STUDY OF AUTOMATED SHUTTLE INTERACTIONS IN CITY TRAFFIC USING SURROGATE MEASURES OF SAFETY

Étienne Beauchamp, M.A.Sc candidate
Civil, Geological and Mining Engineering Department
Polytechnique Montréal, C.P. 6079, succ. Centre-Ville
Montréal (Québec), Canada H3C 3A7
Email: etienne-2.beauchamp@polymtl.ca
ORCID: 0000-0001-5549-3429

Nicolas Saunier, Ph.D., Professor
Civil, Geological and Mining Engineering Department
Polytechnique Montréal, C.P. 6079, succ. Centre-Ville
Montréal (Québec), Canada H3C 3A7
Email: nicolas.saunier@polymtl.ca
ORCID: 0000-0003-0218-7932

Marie-Soleil Cloutier, Ph.D., Professor
Urbanisation Culture Société Research Centre
Institut National de la Recherche Scientifique
385, Sherbrooke Est, Montréal (Québec), Canada H2X 1E3
Email: marie-soleil.cloutier@ucs.inrs.ca
ORCID: 0000-0002-8533-4784

Word Count: 4960 words + 2 table(s) $\times$ 250 = 5460 words

Submission Date: August 1, 2020

ABSTRACT

Driving automation is happening at a rapid pace, with different driver assistance systems already available in mass-market cars. However, this rapid development in driving automation leads to concerns and questions about their impact on safety, in particular for vulnerable road users. While previous studies have been restricted to incident reports and simulation tools, the safety of automated vehicles (AVs) is not clearly demonstrated. Instead of crashes, which are extremely rare events, this study uses surrogate measures of safety (SMoS) to analyze the interactions between road users and low-speed automated shuttles that circulated in Montréal and Candiac, in Canada, during two pilot projects in mid and late 2019. Cameras were placed at seven intersections along the routes of the shuttles. More than 70 hours of footage were processed to extract the road user trajectories using computer vision techniques and compute various safety indicators: speed, acceleration, time headway, time-to-collision (TTC) and post-encroachment time (PET). The Kolmogorov–Smirnov test was used to compare the distributions of interactions involving AVs with the distributions of interactions involving motorized vehicles following paths similar to those of the AVs. The results indicate that these automated shuttles behave generally more safely: their speeds and accelerations are lower and their interactions are characterized by higher TTCs and PETs, notably with vulnerable road users. However, small headway times at one site with high speed differentials between the shuttles and other following vehicles raise concerns that warrant further research into the suitable context for these vehicles.

Keywords: Road Safety, Automated Shuttles, Surrogate Measures of Safety, Video Analysis

INTRODUCTION

Road safety has a considerable impact on public health. Road traffic injuries were among the ten leading causes of death in 2016, with vulnerable road users being overly impacted in the toll (*1*). Moreover, road crashes have repercussions on society in many ways, from the fatalities and injuries to the healthcare costs and impact on congestion. In Canada, driver error was estimated to be a contributing factor in 85 % of road casualties in 2018, mainly because of speeding, distracted driving and impaired driving, respectively in 23 %, 22 % and 19 % of fatal crashes (*2*) and it is estimated that automated vehicles (AVs) could lead to benefits of $ 65 billion per year (*3*).

Among the various ways to improve safety, there has been a lot of interest recently for vehicular technologies, namely advanced driver assistance technologies (ADAS). There has been a growing focus on driving automation in particular during the last decade, with the promise that driverless vehicles will eliminate road crashes. In the meantime, vehicles with varying levels of automation are tested and some are already available on the market. Despite all these developments, the impact of ADAS and driving automation is difficult to evaluate in particular on safety (*1*). Recent improvements in the design of vehicles or road infrastructures might play a bigger role in the decline in traffic injuries in developed countries and result in overestimating the impact of ADAS on safety (*4*). Data on crashes with existing partially automated vehicles (AV) is very rare. The few existing studies on the safety of AVs have mostly examined disengagement reports from AV testing programs, mainly coming from California's Department of Motor Vehicles (DMV) (*5–7*).

While the most common way to study safety is through the analysis of crash data, that approach has many shortcomings (*8*) such as its reactive nature and the low frequency of crashes, which are particularly acute for new technologies with no or limited penetration. In the 1970s, several proactive methods to diagnose safety were developed, initially focused on the detection and characterization of severe traffic conflicts (*9*). There was a renewed interest in the 2000s as new technologies became available to collect data more efficiently and objectively. Traffic conflicts and other events of interest for safety diagnosis are more frequent than crashes, and their observation provides more insight about the traffic processes that may lead to crashes (*8, 10*). The number of such events and other surrogate measures of safety (SMoS) can be used to assess safety more quickly, which is particularly suitable for new technologies like ADAS and AVs, as documented crashes involving AVs are currently scarce.

The objective of this paper is to study the safety of automated low-speed shuttles in real traffic using video data and SMoS. To the authors' knowledge, analyzing video recordings of AVs in an urban setting, under real conditions, and assessing the safety through the use of SMoS has never been done before. Video was recorded at three different sites in Montréal, Canada, during the summer of 2019 and at four different sites in Candiac (a suburb on the South Shore) during the fall of the same year using a portable installation (*11*). The analysis was conducted using several safety indicators; namely speed, acceleration, time headway, time-to-collision (TTC) and post-encroachment time (PET).

The following sections present respectively the literature review, a detailed description of the methodology, the collected data, as well as the experimental results and, finally, the closing remarks of this paper.

LITTERATURE REVIEW

Road Safety Diagnosis Methods

Road safety studies generally rely on one of three main categories of data (*12, 13*): 1. crash data; 2. self-reported crashes, and 3. near-crashes and non-crash observations.

Methods in the first category are the most common and the traditional way to diagnose road safety (*12, 13*). Countless studies at various levels, from individual sites to whole countries, rely on crash data. Although it seems natural to assess safety based on historical crash data, it suffers from many shortcomings, from the quality of the data to its biases for example towards the most severe crashes or some types of road users, to the intrinsic issue that it is a reactive approach that requires to wait for crashes to occur before addressing their causes (*8, 10*). Furthermore, the need to wait for long periods of time to collect sufficient amounts of data makes crash data particularly inappropriate to evaluate new trends and fast evolving technologies.

Information on crashes is collected by various organizations, primarily the police, but there is a growing interest for self-reported crashes and near-crashes. Such data has some advantages over traditional crash data collection: it can include information about other, non-crash, events that may be relevant to safety and more information than is typically collected in crash reports; it can be tailored to specific needs when designing the survey or data collection method; and data is available more quickly, possibly on a continuous basis. Although complementary to historical crash data, it still shares most of the shortcomings of crash data.

