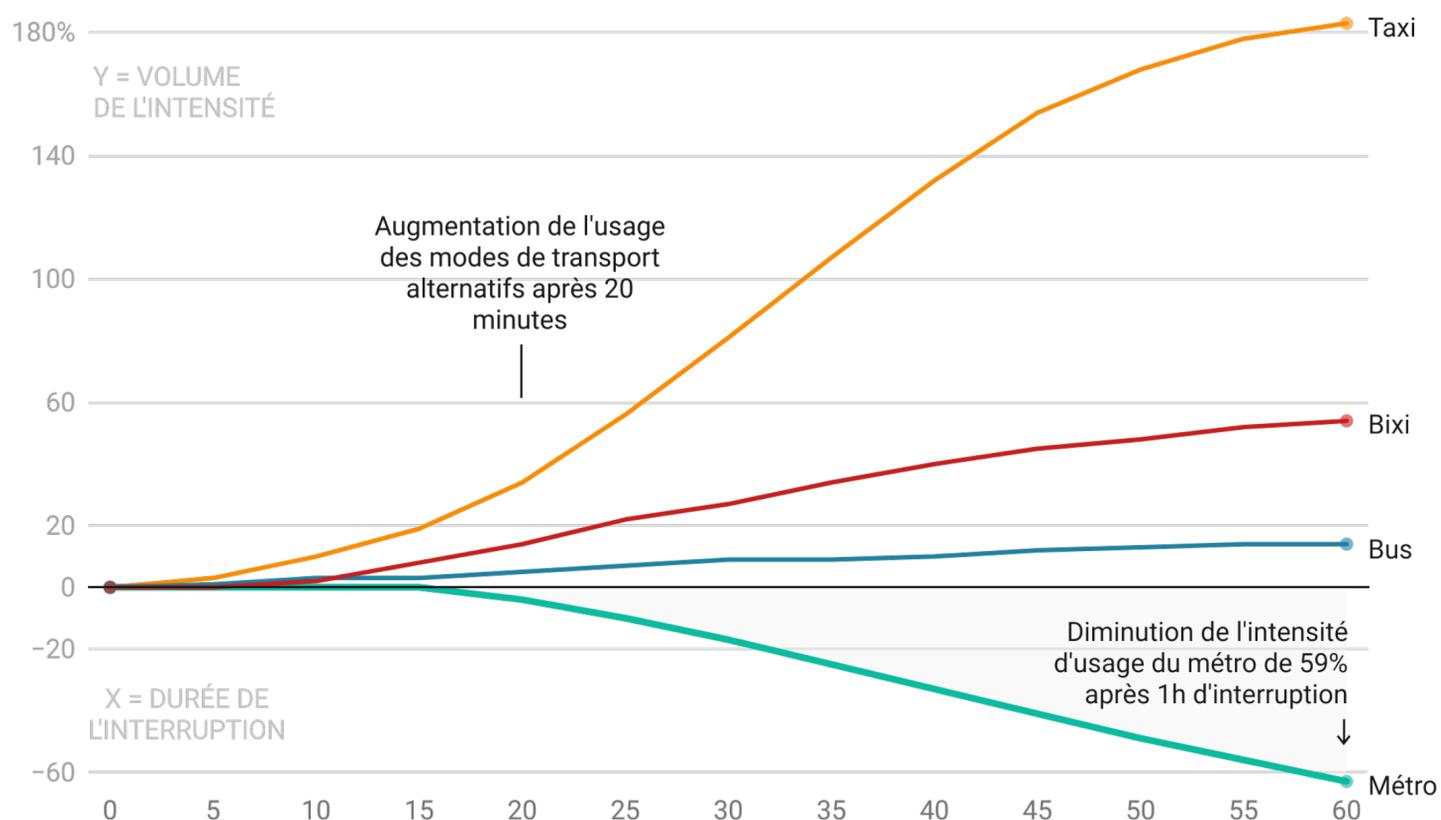


Figure 2. Variation relative de l'intensité par mode selon la durée d'interruption de service
Source: Lepage, 2019

Ces analyses exposent les difficultés liées à l'imprévisibilité des interruptions et les dilemmes cornéliens des utilisateurs des transports collectifs. L'information passager s'avère ainsi un outil efficace pour la prise de décision lors d'interruptions; une communication même imparfaite d'une estimation de l'heure de reprise dans les premières minutes pourrait contribuer à une meilleure expérience client. La fiabilité du service de métro est un facteur clé de la performance d'un réseau de transport ainsi que de la satisfaction de la clientèle (Krishnakumari et al., 2020). Un chapitre subséquent de cette étude introduit notamment le contexte des interruptions de service à Montréal et présente certains indicateurs de performance associés à la fiabilité du réseau de métro de Montréal qui sont détaillés dans les rapports annuels de l'organisation.



Quai
Photo: Steven Wright

3. STRATÉGIE DE RECHERCHE

Cette section introduit les diverses composantes de la stratégie de recherche qui permettront de réaliser les objectifs énoncés ci-dessous.

3.1 Objectifs

Ce projet de recherche comporte deux objectifs principaux:

1. Analyser le contexte des interruptions de service du métro de Montréal
2. Déterminer si les données ouvertes disponibles peuvent aider à prédire la durée des interruptions de service considérant certaines variables indépendantes connues

Cette étude contribuera au corpus scientifique limité portant sur les arrêts de service du métro et permettra aux chercheurs, aux décideurs et à la communauté citoyenne de mieux cerner les enjeux et les possibilités liés aux informations de transport disponibles au public.

3.2 Données

Plusieurs jeux de données ont été utilisés afin de dresser un portrait fidèle des arrêts de service à Montréal et réaliser les régressions pour la période d'étude.

3.2.1 Données principales

Un jeu de données intitulé "Incidents du réseau du métro" a été libéré par la STM et publié sur le portail de données ouvertes de la ville de Montréal. Le jeu de données utilisé, obtenu en décembre 2021, rassemble plus de 18 022 incidents qui ont eu lieu dans les trains ou les stations du métro pour la période allant du 1er janvier 2019 au 31 octobre 2021. Les attributs des incidents sont détaillés dans la figure 3. Ce jeu de données inclut uniquement les incidents qui ont causé des interruptions de plus de 5 minutes. Par ailleurs, seulement une fraction de ces incidents ont causé un délai pour les voyageurs, soit en grande majorité les incidents identifiés "trains". Ainsi, un total de 2 528 incidents ont été répertoriés dans la période d'étude et seront illustrés dans la mise en contexte.

Cependant, pour éliminer l'effet des données extrêmes, seulement les 2 442 incidents qui ont causé un délai de 5 à 60 minutes sur le réseau ont été utilisés pour les analyses et les régressions associées.

Attribut	Description	Exemple 1	Exemple 2
id	Identifiant unique de l'incident	T28021901	T25041901
date	Date de l'incident	2019-02-28	2019-04-25
season	Saison	winter	spring
day	Jour de semaine ou de fin de semaine	weekday	weekday
period	Heure de pointe ou creuse	off-peak	peak
time_start	Heure du début de l'interruption de service	05:35	07:04
time_end	Heure de reprise du service	05:49	07:37
duration	Durée de l'interruption (minutes)	14	33
cause_ini	Cause initiale attribuée au moment de l'incident	Matériel roulant	Clientèle
cause_prim	Cause primaire déterminée suite à l'analyse de l'incident	Matériel roulant	Clientèle
cause_sec	Cause secondaire déterminée suite à l'analyse de l'incident	MR-73	Méfait volontaire
line	Nom de la ligne sur laquelle a eu lieu l'incident	Ligne jaune	Ligne orange
station	Lieu de l'incident	Jean-Drapeau	Rosemont

Figure 3. Liste des attributs liés aux incidents utilisés pour l'analyse

3.2.2 Données complémentaires

Des jeux de données complémentaires ont été employés afin de contre-vérifier les données ouvertes et avoir une meilleure connaissance des informations disponibles au public. D'abord, une demande d'accès à l'information a été effectuée auprès de la STM en novembre 2021. Le jeu de données obtenu comprend 2 082 incidents pour la période d'analyse comparativement aux 2 528 incidents du jeu de données ouvertes.

Ainsi, ces nouvelles informations ont permis de confirmer la grande majorité des incidents, malgré certains attributs divergents. Par exemple, ces nouvelles données incluent le code de la salle de contrôle ainsi que le détail exact de la cause d'une interruption. Un incident décrit comme une "Tentative de mort violente, code 438" dans le jeu de données obtenu suite à la demande d'accès à l'information est représenté comme une interruption pour cause de "Clientèle — Méfait volontaire" dans les informations du portail ouvert. Par ailleurs, seulement la station où l'incident a été déclaré est indiquée, et non la ligne de métro associée, ce qui peut poser problème pour une analyse incluant des incidents aux stations de correspondances où plusieurs lignes se croisent.

Finalement, des données tirées de Twitter ont aussi été utilisées afin d'analyser les communications de la STM durant les interruptions de service. La STM maintient quatre comptes Twitter différents pour ses lignes de métro — @stm_Orange, @stm_Verte, @stm_Bleue et @stm_Jaune — et utilise ces comptes afin de communiquer au public le statut des interruptions de service de plus de 10 minutes. Au-delà de 3,300 tweets publiés depuis 2019 ont été automatiquement collectés via l'outil Twint sur les différents comptes du métro de la STM. Twint est un outil de moissonnage de données écrit en python qui permet d'extraire et de filtrer un grand volume de données de Twitter sans avoir à interagir avec l'API (Application Programming Interface) officiel de la plateforme.

3.3 Méthodologie

Plusieurs étapes ont été nécessaires à la réalisation de cette étude de manière à atteindre les objectifs de recherche définis plus haut.

3.3.1 Traitement des données

Le jeu de données principal utilisé a été nettoyé et enrichi afin de réaliser les analyses nécessaires à ce projet de recherche. Des attributs complémentaires ont aussi été ajoutés au document. Les variables nominales organisées à partir des données brutes, sont:

1. Saison: Selon la date d'incident. Printemps, été, automne, hiver.
2. Jour: Précision si l'incident a eu lieu un jour de semaine ou de fin de semaine.
3. Période: Période de l'incident.

Heure de pointe du matin: incidents entre 6h et 10h

Heure de pointe de l'après-midi: incidents entre 15h et 19h

Heures creuses: incidents survenus durant toute autre période de la journée.

3.3.2 Mise en contexte

Plusieurs visualisations ont ensuite été réalisées afin de mettre en contexte les interruptions de service du métro de Montréal. Un portrait exhaustif pour la période d'étude est introduit. Ce portrait inclut notamment une analyse du nombre d'incidents par lieu et par période, ainsi que les causes principales des interruptions de services. Une estimation probabiliste des interruptions pour une journée de semaine type est aussi présentée.

3.3.3 Régressions

Un des objectifs de cette étude est de déterminer s'il est possible d'estimer la durée approximative d'un arrêt de service en début d'incident en vue d'aider les passagers dans leur prise de décision. Deux techniques de régression ont été employées sur le modèle afin d'estimer la variance des durées des interruptions de service et d'analyser l'effet des attributs disponibles en début d'interruption. Ces variables, toutes nominales, incluent la saison, le jour de la semaine, la période, la ligne de métro ainsi que la cause initiale communiquée.

D'abord, une régression linéaire multiple par la méthode des moindres carrés ordinaire (MCO) a été réalisée. Une régression MCO est possible vu l'apparence continue de la variable dépendante, soit la durée de l'interruption représentée en minutes. Une transformation logarithmique des valeurs de durées a été nécessaire afin de rapprocher le modèle d'une distribution normale étant donné le nombre d'incidents de courte durée. Seulement une régression avec entrée forcée a été réalisée: toutes les variables ont ainsi été évaluées ensemble dans la régression.

Puis, une régression logistique binomiale a permis de poursuivre l'analyse à la lumière des résultats de la régression MCO et d'effectuer des estimations probabilistes en comparant les sous-ensembles d'interruptions courtes (5 à 19 minutes) avec les interruptions longues (20 à 60 minutes).

3.3.4 Analyse et recommandations

L'analyse effectuée permet de déterminer quelle proportion de la variance des durées peut être expliquée par le contexte spatio-temporel de l'incident, et de vérifier si le modèle est apte à effectuer des prédictions utilisables à partir de ces informations de base. Des recommandations propres au réseau de métro de Montréal ont ensuite été élaborées,

basées sur la littérature existante, le portrait des incidents du réseau, ainsi que des résultats des analyses.

3.4 Limites

Une des principales limites de l'étude est l'impact de la baisse de l'achalandage sur le réseau depuis mars 2020 en raison de la pandémie de COVID-19. Cette diminution d'achalandage marquée — 54,2 % sur l'ensemble du réseau de la STM en 2020 par rapport à 2019 — s'est aussi traduite par une diminution du nombre d'incidents sur les lignes de métro. Or, pour une mise en contexte plus représentative de la normale, les données de 2019 seront mises de l'avant. Des choix méthodologiques de simplification des causes pour les données principales ont aussi été faits, et seront détaillés dans l'analyse. Finalement, vu l'intention initiale de réaliser un prototype d'application basé sur le modèle de régression, une simple régression linéaire multiple a été effectuée. Des modèles plus avancés pourraient être développés avec des jeux de données plus étoffés et pour une période plus longue, mais cela dépasse le cadre de cette étude.



Infiltration d'eau au métro Square Victoria-OACI
Photo: STM, 2019

4. MISE EN CONTEXTE

4.1 Lignes et achalandage

Inauguré en 1966 avec 26 stations, le métro de Montréal dessert aujourd'hui 68 stations réparties sur quatre lignes — orange, verte, bleue et jaune. Les quatre lignes totalisent 71 kilomètres, desservant l'île de Montréal et les villes voisines de Longueuil et Laval. Des stations de correspondances entre les différentes lignes (Snowdon, Lionel-Groulx, Berri-UQAM et Jean-Talon) sont localisées stratégiquement dans certains des quartiers centraux les plus denses de la ville et comptent parmi les stations les plus fréquentées du réseau.



Figure 4. Carte du réseau de métro de Montréal

Le métro de Montréal est le réseau le plus achalandé au Canada et l'un des plus importants d'Amérique du Nord. Bien qu'opéré par la STM, c'est l'Autorité régionale de transport métropolitain (ARTM) qui chapeaute la planification et le financement des différents modes de transport, puis assure leur intégration optimale.

Les transports en commun montréalais ont connu plusieurs bouleversements importants depuis les années 1950, affectant de façon significative les choix de déplacements des utilisateurs. Selon les dernières données disponibles, la part modale du transport collectif et des déplacements bimodaux est de 24%, en légère augmentation depuis la fin des années 2000 (ARTM, Enquête Origine-Destination 2018).

Les seules données historiques sur l'achalandage à Montréal regroupent les différents modes opérés par la STM. Néanmoins, on peut imaginer que depuis le début des années 1970, le métro de Montréal connaît une croissance soutenue de son achalandage, suivant la tendance générale des déplacements en transports collectifs.

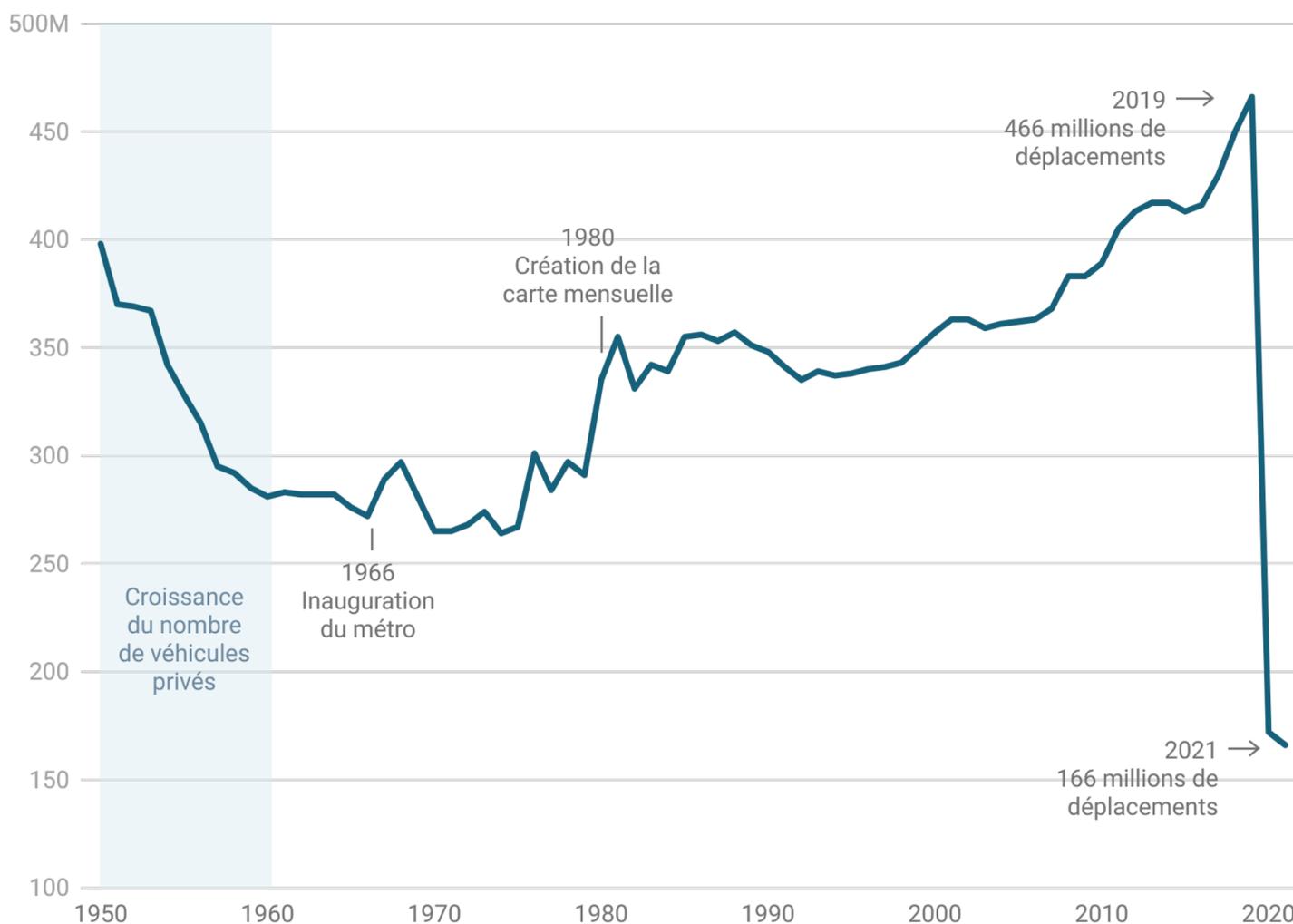


Figure 5. Historique de l'achalandage du transport en commun à Montréal (STM)

La figure 5 introduit certains moments clés qui ont marqué l'utilisation des transports collectifs à Montréal depuis les années 50. L'inauguration du métro a donné une première poussée des déplacements depuis la première partie du 20e siècle. On peut notamment observer une croissance soutenue de l'achalandage depuis les années 90 jusqu'à récemment. Un sommet du nombre de déplacements a été atteint en 2019, avec plus de 466 millions de déplacements réalisés à l'échelle du réseau de la STM. Une chute drastique du nombre des déplacements marque ensuite les années 2020 et 2021 en raison du changement des habitudes de la clientèle en réponse aux mesures liées à la pandémie de COVID-19.

À noter que ces informations se basent sur l'achalandage conventionnel, basé sur un calcul des ventes de titres, dans un souci d'uniformité. C'est seulement en 2018 que la STM a introduit le calcul de l'achalandage électronique à ses rapports annuels, soit le nombre de passages réalisés par les clients sur le réseau de la STM.

4.2 Performance et probabilité

Des investissements dans le maintien des équipements et le développement du réseau sont importants afin d'assurer la fiabilité des opérations et de satisfaire les besoins en déplacement des usagers. Entre 2020 et 2025, plus de 1,6 milliard de dollars seront investis à l'échelle du réseau afin de procéder notamment au remplacement des équipements fixes et à la réfection d'infrastructures (STM, rapport annuel, 2019).

Malgré des investissements colossaux, les pannes de service non planifiées sont inévitables sur les réseaux de métro à travers le monde. La STM mesure la performance face à ces interruptions en chiffrant le nombre d'incidents de 5 minutes et plus par million de kilomètres.

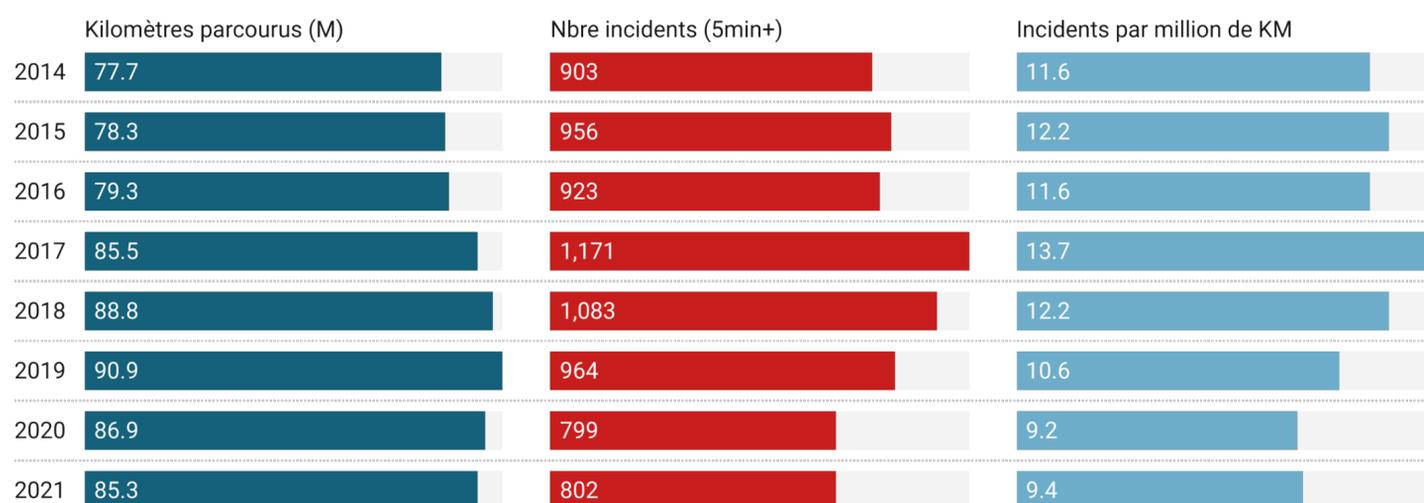


Figure 6. Incidents et millions de kilomètres parcourus sur le réseau de métro.
Données: STM, 2014 à 2021.

Les données publiées dans les rapports annuels de la STM révèlent une diminution du nombre d'incidents de 5 minutes et plus entre 2014 et 2020, passant de 11,6 incidents par million de kilomètres parcourus en 2014 à 9,2 incidents en 2020. Ceci dit, vu la réduction d'achalandage en 2020 causée par la pandémie de COVID-19, la STM précise que "la fiabilité du matériel et des équipements ainsi que la réduction des incidents impliquant des clients blessés ou malades ont grandement contribué à atteindre ce chiffre, le meilleur depuis une dizaine d'années." En 2019, avec 10,9 incidents par million de kilomètres, la STM rapportait que le métro de Montréal performait mieux que la moyenne des métros à travers le monde. Lors de l'annonce des résultats annuels, le président de la STM Philippe Schnobb tenait à démentir la légende urbaine selon laquelle le nombre d'incidents était en hausse et insistait sur l'amélioration progressive de la performance du réseau (CBC, 2019).

Comme mentionné dans la revue de littérature, peu de recherches exhaustives publiques ont été menées à propos des interruptions de service dans les réseaux de métro du pays. Des articles de quotidiens font mention des statistiques annuelles comme discuté, mais une analyse du journal Le Devoir publiée en 2019 sort du lot. À l'aide de visualisations interactives, elle illustre habilement certaines tendances concernant les pannes de service pour la période allant de janvier 2015 à mars 2019.