The last category of data encompasses various kinds of data, particularly traffic events, that are shown or believed to have a relationship to crash occurrence and severity. Traffic conflicts have received the most attention since the late 1960s, with the development of several traffic conflict techniques (TCT), such as the Swedish Traffic Conflict Technique and the Dutch Objective Conflict Technique for Operation and Research (DOCTOR) (*8*). Human observers were trained to identify conflicts in traffic and rate their severity, which is time-consuming, costly and subjective. Recent progress in various sensor technologies, in particular the affordability of quality video recording equipment and computer vision, makes fully or semi-automated video analysis for safety analysis possible (*14–16*). The various types of traffic events and their relationship to safety are famously represented in the safety pyramid popularized by Christer Hydén (*8*), with crashes at the top, the most severe and rare events, and normal traffic at the bottom. The most common SMoS is the number of severe traffic conflicts, or near crashes / misses (*8, 17*). Severity is measured through safety indicators like the speed, TTC and PET (*8*). Conflicts and interactions with lower severity levels may also be interpreted in a safety perspective (*18*).

Studying non-crash events and using SMoS has several advantages such as short data collection periods and the richer data and insights provided by the direct observation of the complete traffic process, on the contrary to what is available in crash reports (*8*). Although some studies have shown correlations between SMoS like the number of severe traffic conflicts and safety (*17, 19–23*), SMoS are still not as widely used as measures derived from crash data.

Safety Assessment of AVs

Even if individual components of AV technology have been extensively tested and their reliability has been proven, AVs as a whole are still a new technology with limited testing in the various real world conditions (*6*). AV safety assessment poses challenges in several areas (*24, 25*) - whether it be the certification of hardware, software or human-machine interface (HMI) - and a single solution to certify AV safety does not exist.

Current AV Testing Frameworks

SAE International issued in 2014 the J3016 standard (*26*) that categorizes the different levels of driving automation. While influential, the document is descriptive and has no legal value. In 2017, Mobileye (subsidiary of Intel) proposed the Responsibility-Sensitive Safety (RSS) (*27*), a mathematical model to help standardize requirements to certify AV safety. However, up to this day, no standards nor international entity exists to regulate and assess the safety of driver assistance technologies.

In the state of California, where many tests of AVs on public roads are conducted, the department of motor vehicles (DMV) requires manufacturers who wish to conduct said tests to obtain permits (*28*). While manufacturers have to report crashes engendering property damage, bodily injury or death (*29*), requirements were found to be too broad and vague to draw clear conclusions on the safety of AV (*5, 6*).

Additionally, AV manufacturers do not provide their data on reported incidents and the DMV reports are the only publicly available datasets known to the authors. In Canada, there is no regulatory requirements for driver automation technologies and it is up to provincial and territorial governments to oversee regulations regarding road operations (such as driver licensing, vehicle registration, motor vehicle insurance and liability, vehicle maintenance standards and traffic laws) (*30*). According to the jurisdiction of the province or territory where they are conducted, manufacturers must: receive permits and any other required authorization; report disengagement, crashes and incidents (*31*). However, Quebec's Highway Safety Code includes no mentions of permits or crash reports (*32*) and no publicly available data on AV incidents or crashes were found by the authors.

Existing AV Safety Studies

Previous studies conducted to evaluate AV safety have mostly used two approaches: 1. miles driven and disengagement reports; and 2. data from microscopic traffic simulations.

AV manufacturers have notably used mileage has an argument to prove the safety of their vehicles' features such as intelligent cruise control and lane assist (*33*). However, such claims require to drive considerable distances, not to mention the fact that distance driven does not take into account the surrounding traffic and environment and the complexity of the driving task (*27*).

Manufacturers who are running tests in California are obliged to report automated disengagement and crashes. Some researchers have explored this publicly available data and found that AVs encountered more frequent crashes per miles driven than human drivers (*6, 7, 34*) and that a strong correlation exists between the number of miles driven and the number of disengagement encountered (*5*). Nonetheless, it was indicated that reports had inconsistencies and a lack of standardization among manufacturers made the analysis and comparison difficult (*5, 6*). This is to be expected, as no strict rules or requirements seem to have been given to AV developers.

While some real-world tests have been conducted in several countries, the scarce data available from these tests has led to the prominence of simulation-based safety assessment studies (*35*). Different studies using the VISSIM microscopic traffic simulation software have shown that the AV safety benefits would depend on the penetration rate of the said vehicles (*36, 37*). Some have argued that crash rates involving conventional users would not necessarily be proportional to the penetration rates (*38*). But invariably, AVs were found to reduce the number of conflicts compared to the scenarios where they are absent and that a full penetration would result in the greatest safety benefits (*36–38*).

Yet, the calibration of driver behavior models can be limiting due to the lack of empirical data (*36, 37, 39*) and it is even more difficult for safety studies since they require to represent the various factors and chains of events, including human factors, that can lead to a crash. Confirming that a simulator accurately represents reality is just as complex as validating the driving policy itself (*27*). Besides, simulation assumes AVs to function as expected on the road, when such vehicles, which are highly dependent on sensors, might encounter problems related to the weather or road infrastructure and occasional unpredictable events (e.g. road repair).

There have been a few real world studies of AV safety by researchers, including some of the pilot projects involving automated shuttles similar to the one presented in this work. Past pilot projects presented little challenge to the AVs as the AVs often run off-road, interacting only with pedestrians. These few published studies generally rely on qualitative data without any systematic evaluation of conflicts and safety indicators like speed, TTC and PET (*40, 41*). To the authors' knowledge, SMoS have yet to be used to evaluate the interactions of AVs with other road users in real city traffic.

METHODOLOGY

The methodology employed is summarized in Figure 1. The steps are discussed in each subsection.

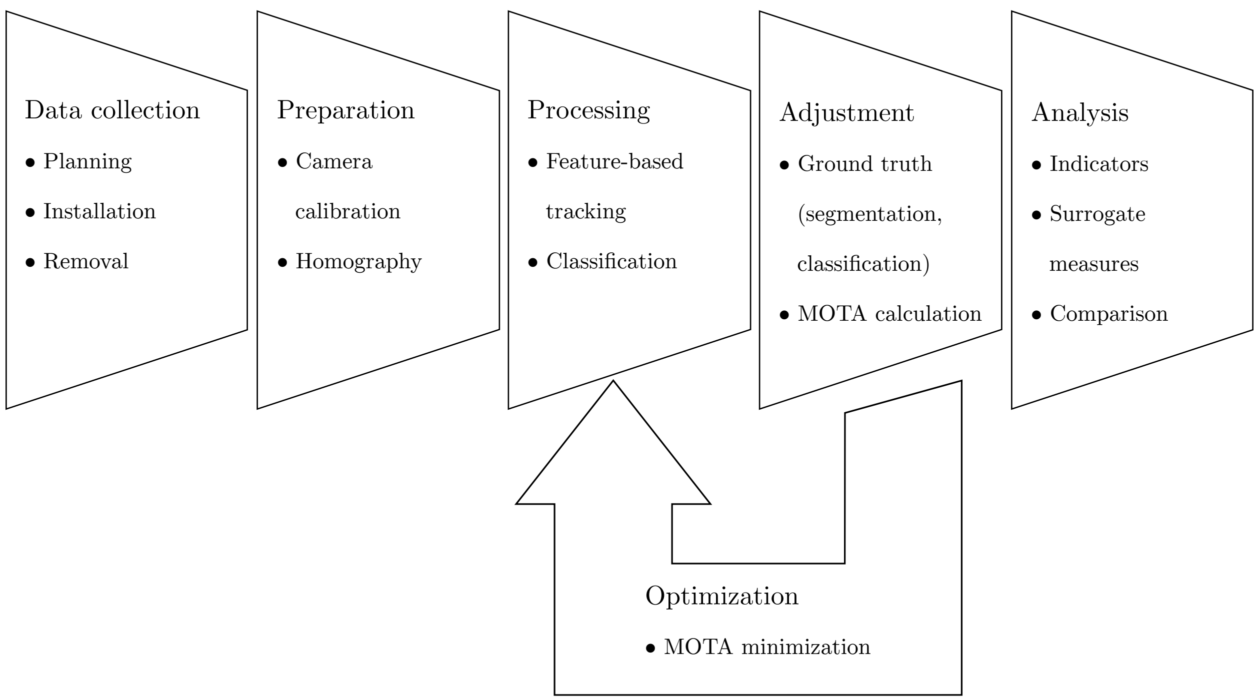

FIGURE 1 Methodology overview.

Site Selection and Video Data Collection

This study takes place in the context of two pilot projects of transit service provided by AV shuttles in the cities of Montréal and Candiac (on the South Shore of the Island of Montréal) in 2019. The AV shuttles are an EZ10 by EasyMile in Montréal and an AUTONOM shuttle by Navya in Candiac. To collect the data needed for this research, both routes shown in Figure 3 were first visited before the start of the projects. Three factors needed to be met to select the data collection locations. Sites had to:

1. include a lamp post or similar pubic structure to attache the video recording equipment (pole and camera);
2. have as few trees and other obstructing objects as possible, to have a clear view of traffic from the camera's viewpoint;
3. provide an interesting view of the shuttle's interaction with other road users, including intersections for more varied movements and maneuvers.

A GoPro camera was installed on a height-adjustable pole attached to a lamp post at each sites. In Montréal, video data was collected during the summer of 2019, for seven days in July and one day in August, between 10:00 AM and 6:00 PM. In Candiac, video data was collected during two days in November 2019 and four days in December 2019, during various time intervals, depending on the shuttle schedule. The data was prepared by correcting the camera lens distortion and computing the homography matrix, which is used to convert the road user coordinates from the image space to real world coordinates at the ground level.

Data Preparation and Processing

The video analysis software from the open source "Traffic Intelligence" project was used to extract road user trajectories and classify them (*11*). All moving objects within a user-defined area in the camera field of view are detected and tracked using the available feature-based tracker (*14*), then classified in three categories as a pedestrian, cyclist or motorized vehicle based on their speed and appearance.

The many tracking parameters are usually adjusted for each site by trial and error (*11, 16, 42*). To further improve the tracking performance, a first dry-run of tracking was made on 30-min videos for each day and the results were manually verified and corrected. The resulting ground truth trajectory databases for the selected videos were then used as an input to adjust the tracking parameters by optimizing the measure of tracking accuracy (MOTA) (*43*) using the Mesh Adaptive Direct Search algorithm (MADS) available in the open source NOMAD tool (*44, 45*). Also, the road user speeds from the ground truth trajectory databases were used to update the user classifiers.

Once the videos were automatically processed with the optimized tracking parameters, the trajectories were manually cleaned by removing false alarms and merging duplicated trajectories for the same road user. The shuttle trajectories were manually annotated with a fourth "automated" road user category.

Trajectories were then clustered using the longest common subsequence similarity (LCSS) in the algorithm available in the "Traffic Intelligence" project, with each cluster being represented by a "prototype" trajectory (*15, 46, 47*). The LCSS parameters are: the similarity measure between user positions, set to the Manhattan distance with a 2-m threshold in this study, and a minimum similarity threshold over the LCSS for trajectories to be clustered together, set to 50 % in this study. The trajectory clusters are used to predict road user motion to compute the TTC and to identify motorized users with trajectories similar to the AV shuttles for comparison.

The manually annotated AV trajectories were used to identify which users followed similar paths. By identifying the clusters these AV trajectories belong to, the other motorized users in these clusters were designated as control users or vehicles. This follows the safety analysis method by movement patterns presented in (*47*). In this way, the safety of AVs can be compared to the safety of human drivers following similar paths. Figure 2 shows an example of the learnt prototypes and the subset of prototypes associated with AVs.

FIGURE 2 Cluster prototypes associated AVs (in green) and other users (in red) at the Montcalm & Inverness site (origins marked with a circle).

Safety Analysis with SMoS

An interaction is created when two road users coexist in the area of study within the field of view. Various positional and velocity-based indicators are extracted and can be used to categorize the interactions at each instant (see (*15*) for more details). Interactions that are mostly in the head-on category were filtered out, as they generated severe safety indicator values that did not reflect their actual safety.

Only the interactions involving at least an AV shuttle or a control vehicle were analyzed: these interactions are split in two subsets. The first is the set of AV interactions, i.e. interactions of one shuttle with another type of road user, since there is only one shuttle at a time. The second is the set of control vehicle interactions that encompasses all other interactions of a control vehicle with any other type of road user (except AV shuttles, counted in AV interactions). The following safety indicators were computed for all interactions: speed, acceleration, TTC and PET. Speed and acceleration are derived from the road user positions, using the Savitsky-Golay filter for smoothing. TTC is "the time until a collision between the vehicles would occur if they continued on their present course at their present rates" (*8, 9*). PET is "the time between the moment that the first road user leaves the path of the second and the moment that the second reaches the path of the first" (*8*). TTC and PET are computed using a distance threshold of 1.7 m to account for the average road user size. A method to predict the future positions is necessary to compute TTC at each instant: instead of assuming that road users keep moving with constant speed and direction, the probabilistic framework first presented in (*48*) was used with the learnt trajectory prototypes to account for the various paths a road user may take.