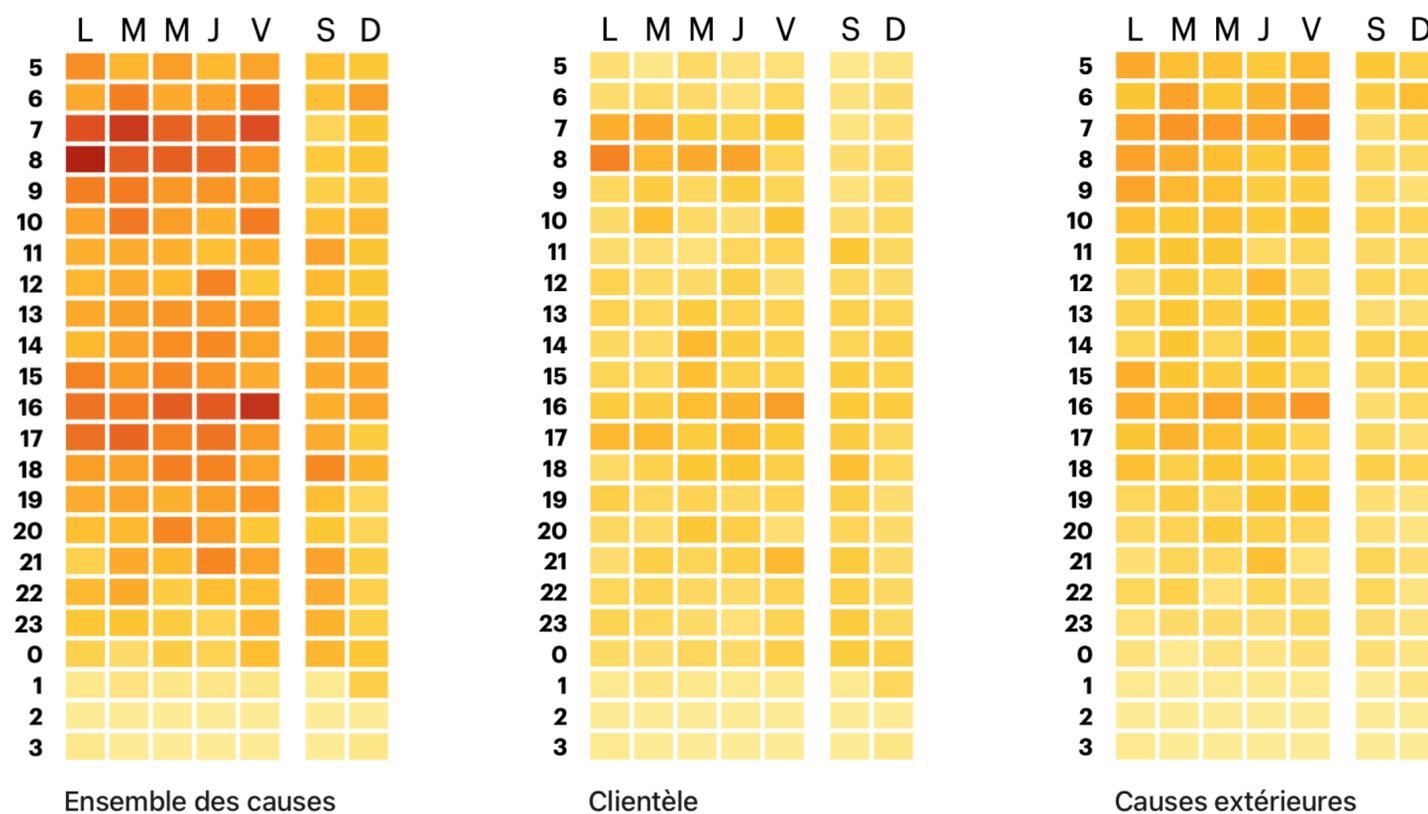
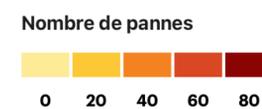


Figure 7. Causes des pannes de service, affichées par période et jour de la semaine
Source: Le Devoir, 2019



Un des enseignements de l'analyse est la présence d'un plus grand nombre d'incidents aux heures de pointe d'achalandage, où les données de Google ont été utilisées. Une modélisation des pannes de service par station permet de visualiser par année la concentration des interruptions. On peut observer qu'un grand nombre d'interruptions dites de "Cause usagers" semblent se produire sur les segments de lignes au centre-ville de Montréal, contrairement aux "Causes extérieures" qui elles se produisent plutôt en fin de ligne. Des enseignements similaires ont pu être observés dans les jeux de données utilisés pour la présente étude, bien que le traitement des informations ait été effectué différemment.

Tel que mentionné précédemment, il y a souvent un biais de négativité associé à la performance du réseau de métro. Les utilisateurs de transport en commun ont souvent l'impression que la probabilité de faire face à une interruption de service est plus forte que la réalité. Même certains utilisateurs qui ont eu une mauvaise expérience par le passé justifient leur choix d'un mode alternatif, tel que la voiture solo, afin d'éviter de faire face à une autre mauvaise expérience de transport en commun (Donovan, 2015).

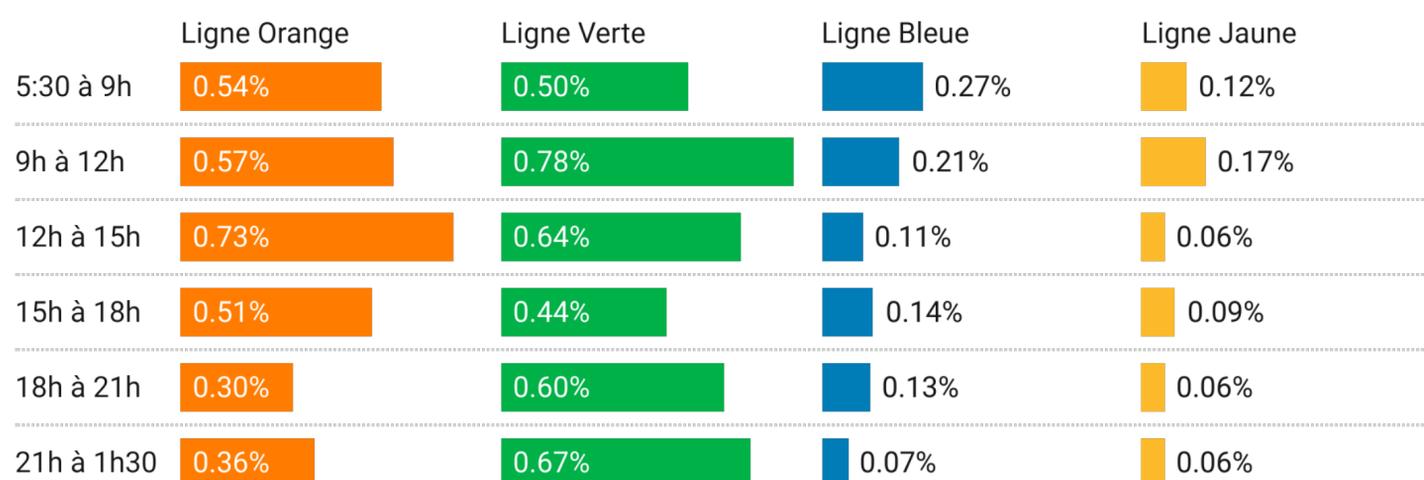


Figure 8. Probabilité de faire face à une interruption de service un jour de semaine en 2019. Mise en situation: une passagère qui emprunte la ligne verte à 9h30 a en moyenne 0.78% de chances de faire face à une interruption de plus de 5 minutes, comparativement à 0.21% à la même heure sur la ligne bleue. Ces estimations ont été calculées en divisant la somme des interruptions en 2019 sur la ligne pour la période en jour de semaine, par la somme des départs de métro pour la même période à une station de la ligne.

Données: STM, 2019.

Une statistique qui n'a — à priori — pas été communiquée ni calculée dans la littérature pour le réseau de Montréal est la probabilité pour un utilisateur de faire face à une interruption de service sur une ligne donnée, à un moment de la journée. En combinant les données des départs planifiés des lignes et le nombre d'incidents par période, il est possible d'estimer la

probabilité, toutes durées confondues, de subir une interruption pour un jour de semaine. La figure 8 montre que les périodes avec la plus forte probabilité sont sur la ligne verte entre 9h et 12h, puis sur la ligne orange entre 12h et 15h. Malgré le nombre réduit de départs sur les lignes bleue et jaune, les chances sont tout de même plus minces d'éprouver un délai causé par une panne de service. Il y a généralement une probabilité plus réduite de faire face à une interruption en soirée.

4.3 Portrait et tendances des interruptions de service

Le jeu de données ouvertes concernant les interruptions de service ainsi que les données obtenues en vertu d'une demande d'accès à l'information permettent d'avoir un portrait exhaustif des pannes sur le réseau entre le 1er janvier 2019 et le 31 octobre 2021. Cette section détaille les tendances tirées de ces données en appui aux choix d'analyses qui ont été effectués.

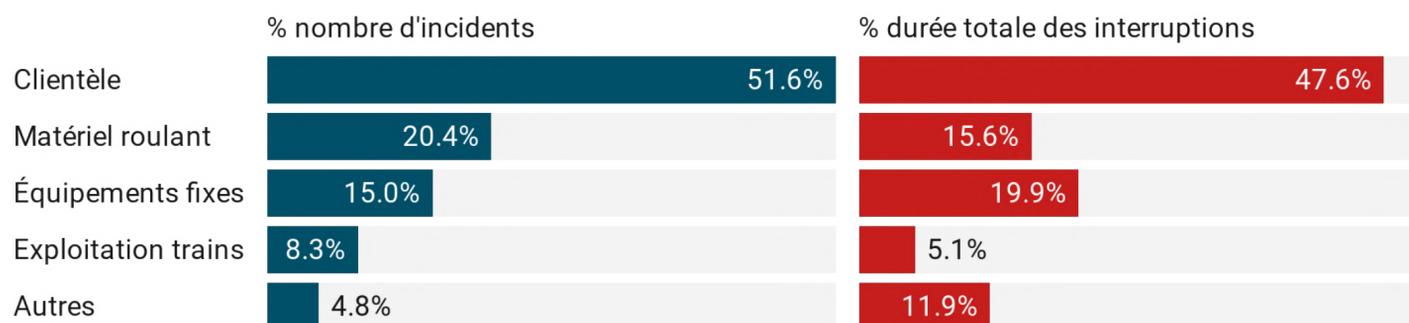


Figure 9. Comparaison entre la proportion du nombre d'incidents d'une cause et la part de la durée totale des interruptions pour la période étudiée.

Comme discuté dans la revue de littérature, les causes des incidents sont parmi les plus importants déterminants de la durée des interruptions. L'analyse de Barron & al (2013) pointe l'importance d'étudier l'impact réel des différentes causes sur les usagers, et ne pas s'en tenir qu'au nombre afin d'étudier la performance d'un réseau. Les données ouvertes concernant les usagers de la STM étant limitées, il n'a pas été possible d'évaluer comme Barron le pourcentage de passagers affectés par les différentes causes à Montréal. Néanmoins, comparer la proportion du nombre d'incidents d'une cause à sa part de durée totale démontre l'idée que les causes n'affectent pas les usagers du réseau de la même façon.

La figure 9 montre avec évidence que les incidents dus à la clientèle impactent le réseau autant en fréquence qu'en durée: 291 heures d'interruptions (47,6% de la durée totale entre janvier 2019 et octobre 2021) sont ainsi dues aux usagers eux-mêmes. Les incidents dus au matériel roulant représentent pour leur part 20,4% du nombre total d'incidents de la période étudiée, mais occupent seulement 15,6% de la durée totale. Les interruptions dues aux motifs autres — tels que les incidents d'émanation — sont parmi celles qui arrivent le moins fréquemment (4,8%), mais qui représentent plus du double de leur fréquence en durée (11,9%).

La figure 10 présente les tendances au fil des mois en illustrant la durée cumulée (minutes) des interruptions de service ventilées par cause. Certains mois connaissent des sommets évidents, notamment en 2019. Parmi les incidents les plus connus et médiatisés: les infiltrations d'eau qui ont paralysé la ligne verte (mai) et orange (octobre) pendant de nombreuses heures et un méfait dit "volontaire" qui a mis à l'arrêt des stations de la ligne bleue en avril 2019 pendant 14 heures. Ces incidents sont visiblement l'exception, mais ont néanmoins un impact considérable sur les itinéraires de centaines de milliers de clients lorsqu'ils surviennent. Ainsi, rapidement identifier et communiquer aux passagers la cause en début d'interruption est d'une importance capitale.

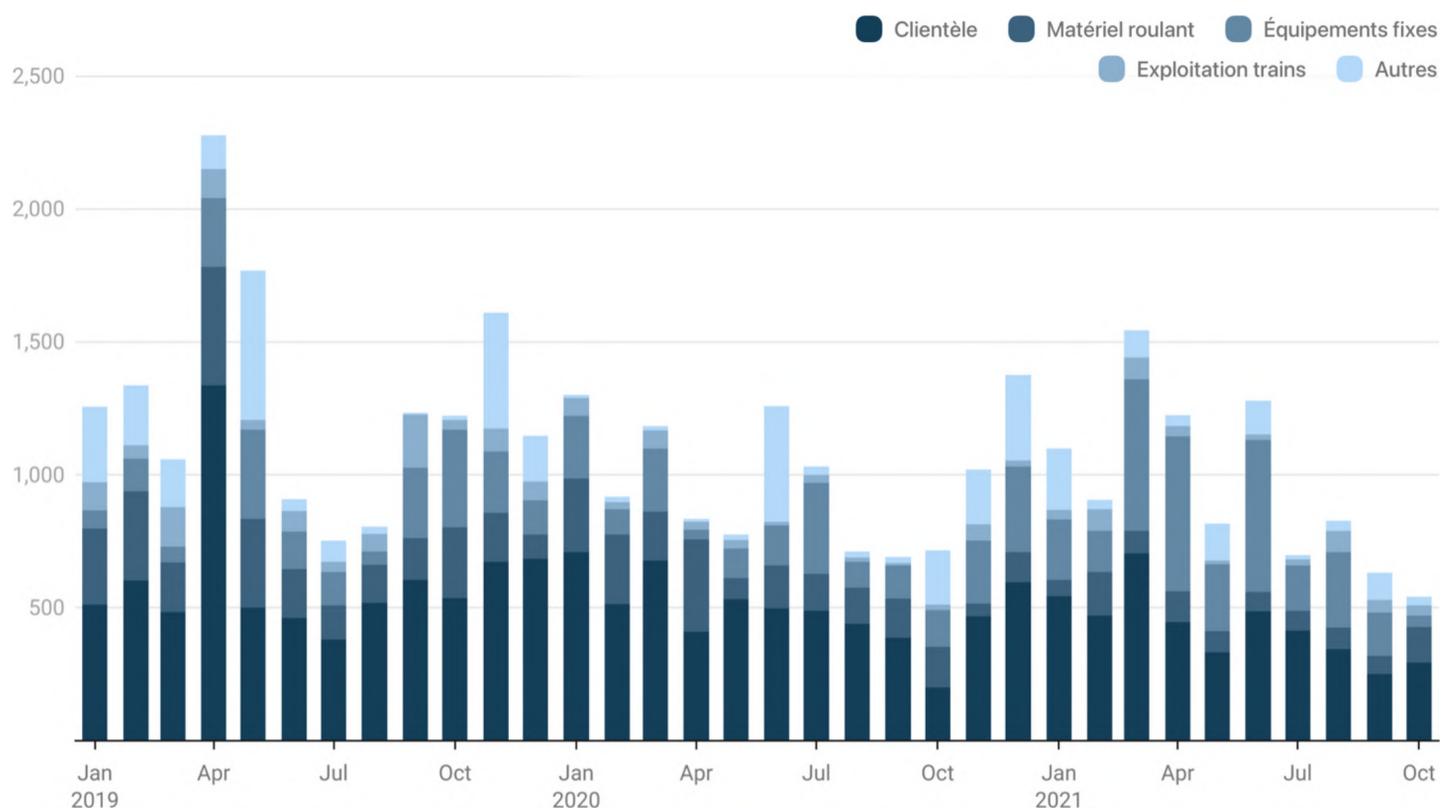


Figure 10. Durées cumulées (minutes) des interruptions de service, ventilées par cause

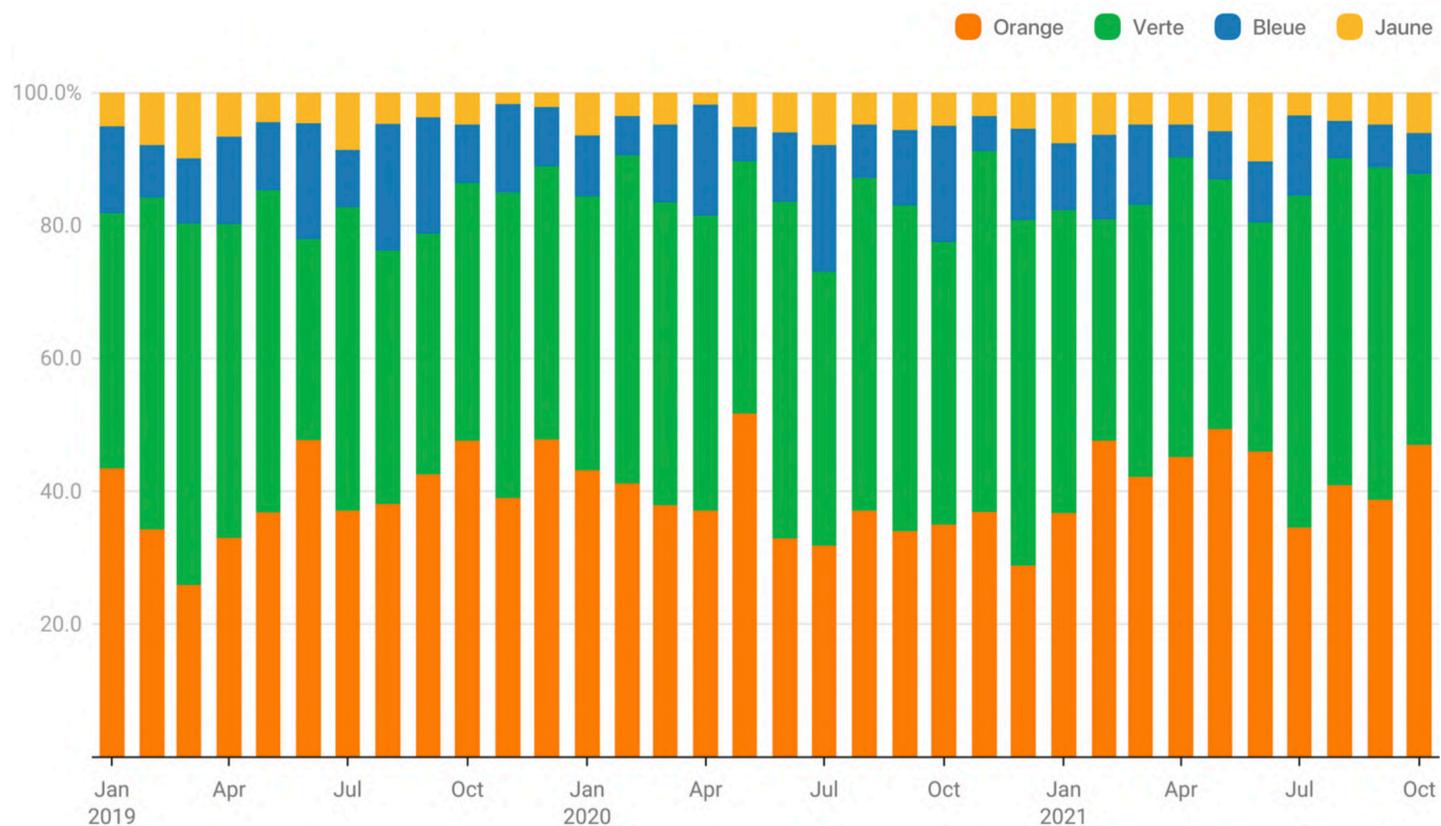


Figure 11. Proportion du nombre mensuel d'interruptions de service, ventilées par ligne

Une autre manière d'illustrer les tendances historiques est de ventiler les incidents par ligne de métro. La figure 11 révèle la proportion du nombre total mensuel d'interruptions: les lignes verte et orange sont habituellement coude-à-coude, comptant respectivement 44% et 40% du nombre total d'interruptions pour la période d'étude. Les lignes bleue et jaune se placent au 3e et 4e rang, avec respectivement 11% et 5% du nombre d'incidents total.

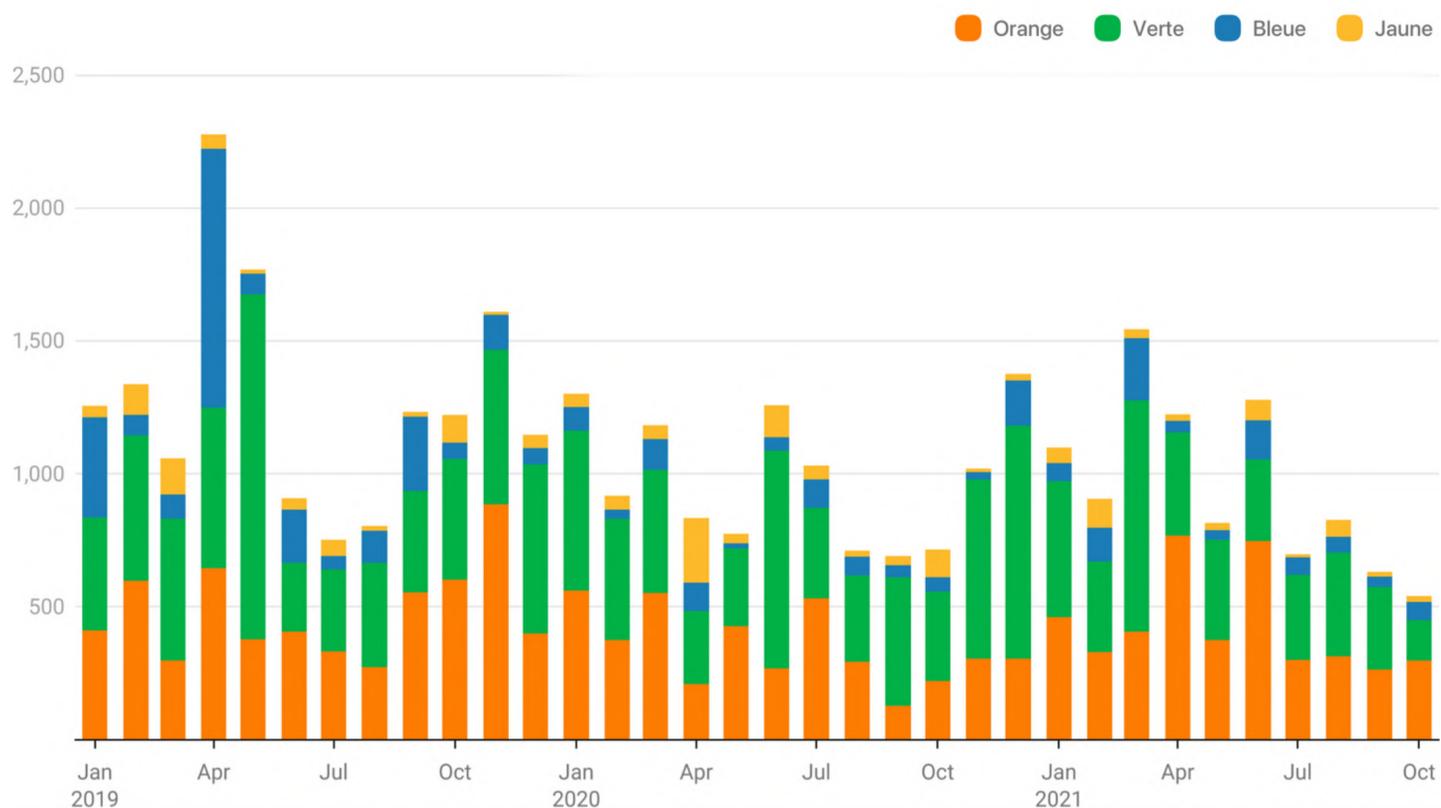
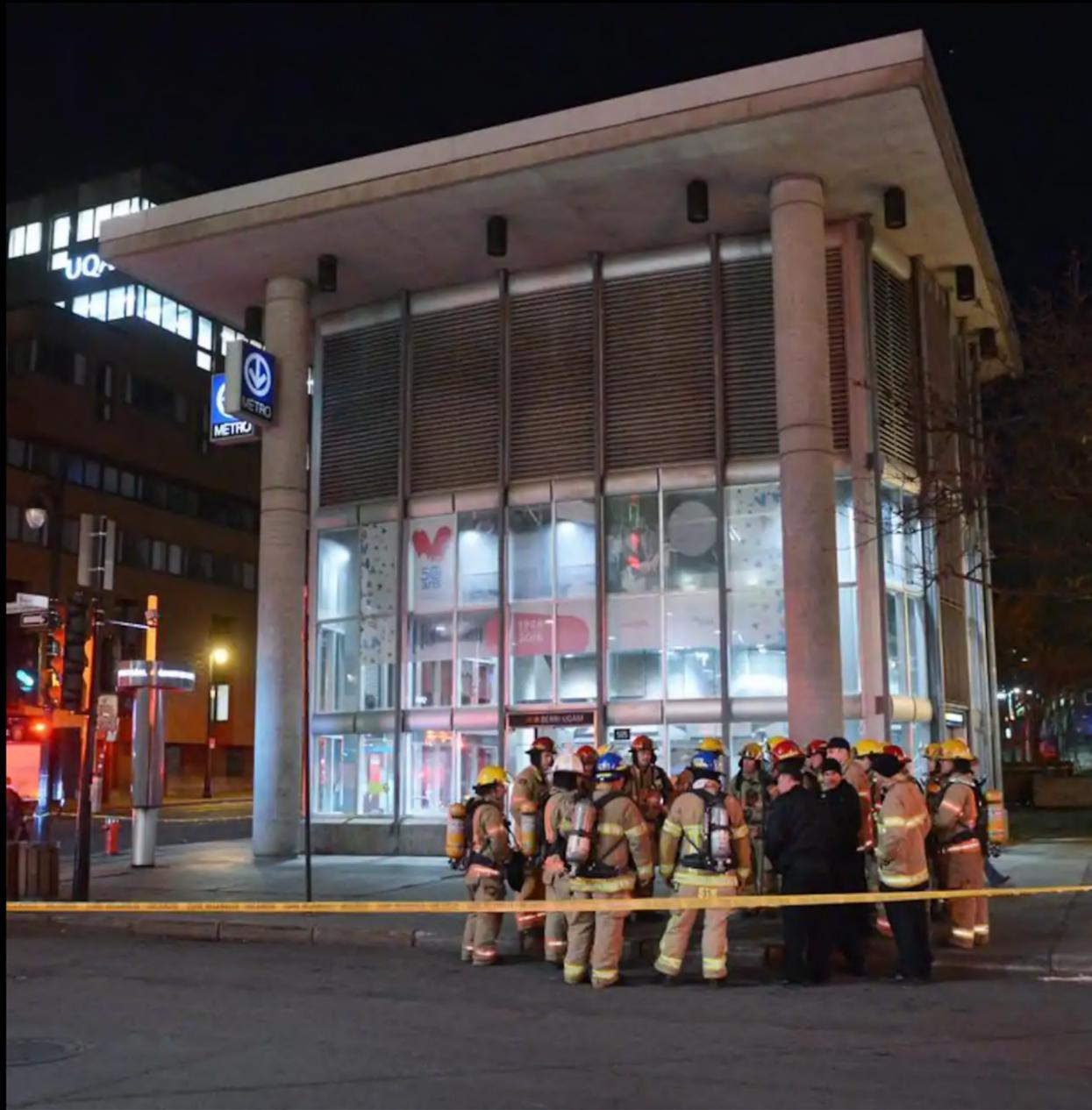


Figure 12. Durées cumulées (minutes) des interruptions de service, ventilées par ligne

Enfin, les durées ventilées par ligne dévoilent des tendances similaires au nombre d'incidents. Malgré les quelques sommets d'interruptions sur les différentes lignes, les sommes sur la période indiquent que les proportions des durées totales sont égales aux proportions du nombre d'incidents. Par ailleurs, à première vue, les tendances ne semblent pas connaître d'importantes variations au fil des saisons. Une analyse plus détaillée des valeurs temporelles permettra de vérifier l'importance des différentes variables sélectionnées.