Finally, the time headway was also computed for rear-end interactions (car following situations) at three sites in Candiac on the Montcalm Boulevard where there are few turning movements and car following could be observed for sufficient amounts of time (time headway is PET for car following situations). Although time headway is usually measured at a fixed location on the road, it was computed continuously in this study: for each position x_l of a leading user at instant t_1 in a rear-end interaction, the time t_2 at which the following user reached x_l was recorded to compute the time headway $t_2 - t_1$.

All these safety indicators except PET are continuously measured and are aggregated to characterize each interaction with a single value: the mean speeds and accelerations are computed over the trajectory, while the 15th centile is used for TTC (denoted TTC_{15}) and the time headway (denoted h_{15} to reflect the minimum values and avoid outliers (*49*). The nonparametric Kolmogorov-Smirnov (KS) test is used to determine whether the indicator distributions for the AV shuttles and the control vehicles are different.

EXPERIMENTAL RESULTS

To minimize the computing time, videos before and after the AV shuttles' operating hours were not processed. The tables 1 and 2 summarize the information about the sites in Montréal and Candiac, the amount of processed video data, the number of trajectories per road user category, site characteristics and the number of interactions. The shuttle routes are shown in Figure 3. Sample frames from the video data are shown for each site in Figure 4. All indicator distribution plots, or "violin" plots, for speed, acceleration, time headway, TTC and PET such as in Figure 5 include the number of observations ("n:") and the three quartiles represented as dashed lines.

TABLE 1 Summary of processed video data collected in Montréal and Candiac per site.

Sites	Hours	#AV	#Cars	#Ped.	#Cyc.	Comments
Montréal						
Letourneux & Coubertin	≈11h50	104	5474	2516	408	Traffic light, shuttle makes a turn
Letourneux & Hochelaga	≈16h40	84	22319	1373	440	Traffic light
Letourneux & Ontario	≈14h10	125	2693	724	224	Shuttle turns into back alley
Candiac						
Montcalm & Residence	≈8h00	48	2987	37	0	Stop sign, shuttle stops in front of a retirement home
Montcalm & Inverness	≈10h20	61	3623	107	29	Stop sign on Inverness
Montcalm & Victorin	≈9h25	91	7717	71	30	Traffic light
Montcalm & Rail	≈5h40	42	1786	39	2	Railway crossing

TABLE 2 Summary of the number of interactions in Montréal and Candiac.

Cities	# AV Interactions	# Control Vehicle Interactions
Montréal	778	1318
Candiac	507	707

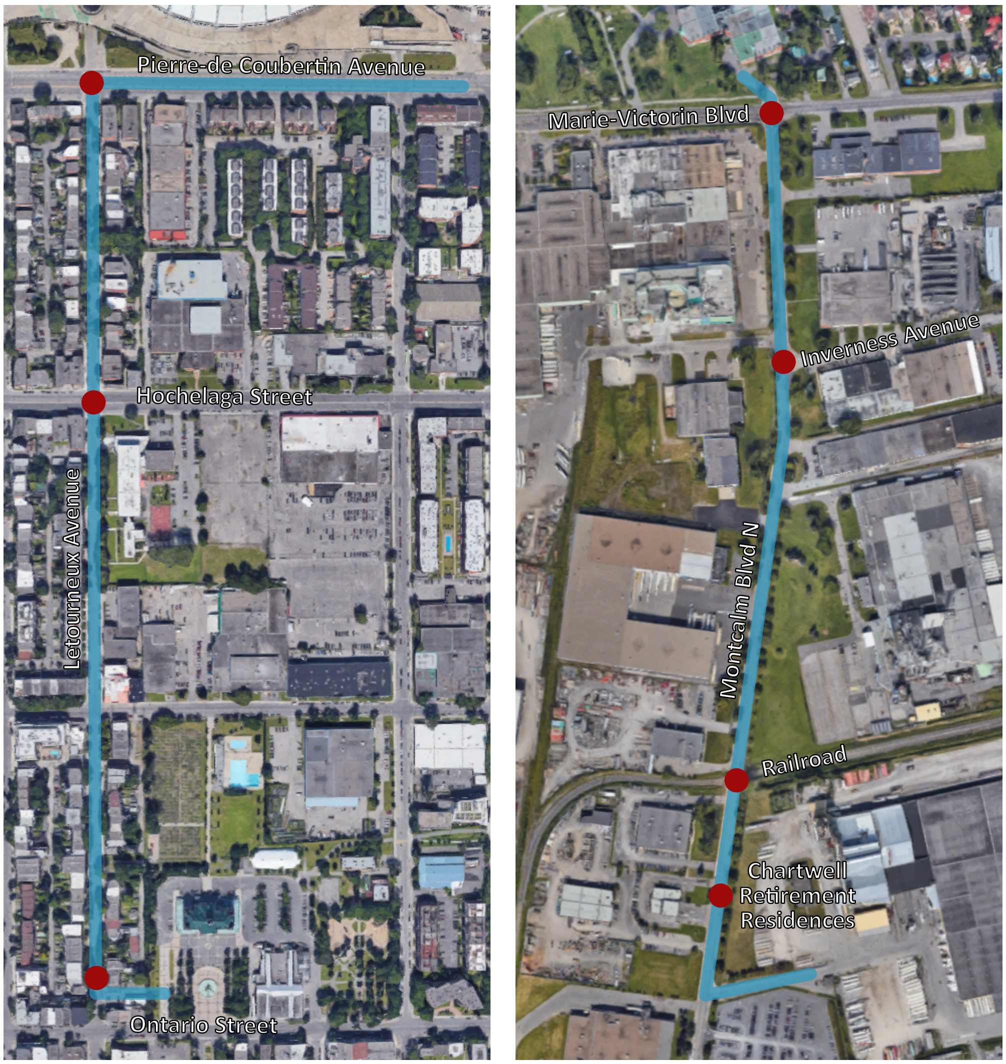

FIGURE 3 Maps of the routes (in blue) and selected video data collection sites (red dots) in Montréal (left) and Candiac (right).

Speed and Acceleration

The Figures 5 and 6 show respectively the distributions for the mean user speed and acceleration for AV shuttles and control vehicles per site. The mean, standard deviations and all quartiles of the shuttle speeds are much lower than those of the control vehicles. This is confirmed by the KS test which is significant for all sites (p-value < 10^{-10}). All the speeds, independently for control vehicles and shuttles, are slightly higher at the sites in Candiac compared to Montréal, which may

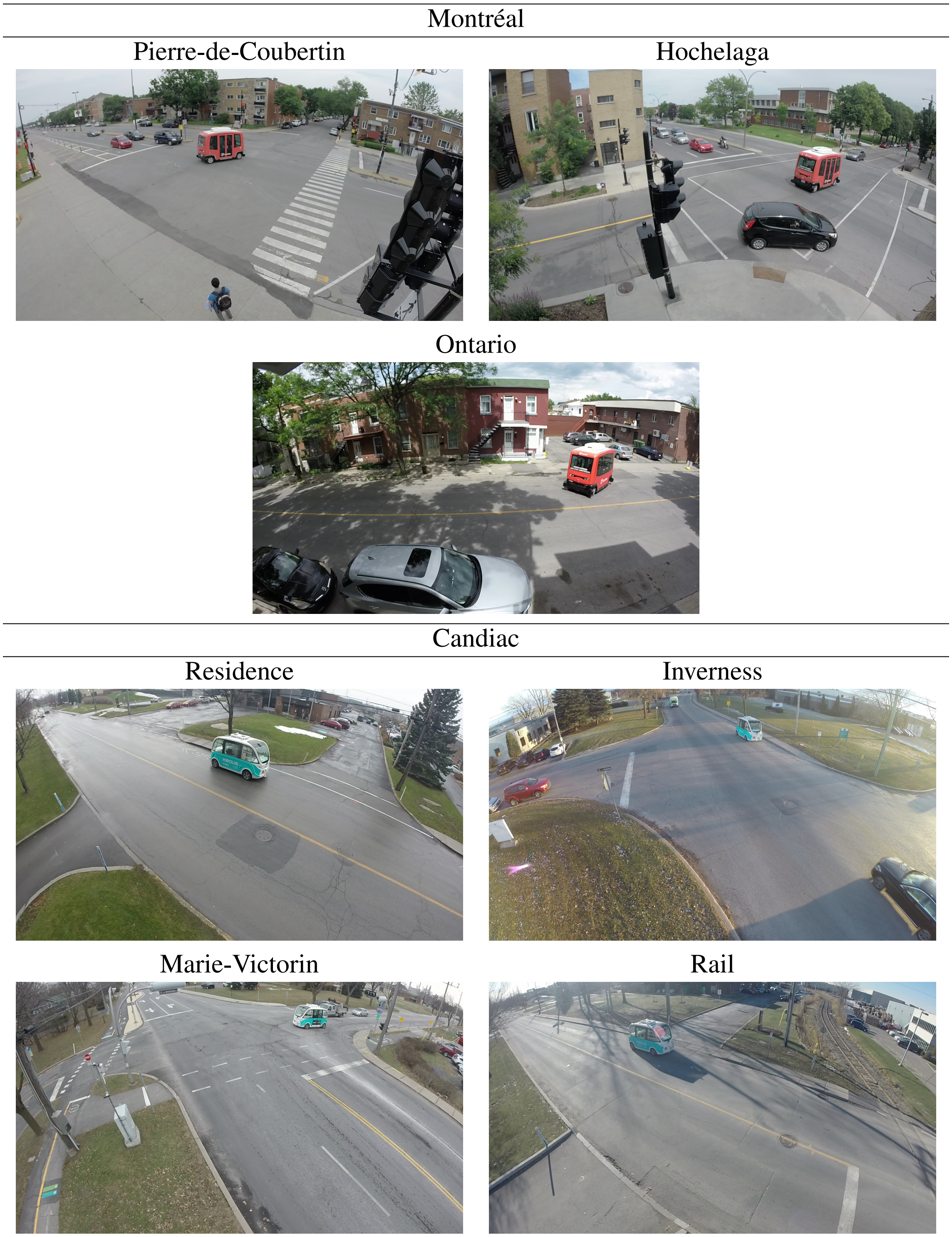

FIGURE 4 Frames from the collected video data at each site with the AV shuttle visible.

be attributed to the wider road and lighter traffic.

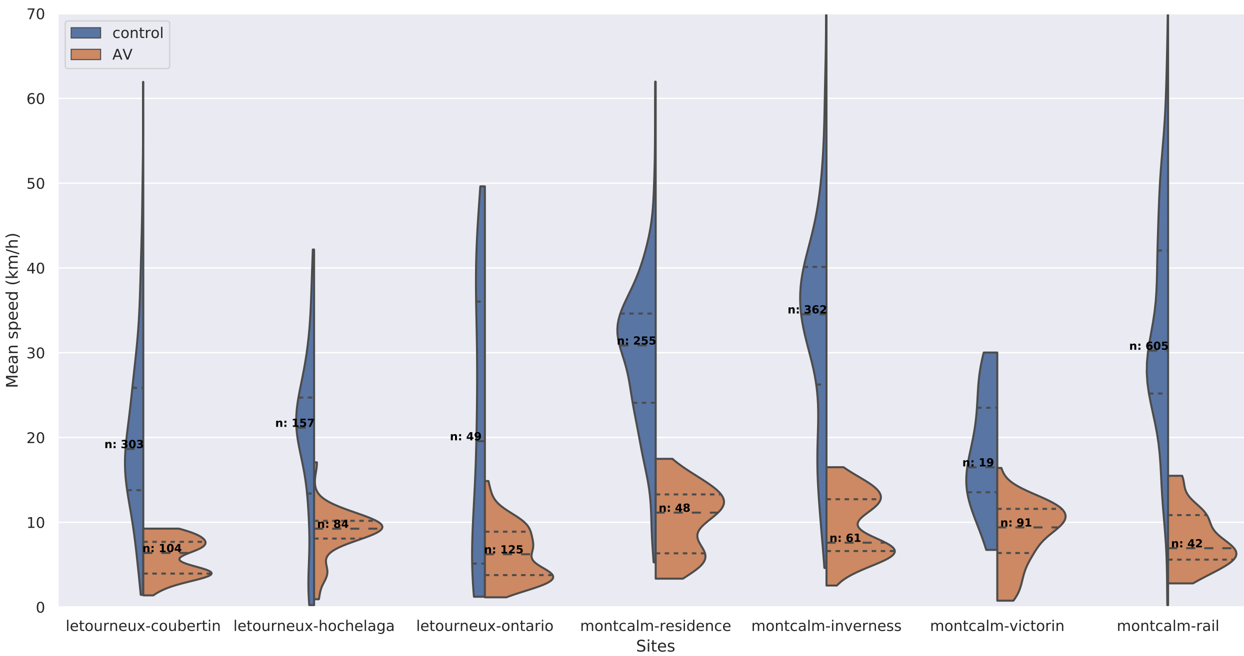

FIGURE 5 Mean user speed distributions of control vehicles and AV shuttles at each site.