Les tendances présentées indiquent déjà que les causes semblent d'importants déterminants de la prédiction des durées des interruptions pour le développement du modèle. Les autres variables spatio-temporelles d'intérêt — soit les lignes associées, la saison, le jour de la semaine, ainsi que l'heure des incidents — seront étudiées davantage dans l'analyse afin de déterminer leur rôle dans la prédiction des durées.



Intervention des services d'urgence à la station Berri-UQAM
Photo: Pascal Girard - QMI, 2019

5. ANALYSE

5.1 Précisions

L'étude du jeu de données brut a permis d'identifier certaines modifications nécessaires afin de procéder à une analyse fidèle et représentative des interruptions de service et de répondre aux objectifs de recherche.

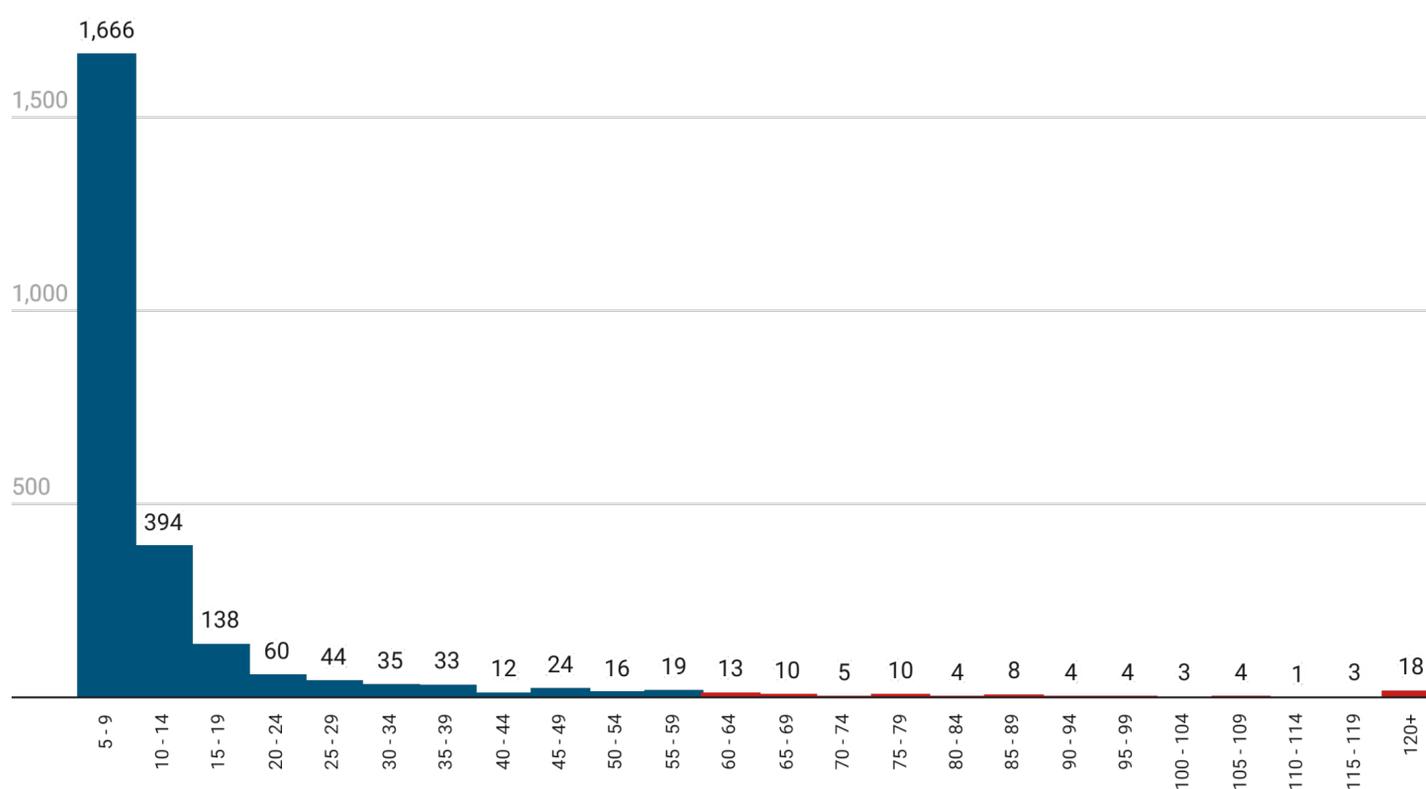


Figure 13. Histogramme de la durée des interruptions de service du jeu de données brut, par intervalles de 5 minutes. Les incidents de plus de 60 minutes ont été retirés du jeu de données final.

D'abord, lorsqu'on visualise l'histogramme de la durée des observations, on constate rapidement que la distribution est désaxée vers la droite. La durée moyenne des interruptions sur l'ensemble du réseau (14 minutes) est nettement supérieure à la médiane (7 minutes). Une minorité de résultats d'une durée de plus d'une heure, représentant 3,44% de l'ensemble du jeu de données, semblent avoir un effet sur le déséquilibre de la distribution. Vu l'objectif de déterminer si les informations inférées du contexte spatio-temporel peuvent être utilisées afin d'estimer la durée d'une interruption de service dans la grande majorité des cas, il a été décidé de retirer les incidents de plus d'une heure pour le modèle final — intervalles identifiés en rouge sur la figure 13.

On obtient ainsi une durée moyenne de 10 minutes pour la même médiane de 7 minutes. Afin de se rapprocher le plus possible d'une distribution normale, une transformation logarithmique des valeurs de durées a été effectuée. La figure 14 montre les histogrammes des valeurs en minutes — modèle 1 — et en $\log(\text{minutes})$ — modèle 2. Une asymétrie semble toujours présente suite à la transformation. Néanmoins, l'analyse détaillée permettra de déterminer si ces modifications permettront de développer un modèle prédictif fiable.

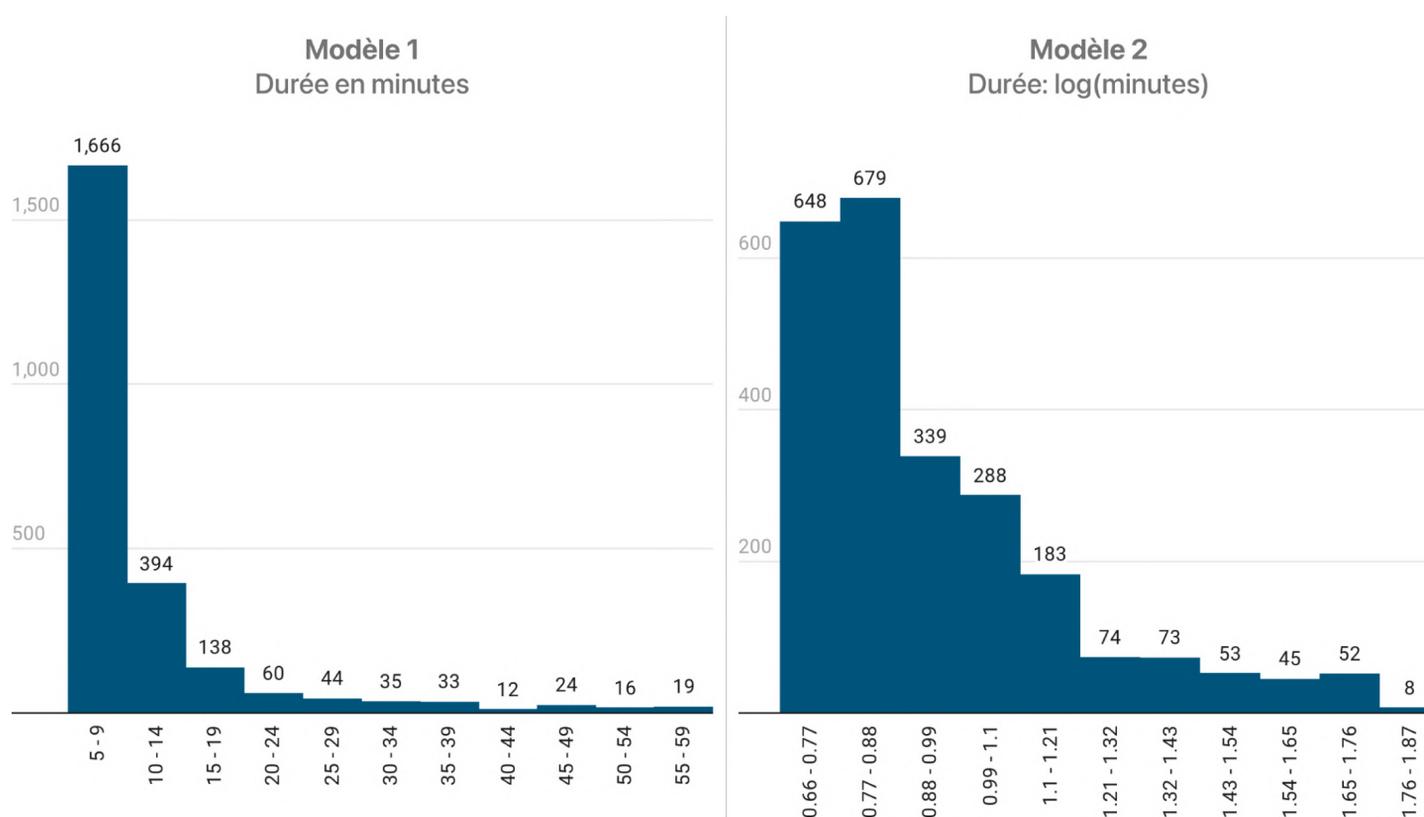


Figure 14. Histogramme de la durée des interruptions de service: modèle 1 en minutes, modèle 2 en $\log(\text{minutes})$ avec le même nombre d'intervalles (11).

De plus, étant donné les causes variées contenues dans le jeu de données brut et le manque d'observations pour certaines d'entre elles, une nouvelle catégorisation des causes a été réalisée afin d'obtenir des résultats acceptables. Ainsi, les observations du jeu de données ont été combinées et organisées en quatre classes distinctes, présentées dans la figure 15.

En ce qui concerne la période de la journée — heure de pointe ou creuse — la fréquence horaire des métros a été utilisée comme proxy afin d'effectuer la classification. Les incidents identifiés "heure de pointe" ont eu lieu soit entre 6h et 10h du matin, ou entre 15h et 19h inclusivement. Les incidents identifiés "heure creuse" se sont produits durant les heures d'ouverture du métro — 5h30 à 1h30 — mais à l'extérieur des heures de pointe.

Causes - Nouvelle catégorisation	Causes - Jeu de données brut	Nombre d'incidents
Clientèle	Clientèle	869
Matériel roulant	Matériel roulant	908
Incident d'émanation	Feu, fumée, odeur, produit, etc.	61
Divers	Équipements fixes	448
	Exploitation trains	128
	Divers	28
	Total: Divers	604
	Total	2,442

Figure 15. Tableau présentant le résultat de la nouvelle catégorisation des causes pour le jeu de données final.

5.2 Statistiques descriptives

Après traitement, ce sont un total de 2 442 interruptions de service survenues entre le 1er janvier 2019 et le 31 octobre 2021 qui ont été utilisées afin de réaliser les régressions.

	N	Moyenne	Médiane	SD	Min	Max
Modèle 1	2,442	10.5	7	9.3	5	60

Figure 16. Statistiques descriptives du modèle. 2,442 incidents de 5 à 60 minutes. Durée en minutes.

Les statistiques des variables indépendantes spatio-temporelles retenues pour l'analyse sont détaillées dans la figure 17. Elles incluent la saison, le jour de semaine, la période de la journée, la ligne de métro ainsi que la cause confirmée suite à une interruption de service.