Absolute mean accelerations and the acceleration standard deviation are also lower for the AV shuttles, which is consistent with their more predictable and cautious behavior. The means and the medians (second quartile) of the shuttle accelerations tend towards 0, which is also consistent with the tendency of shuttles to keep a constant speed and make less abrupt accelerations. As with the speed distributions, the results of the KS tests are statistically significant for all the sites (p-value < 10^{-5})

Time Headway

Figure 7 displays a side-by-side comparison of time headways (15th centile h_{15}) for rear-end interactions with either a leading control vehicle or AV shuttle. The comparison is different at the three sites: while there is no statistical difference for the KS test between the h_{15} distributions at the rail crossing (D = 0.2074, p-value = 0.2199), the difference is statistically different at the two other sites (Inverness: D = 0.209, p-value = 0.00443 and Residence: D = 0.235, p-value = 0.02049), albeit in opposite directions. Headway times for interactions with an AV leader tend to be smaller (all quartiles) at the Inverness intersection, while it is the opposite in front of the retirement home. This is directly related to the speeds (see Figure 5): the Inverness intersection has the highest differential between the control vehicle and AV speeds, which can lead to driver impatience as they are forced to follow the shuttle if they cannot pass it. The shuttle stops in front of the retirement home at the other site, which may be reflected in the bimodal speed and h_{15} distributions.

Time-to-Collision (TTC)

Figure 8 shows a side-by-side comparison of the TTC_{15} distributions for control vehicle and AV interactions. It is clear that the TTC_{15} distributions for AV interactions are either similar or shifted toward larger TTC_{15} values based on the distribution shapes and quartiles. Sites with higher control

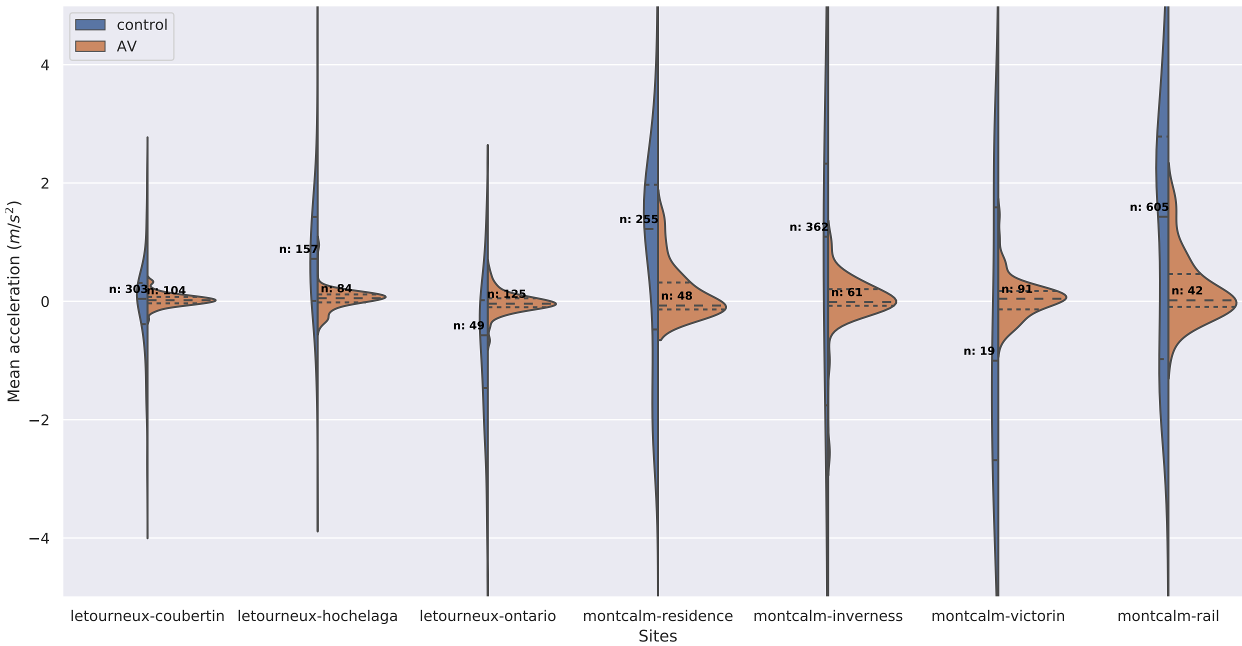

FIGURE 6 Mean user acceleration distributions of control vehicles and AV shuttles at each site.

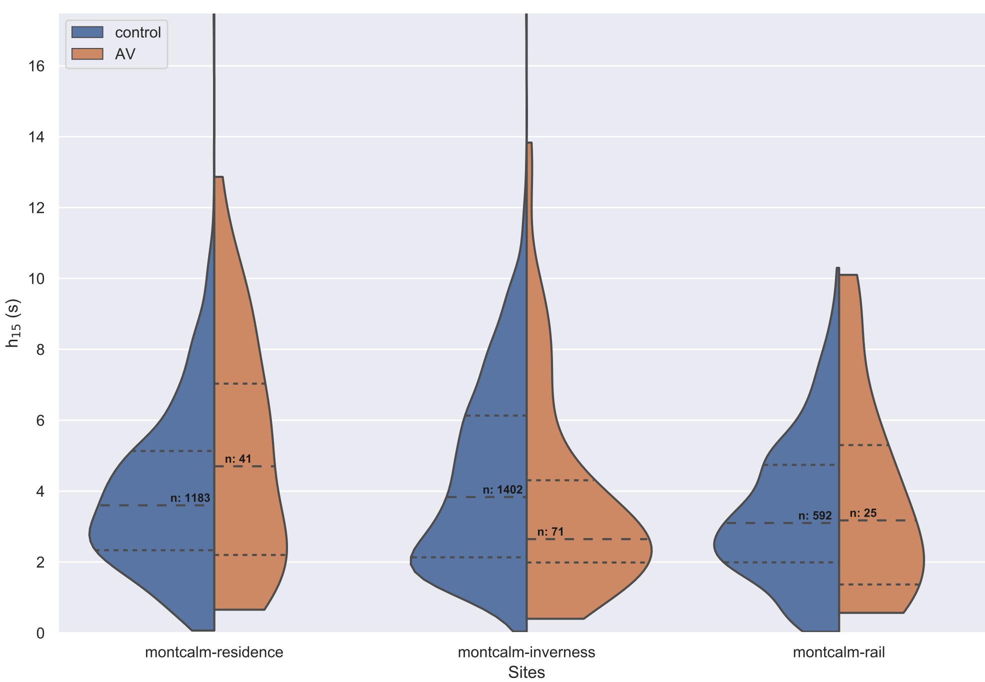

FIGURE 7 h_{15} distributions with a leading control vehicle and AV shuttle for three sites in Candiac.

vehicle and AV speeds, namely Hochelaga, in Montréal, and Inverness, in Candiac, also have lower TTC_{15} values. In Montréal, at the intersections of Letourneux Street with Pierre-De Coubertin Avenue and Ontario Street, AVs generally had to stop to yield before making a turn. Consequently, control vehicles (selected to have similar trajectories) are also likely to have similar movements, which lead to safer interactions in general. AVs appear to have even safer interactions, which may be attributed to the cautious driving behavior of these vehicles. Similarly, in Candiac, the observation site in front of the retirement home, which is close to a stop sign, and the slightly raised intersection with the railway and stop lane markings have higher TTC_{15} compared to the other sites in Candiac.