En observant la figure 17, on constate que les moyennes des durées sont supérieures aux médianes pour l'ensemble des variables, renforçant l'idée que la distribution est désaxée vers la droite avec une majorité des observations d'une durée de 7 minutes ou moins. Les seules variations importantes concernent les causes d'interruptions. Les incidents provoqués par le matériel roulant, bien qu'ils représentent le motif avec le plus grand nombre de pannes pour la période, sont habituellement de courte durée. Avec une moyenne de 8,6 minutes et un écart type (SD) de 5,4 minutes, on peut assurer que 68% des incidents dus au matériel roulant sont d'une durée de moins de 14 minutes. À l'opposé, les incidents d'émanation semblent se produire peu fréquemment à l'échelle du réseau, mais entravent les déplacements des usagers pour des durées significatives. Avec seulement 61 incidents, la médiane est la plus haute de l'ensemble des variables à 23 minutes.

Variable	Valeur	N	Moyenne	Médiane	SD	Min	Max
Ligne	Orange	972	10.4	7	9.9	5	60
	Verte	1,068	10.6	7	8.9	5	59
	Bleue	270	10.8	8	9.3	5	58
	Jaune	132	10.0	6	8.7	5	47
Cause	Clientèle	869	11.1	7	10.1	5	60
	Matériel roulant	908	8.6	7	5.4	5	53
	Incident d'émanation	61	25.0	23	17.7	5	59
	Divers	604	11.0	7	10.1	5	59
Saison	Printemps	630	10.3	7	9.2	5	58
	Été	575	10.1	7	8.4	5	59
	Automne	521	10.4	7	9.4	5	59
	Hiver	716	11.1	7	10.1	5	60
Jour	Semaine	1,876	10.6	7	9.5	5	60
	Fin de semaine	566	10.3	7	8.9	5	59
Période	Heure creuse	1,294	10.9	7	9.8	5	60
	Pointe	1,148	10.1	7	8.8	5	59

Figure 17. Statistiques descriptives détaillées. Incidents de 5 à 60 minutes. Durée en minutes.

La figure 18 permet de voir en détail les statistiques descriptives des causes ventilées par ligne de métro. On observe ainsi que les moyennes et durées des causes sont relativement constantes à travers le réseau: les causes de matériel roulant prennent en moyenne le même temps court à être résolus sur les différentes lignes (environ 8 minutes) comparativement aux incidents d'émanation qui prennent le plus de temps, soit entre 21 et 28 minutes en moyenne. On note aussi que les pannes de service dues aux incidents d'émanation ont l'écart type le plus élevé, supportant déjà l'hypothèse qu'il est difficile d'estimer avec précision la durée de certains types d'interruptions.

Ligne	Cause	N	Moyenne	Médiane	SD	Min	Max
Orange	Clientèle	428	10.7	7	10.2	5	60
	Matériel roulant	238	8.2	6	6.1	5	53
	Incident d'émanation	26	27.8	22	20.6	5	59
	Divers	280	10.2	7	9.0	5	57
Verte	Clientèle	360	11.4	8	9.9	5	59
	Matériel roulant	441	8.8	7	5.0	5	48
	Incident d'émanation	25	22.6	18	16.3	5	55
	Divers	242	11.3	7	10.5	5	59
Bleue	Clientèle	51	13.6	9	11.6	5	54
	Matériel roulant	163	8.9	7	5.6	5	45
	Incident d'émanation	8	24.4	30	13.0	5	38
	Divers	48	12.3	7	12.6	5	58
Jaune	Clientèle	30	8.9	6	9.1	5	46
	Matériel roulant	66	8.1	6	5.2	5	36
	Incident d'émanation	2	21.5	22	16.3	10	33
	Divers	34	14.2	9	11.7	5	47

Figure 18. Statistiques descriptives détaillées, ventilées par ligne et par cause. Incidents de 5 à 60 minutes. Durée en minutes.

5.3 Résultats et discussion

Le test de signification des modèles — où $p < 0.001$ — permet de rejeter l'hypothèse nulle et affirmer que les variables indépendantes choisies peuvent constituer un modèle admissible pour l'hypothèse alternative.

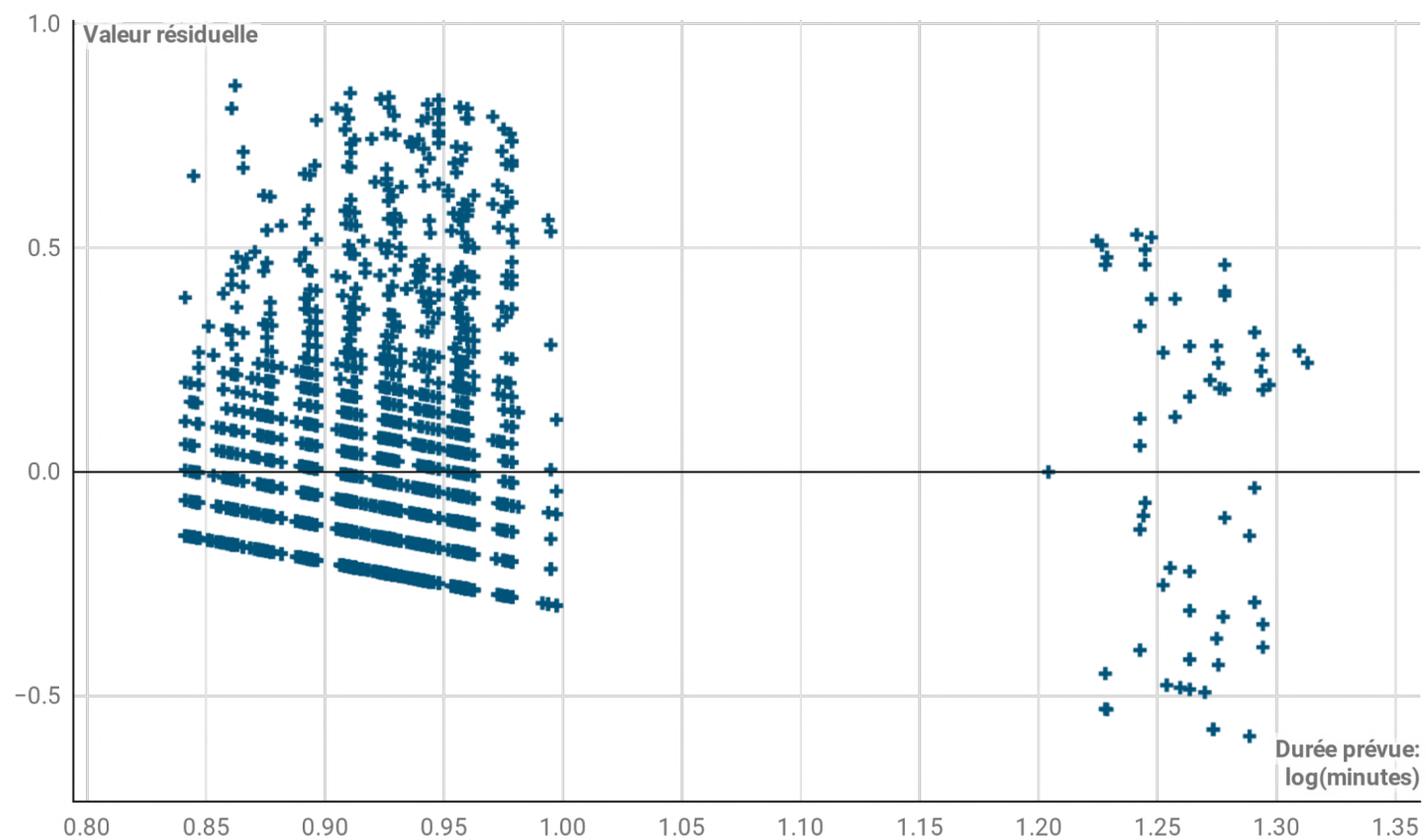
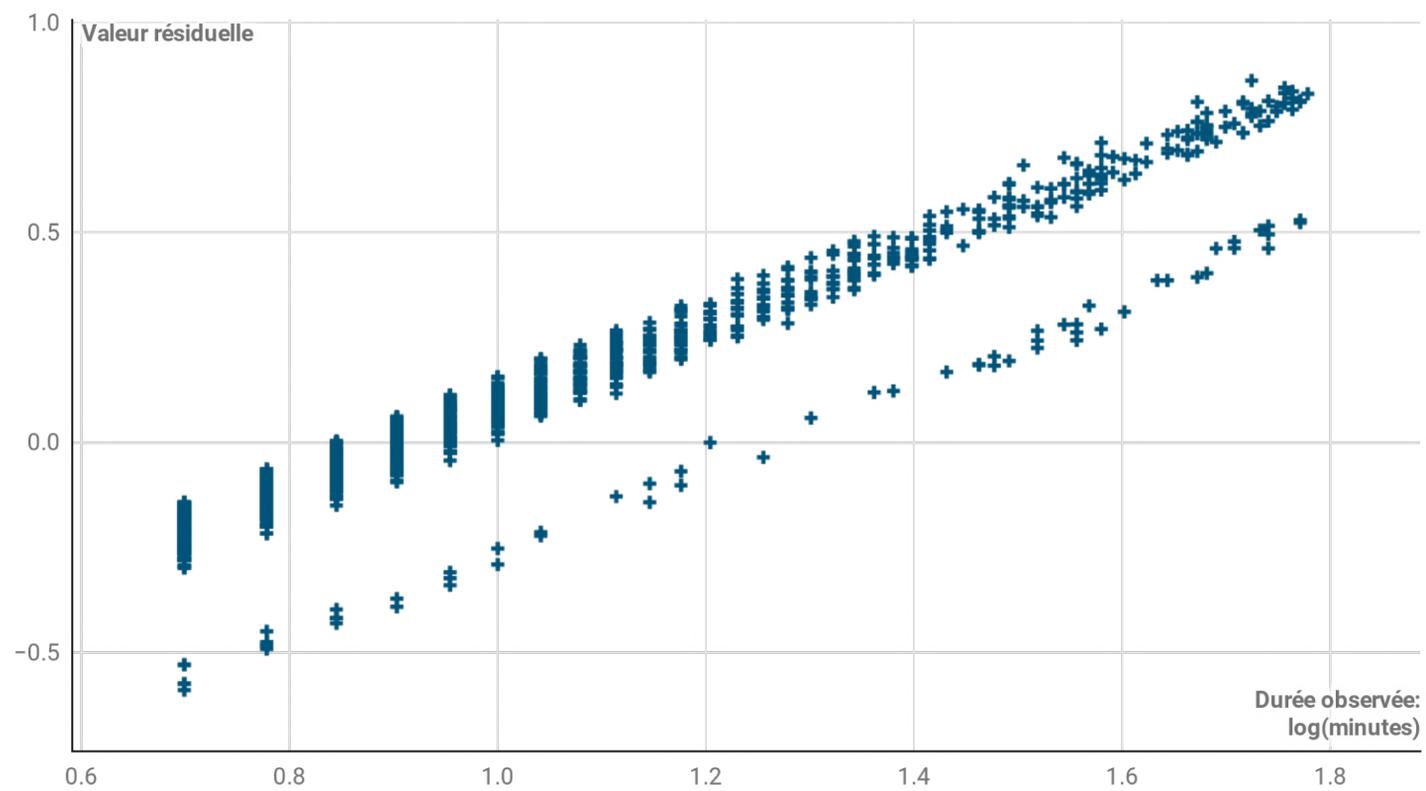
Modèle	R	R ²	R ² ajusté	Test de signification			
				F	df1	df2	p
1 - Durée (minutes)	0.287	0.0823	0.0782	19.8	11	2430	< .001
2 - Durée (logarithmes b10)	0.255	0.0651	0.0609	15.4	11	2430	< .001

Figure 19. Statistiques de régression et test de signification.

Modèle 1 — Incidents avec la durée en minutes.

Modèle 2 — Incidents avec la durée en log(minutes).

Néanmoins, les résultats indiquent que les variables étudiées expliquent à peine 6% (R^2 ajusté de 0.0609) de la variance des logarithmes des durées des interruptions, laissant présager que les informations disponibles en début d'interruption de service sont de faibles prédicteurs de durée, malgré les modifications apportées au jeu de données.



Haut — Figure 20. Graphique de valeurs résiduelles des durées observées
Bas — Figure 21. Graphique de valeurs résiduelles des durées prévues (log)

Les graphiques des valeurs résiduelles exposent les lacunes du modèle d'analyse. On peut constater une différence significative entre les durées observées et les durées prévues estimées par le modèle (figure 20). La disparité est notamment importante pour les incidents avec les durées observées les plus longues. Ceci dit, la pente indique qu'il existe un certain lien entre les valeurs résiduelles et observées.