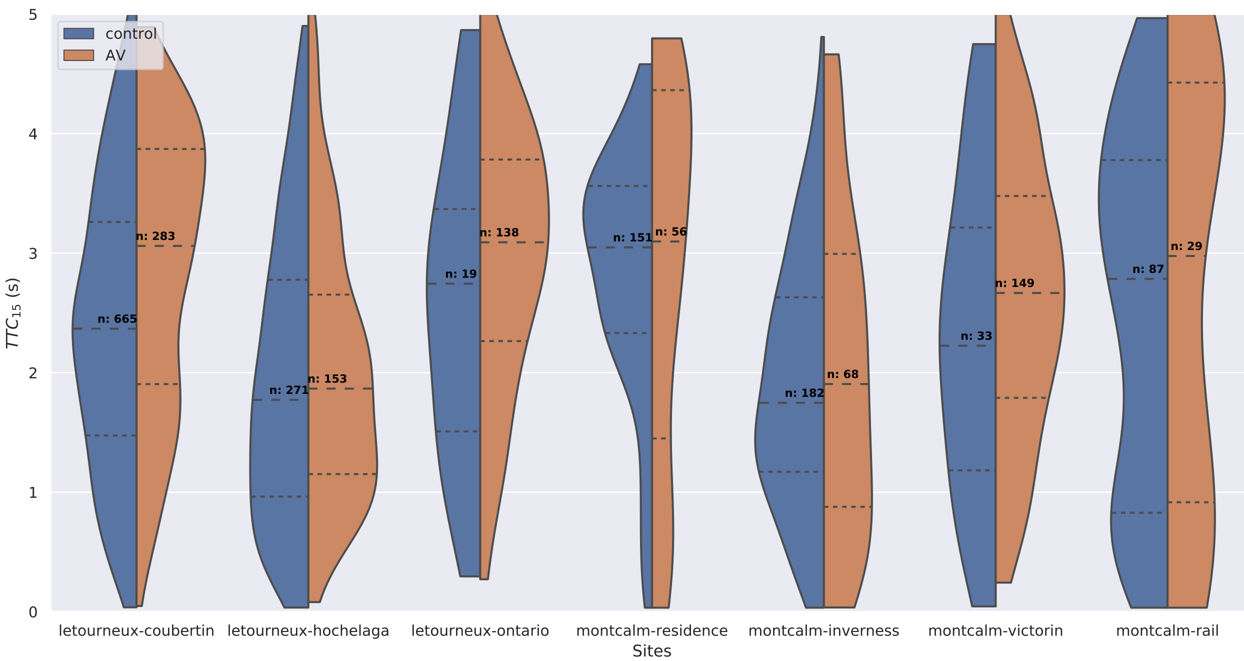

FIGURE 8 TTC_{15} distributions for control vehicle and AV interactions at each site.

Result of the KS tests comparing TTC_{15} distributions indicate a statistically significant difference (p-value < 0.05) at the following sites: Pierre-de Coubertin Avenue (D = 0.166, p-value = 3.13e-07) in Montréal, and three of the four sites in Candiac, at the retirement home (D = 0.244, p-value = 0.00303), Marie-Victorin Boulevard (D = 0.228, p-value = 0.04636) and the rail crossing (D = 0.231, p-value = 0.0322). The KS test result at the Hochelaga intersection is significant at the 0.1 level (D = 0.119, p-value = 0.0681). The distributions are similar at other sites. Few control vehicles turn at the intersection with Ontario, and the smaller interaction sample (only 19 control vehicles interactions) partly explains the lack of significance.

Interactions with vulnerable road users were investigated separately. Unfortunately, pedestrian and cyclist traffic was too low at the sites in Candiac. There were too few interactions with cyclists in Montréal to be able to draw conclusions: the TTC_{15} were higher for AV interactions with cyclists at the Pierre-de Coubertin and Hochelaga intersections, but the difference between the distributions was not significant.

Regarding the interactions with pedestrians shown in Figure 9, the AV and control vehicle TTC_{15} distributions at Pierre-De Coubertin and Ontario are similar, with most TTC_{15} above 1.5 or 2 s. There is no significant difference (p > 0.1). On the other hand, there is an important proportion of interactions with low TTC_{15} at the Hochelaga site, and the KS test is statistically significant (D = 0.3509, p-value = 0.0266), with the shuttle distribution shifted toward higher, safer TTC_{15}

values.