Le graphique mettant en contraste les données prévues et les valeurs résiduelles (figure 21) révèle un écart important à la droite de zéro. Scindé en deux, on observe qu'une minorité des interruptions sont le résultat de prédictions beaucoup plus longues, au-delà de 1.20 log(minutes), soit 16 minutes en valeur absolue. La vérification des variables des observations irrégulières révèle qu'elles sont toutes associées à des incidents de cause d'émanation. Le tableau des coefficients à la figure 22 montre par ailleurs que la variable des incidents d'émanation, avec un coefficient élevé de 0.381, est de loin la plus significative.

	Variab	Coefficients	SE	t	p
Intercept ^a	-	0.927	0.020	45.876	< .001
Saison	Printemps	-0.021	0.013	-1.566	0.118
	Été	-0.019	0.014	-1.367	0.172
	Automne	-0.019	0.014	-1.331	0.183
	Hiver	0.000	0.000	NaN	NaN
Jour	Semaine	0.004	0.012	0.313	0.754
	Fin de semaine	0.000	0.000	NaN	NaN
Période	Heure de pointe	-0.016	0.010	-1.587	0.113
	Heure creuse	0.000	0.000	NaN	NaN
Ligne	Orange	-0.049	0.017	-2.857	0.004
	Verte	-0.019	0.017	-1.117	0.264
	Bleue	0.000	0.000	NaN	NaN
	Jaune	-0.040	0.026	-1.536	0.125
Cause	Clientèle	0.066	0.012	5.477	< .001
	Matériel roulant	0.000	0.000	NaN	NaN
	Incident d'émanation	0.381	0.032	11.828	< .001
	Divers	0.063	0.013	4.834	< .001

Figure 22. Tableau des coefficients de l'ensemble des variables du modèle 2: durées en log(minutes)

Ainsi, on peut conclure la présence d'hétéroscédasticité, soit une variation dans les erreurs observées, indiquant la difficulté du modèle à estimer la durée des interruptions. Le tableau des coefficients révèle qu'avec l'ensemble des variables modélisées dans la régression, seulement quelques-unes semblent affecter la durée de façon significative. On peut statuer avec les résultats que les variables temporelles — saison, jour de semaine, période — ne semblent pas déterminantes dans les prévisions de la durée. Ce sont les causes associées qui sortent du lot dans le tableau. Comparativement au motif de référence (matériel roulant), les autres causes — soit la clientèle, les incidents d'émanation, et les pannes de causes diverses — semblent déterminantes ($p < 0.01$) dans l'effet sur la durée des interruptions, et ajoutent entre 1 et 10 minutes au calcul prévisionnel.

Malgré la faible significativité du modèle, deux mises en situation ont été élaborées à la figure 23 afin d'illustrer l'effet potentiel des différentes variables sur le calcul de la durée prévue.

Situation #1 🌞👤		Situation #2 🍁🔥	
Un incident au printemps, un jour de semaine à l'heure de pointe, sur la ligne orange. L'incident est dû à un objet sur la voie (Clientèle).		Un incident à l'automne, la fin de semaine à une heure creuse, sur la ligne jaune. Un dégagement de fumée a été repéré (Incident d'émanation).	
Variables	Coefficients	Variables	Coefficients
Printemps	-0.021	Automne	-0.019
Semaine	0.004	Fin de semaine	0.000
Heure de pointe	-0.016	Heure creuse	0.000
Ligne orange	-0.049	Ligne Jaune	-0.040
Clientèle	0.066	Incident d'émanation	0.381
Constante	0.927	Constante	0.927
SommeCoeff	0.911	SommeCoeff	1.250
Durée prévue (minutes)	8.15	Durée prévue (minutes)	17.80
10^SommeCoeff		10^SommeCoeff	

Figure 23. Mises en situation afin d'illustrer la prévision de la durée d'une interruption de service

Dans la première situation, un objet sur la voie entraîne une interruption de service à l'heure de pointe. Suite à une transformation du logarithme de durée en valeur absolue, on peut estimer que la durée prévue de l'interruption serait de 8 minutes, avec une confiance de 75% que l'interruption serait de 19 minutes ou moins. C'est le type de situations où, dans la plupart des cas, la majorité des passagers devraient rester à bord du train et attendre patiemment quelques minutes avant que la situation se règle afin de poursuivre leur itinéraire.

La deuxième situation est plus problématique, où un dégagement de fumée à une station provoque une interruption. Le modèle prévoit une panne de 18 minutes avec une confiance de 75% que l'interruption serait de moins de 30 minutes. Dans ce cas-ci, les clients du métro devraient considérer un itinéraire alternatif dépendamment de leur destination.

Ces mises en situation reflètent les possibles comportements des usagers lors d'incidents similaires, supportées par l'analyse de Lepage à Montréal (2019) qui révèle une augmentation de l'intensité d'usage des autres modes de transport lors d'interruption d'une durée de plus de 20 minutes. Ainsi, une régression binomiale a été réalisée en utilisant deux sous-ensembles avec les mêmes attributs: le premier avec les interruptions de moins de 20 minutes (2 198 observations) et le second incluant les interruptions de 20 minutes ou plus (244 observations). Ce modèle est robuste, mais les informations fournies — qui sont la probabilité qu'un incident mène à un arrêt de plus de 20 minutes — sont difficiles à communiquer.

Ces nouveaux résultats corroborent les constats précédents et montrent que quelques variables — précisément reliées au motif des arrêts et à la ligne touchée — affectent les deux sous-ensembles de moins et plus de 20 minutes. Avec un pseudo R2 de 0.113, ce nouveau modèle semble un peu plus puissant et adapté à la réalité du jeu de données. Sans permettre des estimations probabilistes plus précises sur la durée potentielle d'une interruption, les résultats peuvent être utilisés afin d'effectuer des prédictions par rapport à une alternative.

Rapport de cotes		Intervalle de confiance (95%)		
	Variables	Estimation	Min	Max
Saison	Printemps	0.959	0.668	1.375
	Été	0.807	0.548	1.187
	Automne	0.782	0.525	1.163
Jour	Fin de semaine	0.925	0.666	1.285
Période	Heure creuse	1.129	0.853	1.495
Ligne	Orange	0.561	0.350	0.897
	Verte	0.709	0.448	1.121
	Jaune	1.037	0.526	2.045
Cause	Clientèle	3.893	2.590	5.852
	Incident d'émanation	29.043	15.686	53.775
	Divers	3.656	2.382	5.612

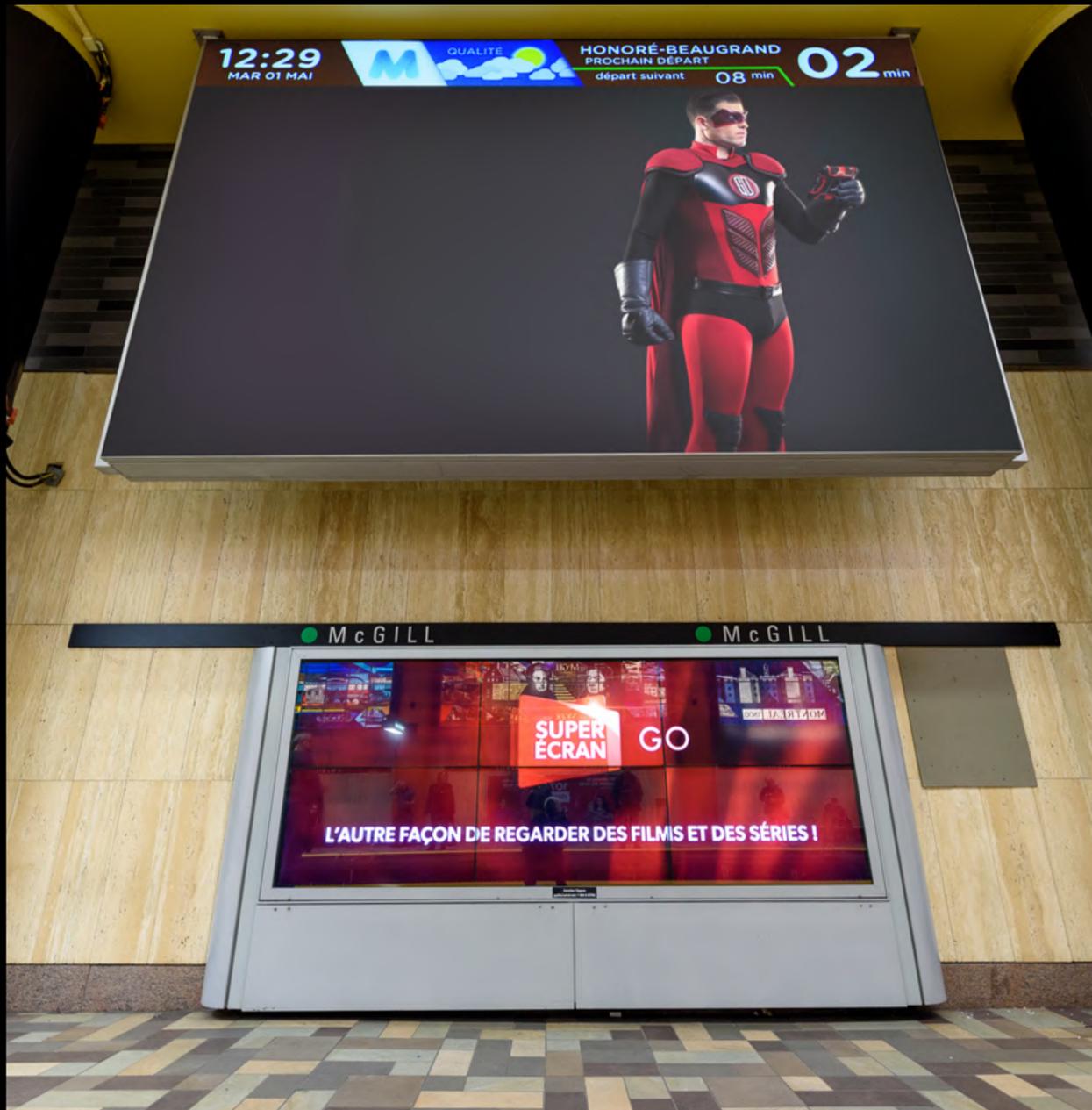
Figure 24. Tableau des estimations du rapport de cotes, avec un intervalle de confiance de 95%

Les variables significatives ($p < 0.001$) sont en gras à la figure 24. Les estimations du rapport de cotes peuvent être interprétées de manière à évaluer la probabilité d'un incident de durer plus de 20 minutes.

On peut interpréter les résultats significatifs de la figure 24 ainsi:

- Un incident au motif '**Divers**' à 3.7 fois plus de chance de durer plus de 20 minutes qu'un incident dû au matériel roulant;
- Un incident au motif '**Clientèle**' à 3.9 fois plus de chance de durer plus de 20 minutes qu'un incident dû au matériel roulant;
- Un incident pour **cause d'émanation** à 29 fois plus de chance de durer plus de 20 minutes qu'un incident dû au matériel roulant;
- Un incident sur **la ligne orange** à 0.56 fois moins de chance de durer plus de 20 minutes qu'un incident survenu sur la ligne bleue.

Finalement, avec les modèles établis et l'analyse effectuée, on peut conclure que les variables spatio-temporelles sont de faibles prédicteurs de la durée et qu'il est difficile de prévoir la durée d'une interruption avec un intervalle de confiance convenable qui pourra aider les usagers dans leur décision au moment venu. Les résultats les plus justes sont associés aux prédictions probabilistes qu'une interruption d'un certain attribut soit de plus ou moins 20 minutes. Néanmoins, il est évident à travers l'ensemble des modèles examinés que la cause d'un incident est le facteur le plus significatif affectant la durée potentielle. Or, comme nous le verrons dans les recommandations, l'optimisation des processus communicationnels, de l'identification de la cause jusqu'à l'information aux passagers, est centrale à l'amélioration de l'expérience client.



Écran sur le quai, station McGill
Photo: Bell Média, 2018

6. RECOMMANDATIONS

Ce projet de recherche portant sur les interruptions de service du métro révèle des apprentissages essentiels aux preneurs de décisions au sein des sociétés de transport, mais aussi aux fournisseurs et organisations qui collaborent aux opérations et à la mise en œuvre des stratégies de mobilité. À la lumière des recherches et des analyses effectuées, quatre recommandations ont été formulées afin de contribuer aux réflexions des diverses parties prenantes quant aux défis liés aux incidents dans les transports publics lourds.

6.1 Amélioration du processus d'identification des causes en début d'incident

Lors d'interruptions de service, les sociétés de transport — dont la STM — se basent sur des données historiques afin d'estimer l'heure de reprise pour une cause initiale déterminée. (STM, 2018). Ainsi, une identification rapide de la catégorie exacte des causes en début d'interruption est primordiale afin de non seulement communiquer la bonne information aux passagers, mais aussi de mobiliser les bonnes équipes dans la résolution des incidents.

L'analyse des données ouvertes disponibles révèle que, pour la période étudiée, plus de 27% des observations ont une cause initiale — attribuée au moment de l'incident par la STM — différente de la cause déterminée suite à l'analyse de l'incident. Cette donnée expose le défi associé à l'identification rapide et juste des causes en début d'interruption, puis aux enjeux d'information passager associés.

Toutefois, les raisons exactes de cette discordance sont incertaines. Plus de détails sur les processus internes de résolution sont nécessaires afin de renforcer les stratégies en place. Un nouvel indicateur de performance lié à l'identification juste et rapide des causes pourrait contribuer à mesurer la réduction de cet écart au sein des sociétés de transport.

6.2 Une stratégie d'information passager multicanale optimisée

La disponibilité d'information passager, notamment lors d'interruptions de service, est un facteur déterminant de la satisfaction des usagers (Gentile and Noekel, 2016). Ces derniers s'attendent à obtenir des informations dites "parfaites" sur les interruptions (Binder et al., 2017) et de pouvoir aisément trouver, lors d'interruptions prolongées, une alternative adéquate pour se rendre à destination (Cadarso et al., 2015).

En avril 2022, la satisfaction des usagers du métro de la STM était à son plus bas en deux ans, avec une satisfaction de 67% (STM, 2022). La raison principale mentionnée par les participants au sondage était la perturbation fréquente du service. Bien qu'une interruption ne dure que quelques minutes dans la grande majorité des cas, les usagers ont des attentes élevées d'obtenir des informations précises et régulières lorsqu'une interruption affecte le réseau.

Une analyse sommaire comparative des données de l'étude et des informations de la STM partagées sur Twitter sur la ligne orange révèle que de nombreuses interruptions de plus de 10 minutes n'ont pas eu d'estimations d'heure de reprise et n'ont pas été communiquées en dehors des trains et des stations du réseau. Même si la durée d'une interruption est difficile à prédire, les opérateurs de transports peuvent tout de même améliorer leur processus de deux façons:

1. Une communication multicanale effective: si une interruption persiste pendant plus de 10 minutes, l'opérateur devrait systématiquement diffuser cette information au sein de son réseau et via les canaux digitaux (Twitter, alertes dans les applications mobiles, bornes d'informations et afficheurs dans les autobus) afin de gérer les attentes des clients et de proactivement les aider à choisir un itinéraire alternatif.

2. Une hausse de la fréquence des messages à la clientèle: un standard pour plusieurs sociétés de transport dont la STM est d'informer les clients du statut de l'interruption à chaque cinq minutes. Or, des messages plus réguliers, à chaque deux ou trois minutes par exemple, permettraient d'apaiser le sentiment de frustration et d'incertitude des clients provoqués par l'interruption en cours (Caulfield et O'Mahony, 2007).

6.3 Des indicateurs de performance centrés sur la clientèle

La plupart des sociétés de transports adoptent des indicateurs de performance opérationnels, comme le nombre d'incidents de 5 minutes et plus par million de km. Un des enjeux avec cet indicateur populaire est qu'il est détaché de la réalité de la clientèle: un métro peut être vide lors d'incidents, ou encore complètement bondé. Ainsi, des indicateurs de performances complémentaires sont à considérer afin de centrer les interventions sur la clientèle: par exemple, en mesurant le nombre de passagers affectés par des incidents de plus de 5 minutes, ou en estimant la durée cumulée des retards pour la clientèle. Ces indices permettent de centrer les stratégies d'intervention sur les incidents qui affectent directement les passagers, et contribuent à une meilleure priorisation des équipes de réponse en fonction des causes déclarées.

6.4 Données ouvertes sur la performance et la fiabilité du réseau

La STM a une politique de données ouvertes novatrice comparativement à plusieurs sociétés de transport nord-américaines. Sur le portail de données ouvertes de la ville de Montréal, sept jeux de données historiques, notamment celui des incidents du réseau ainsi que celui du kilométrage réalisé, sont mis à jour périodiquement par l'agence de transport. Un jeu de données complémentaire qui serait pertinent pour les journalistes et chercheurs en transports serait précisément celui de la fiabilité du réseau. Un indicateur comme celui du pourcentage de passagers qui attendent la durée prévue entre les départs de bus ou de métro est essentiel à l'évaluation de la performance d'un réseau. La Massachusetts Bay Transportation Authority (MBTA) qui opère les transports publics dans la grande région de Boston a notamment mis à la disposition du public un tableau de bord avec des indicateurs de performance et des statistiques mis à jour quotidiennement (*mftabackontrack.com*). Une approche similaire pour la STM permettrait une plus grande transparence et indiquerait un engagement plus ferme envers les efforts dédiés à l'amélioration de la fiabilité du réseau.



Train Azur
Photo: Ilia Usmanov

7. CONCLUSION

Les objectifs de ce projet de recherche étaient doubles: d'abord dresser le portrait des interruptions de service du métro de Montréal, puis déterminer si les données ouvertes peuvent être utilisées afin de prédire la durée des interruptions de service.

L'analyse préliminaire des données contextuelles met en lumière certaines tendances peu étudiées du réseau de métro montréalais. On note une diminution du nombre total d'incidents entre 2014 et 2021 malgré une hausse du nombre de kilomètres parcourus. Les incidents causés par la clientèle et le matériel roulant représentent ensemble 72% du nombre total d'incidents et 63.2% de la durée cumulée des interruptions pour la période d'étude. Quant à l'emplacement, la grande majorité des incidents (84%) entre 2019 et 2021 ont affecté les lignes orange et verte.

Le fait que la plupart (66%) des interruptions sont résolues en moins de 10 minutes est un indicateur du travail accompli par la STM afin d'optimiser les processus d'intervention des différentes équipes lors d'incidents. Il est néanmoins difficile de comparer les résultats de la STM avec ceux d'autres agences de transport vu le manque de données disponibles.

Ce projet de recherche tentait aussi d'évaluer si les données ouvertes libérées par la STM peuvent être utilisées afin de prédire les durées des interruptions de service de manière à aider les utilisateurs du métro à effectuer des choix d'itinéraires éclairés plus rapidement. Les modèles développés révèlent que la plupart des informations spatio-temporelles disponibles en début d'interruption sont de faibles prédicteurs de la durée des incidents: ainsi, la saison, la période, la journée et le lieu n'ont que très peu d'effet sur la durée. La cause associée est sans équivoque le prédicteur le plus significatif, avec les incidents pour cause d'émanation entraînant les durées prédites les plus longues. La régression binomiale renforce les résultats du premier modèle et indique que les incidents d'émanation ont 29 fois plus de chance de durer plus de 20 minutes qu'un incident associé au matériel roulant.

Les analyses effectuées démontrent la difficulté pour toute société de transport de prédire avec un niveau de confiance élevé la durée des pannes de service des modes de transport lourds, au grand dam des usagers. Des données granulaires concernant les interruptions

couvrant une plus longue période pourraient permettre d'effectuer des analyses prédictives plus robustes.

Néanmoins, plusieurs actions peuvent être mises en place afin de consolider les stratégies des acteurs de transport et d'améliorer l'expérience de la clientèle lors d'interruptions. Les avancées technologiques permettent de multiplier les canaux de communication aux usagers en temps réel: les sociétés de transport devraient utiliser tous les moyens à leur disposition afin de renforcer le lien de confiance et réduire l'incertitude lors de pannes de service.

Finalement, de nouveaux indicateurs de performance centrés sur l'impact des interruptions sur les usagers plutôt que sur les infrastructures des réseaux sont essentiels au développement de stratégies de transport collectif axées sur l'humain. Des efforts soutenus de l'ensemble des parties prenantes sont indispensables à l'amélioration de la fiabilité des réseaux afin de surmonter le biais de négativité de la population et favoriser l'adoption du transport collectif.

BIBLIOGRAPHIE

- Barron, A., Canavan, S., Anderson, R., Cohen, J., 2018. Operational Impacts of Platform Doors in Metros. *Transportation Research Record* 2672, 266–274. <https://doi.org/10.1177/0361198118784386>
- Barron, A., Melo, P.C., Cohen, J.M., Anderson, R.J., 2013. Passenger -Focused Management Approach to Measurement of Train Delay Impacts. *Transportation Research Record* 2351, 46–53. <https://doi.org/10.3141/2351-06>
- Binder, S., Maknoon, Y., Bierlaire, M., 2017. The multi-objective railway timetable rescheduling problem. *Transportation Research Part C: Emerging Technologies* 78, 78–94. <https://doi.org/10.1016/j.trc.2017.02.001>
- Cadarso, L., Maroti, G., Marin, A., 2015. Smooth and Controlled Recovery Planning of Disruptions in Rapid Transit Networks. *IEEE Trans. Intell. Transport. Syst.* 16, 2192–2202. <https://doi.org/10.1109/TITS.2015.2399975>
- Caulfield, B., O'Mahony, M., 2007. An Examination of the Public Transport Information Requirements of Users. *IEEE Trans. Intell. Transport. Syst.* 8, 21–30. <https://doi.org/10.1109/TITS.2006.888620>
- Diab, E., Shalaby, A., 2018. Subway Service Down Again? Assessing the Effects of Subway Service Interruptions on Local Surface Transit Performance. *Transportation Research Record* 2672, 443–454. <https://doi.org/10.1177/0361198118791665>
- Donovan, J., 2015. Overcoming the tyranny of distance: making walking and cycling possible and preferable. *Proceedings of the Institution of Civil Engineers - Municipal Engineer* 168, 115–124. <https://doi.org/10.1680/muen.14.00014>
- Gentile, G., Nökel, K., 2016. Modelling public transport passenger flows in the era of intelligent transport systems. *Springer Tracts on Transportation and Traffic* 10.
- Krishnakumari, P., Cats, O., van Lint, H., 2020. Estimation of metro network passenger delay from individual trajectories. *Transportation Research Part C: Emerging Technologies* 117, 102704. <https://doi.org/10.1016/j.trc.2020.102704>
-

- Leng, N., Corman, F., 2022. Communicating delays and adjusted disposition timetables: Modelling and evaluating the impact of incomplete information to passengers. *Expert Systems with Applications* 191, 116265. <https://doi.org/10.1016/j.eswa.2021.116265>
- Leng, N., Corman, F., 2020. The role of information availability to passengers in public transport disruptions: An agent-based simulation approach. *Transportation Research Part A: Policy and Practice* 133, 214–236. <https://doi.org/10.1016/j.tra.2020.01.007>
- Lepage, S., 2019. Modélisation des interactions entre les modes de transport aux abords des stations de métro. Université de Montréal 173.
- Louie, J., Shalaby, A.S., Habib, K.M.N., 2016. Modelling Disruption Duration for Toronto's Subway System: An Empirical Investigation Using Lognormal Regression and Hazard Models. Presented at the Transportation Research Board.
- Melo, P.C., Harris, N.G., Graham, D.J., Anderson, R.J., Barron, A., 2011. Determinants of Delay Incident Occurrence in Urban Metros. *Transportation Research Record* 2216, 10–18. <https://doi.org/10.3141/2216-02>
- Pender, B., Currie, G., Delbosc, A., Shiwakoti, N., 2014a. International Study of Current and Potential Social Media Applications in Unplanned Passenger Rail Disruptions. *Transportation Research Record* 2419, 118–127. <https://doi.org/10.3141/2419-12>
- Pender, B., Currie, G., Delbosc, A., Shiwakoti, N., 2012. Planning for the unplanned: an international review of current approaches to service disruption management of railways.
- Piner, D., Condry, B., 2017. International best practices in managing unplanned disruption to suburban rail services. *Transportation Research Procedia* 25, 4403–4410. <https://doi.org/10.1016/j.trpro.2017.05.331>
- Rahimi, E., 2020. Transportation Resilience: Behavioral Impact Analysis of Disruptions and Pandemic on Public Transit (PhD Thesis). University of Illinois at Chicago.
- Rahimi, E., Shamsiripour, A., Shabanpour, R., Mohammadian, A., Auld, J., 2019. Analysis of transit users' waiting tolerance in response to unplanned service disruptions. *Transportation Research Part D - Transport and Environment* 77, 639–653. <https://doi.org/10.1016/j.trd.2019.10.011>
-

AUTRES RÉFÉRENCES

The Community of Metros (COMET)

<https://communityofmetros.org/>

Données ouvertes — Ville de Montréal

Incidents du réseau du métro

<https://donnees.montreal.ca/societe-de-transport-de-montreal/incidents-du-reseau-du-metro>

Le Devoir — Plus de 4000 interruptions de service depuis quatre ans dans le métro

Publié le 3 juin 2019

<https://www.ledevoir.com/societe/transports-urbanisme/555860/plus-de-4000-interruptions-de-service-depuis-quatre-ans-dans-le-metro>

Le Devoir — Portrait des pannes de métro

Publié en mai 2019

https://www.ledevoir.com/documents/special/19-05_pannes_STM/index.html

The jamovi project (2021). jamovi (Version 1.6) [Computer Software].

<https://www.jamovi.org>

Société de transport de Montréal — L'ABC des interruptions de service

<https://www.stm.info/fr/infos/conseils/les-arrets-de-service/labc-des-interruptions-de-service>

TWINT - Twitter Intelligence Tool

<https://github.com/twintproject/twint>