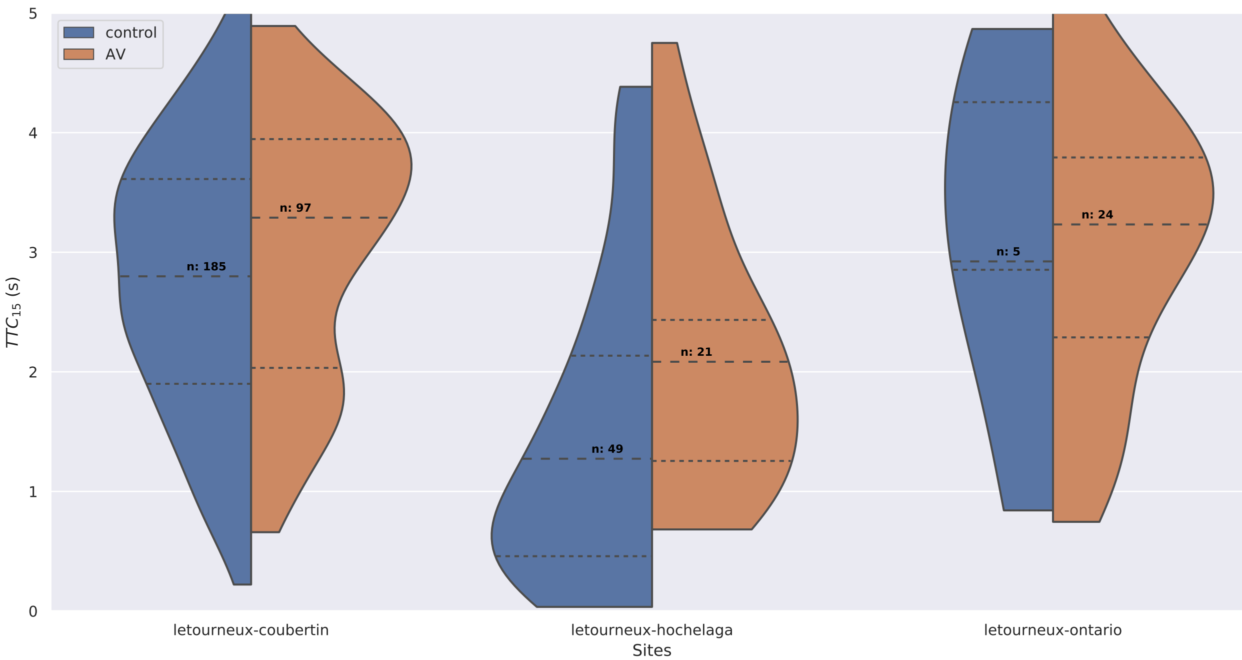

FIGURE 9 TTC_{15} **distributions for AV and control vehicle interactions with pedestrians at three sites in Montréal.**

Post-Encroachment Time (PET)

The PET distributions for AV and control vehicles interactions are shown in Figure 10. A first thing to note is that much higher PET values are recorded than TTC, as there is no upper bound on the maximum duration between the passing times of two road users at the same location (TTC is by construction limited to the motion prediction time horizon of 5 s used in this study). Similarly to the previous TTC analysis, the PETs are greater overall for shuttle interactions. Here, the KS test indicates a statistically significant difference between the distributions (p-value < 0.05) for all sites, with the exception of the Marie-Victorin intersection in Candiac (D = 0.1307, p-value = 0.5511), which, despite the distribution shift, is related to the small sample size.

Vulnerable road user safety was also studied specifically at two sites in Montréal and the PET distributions are shown in Figure 11. The KS test shows a statistically significant difference between the AV and control vehicle PET distributions at Pierre-De Coubertin site (D = 0.5876, p-value = 3.509e-06) and less clearly at the Hochelage site (D = 0.600, p-value = 0.0525), with clear shift toward higher PETs for the interactions with the shuttles. It is more difficult to draw clear conclusions for interactions with pedestrians. The distributions are statistically different at the Pierre-De Coubertin site (D = 0.2246, p-value = 0.000252), with a higher proportion of low PET values for control vehicle interactions (lower first quartile), but also a higher median and third quartile. The comparison is similar at the Hochelaga site, but the KS test is not significant because of the small sample size.

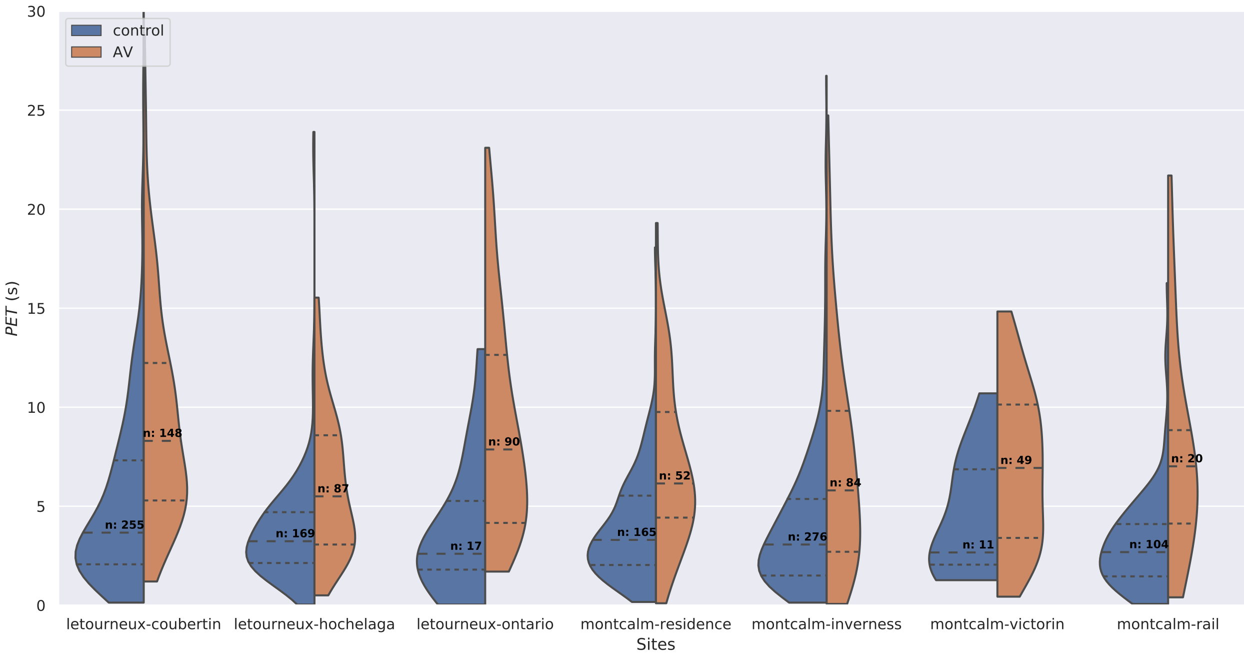

FIGURE 10 PET distributions for control vehicle and AV interactions at each site.

CONCLUSIONS

This study investigated the safety of automated shuttles with an approach that has, to the authors' knowledge, never been used. Using road user trajectories extracted from video data, the speed, acceleration, time headway, TTC and PET were computed for all road users. The results confirm the cautious driving behavior of AVs, in particular the low-speed shuttles used in these two pilots. Their average speeds and accelerations are lower, although higher at the Candiac sites where general traffic was also faster. The time headway distributions were different at each site in Candiac. The site with the highest control vehicle speeds and highest differential with AV speeds, showed the lowest time headways. This is expected in such a situation, but points to potential issues for low-speed shuttles on higher speed roads. The TTC and PET results all point toward safer interactions of AV shuttles with all road users, including vulnerable road users, compared to control vehicles. The KS tests shows significant differences and clear shifts toward safer values for several sites. This is again consistent with the lower speeds and accelerations of the automated shuttles.

A safety concern that was not addressed is the lack of cautiousness from surrounding users interacting with AVs. This was observed repeatedly on site and in the videos. This is known in the literature as compensatory behavior or risk compensation. Studies have suggested increased risk-taking behaviors might be observed from travelers who feel safer (*50*). This unique dataset will be used to further investigate the impact of automated shuttles on road user behaviour.

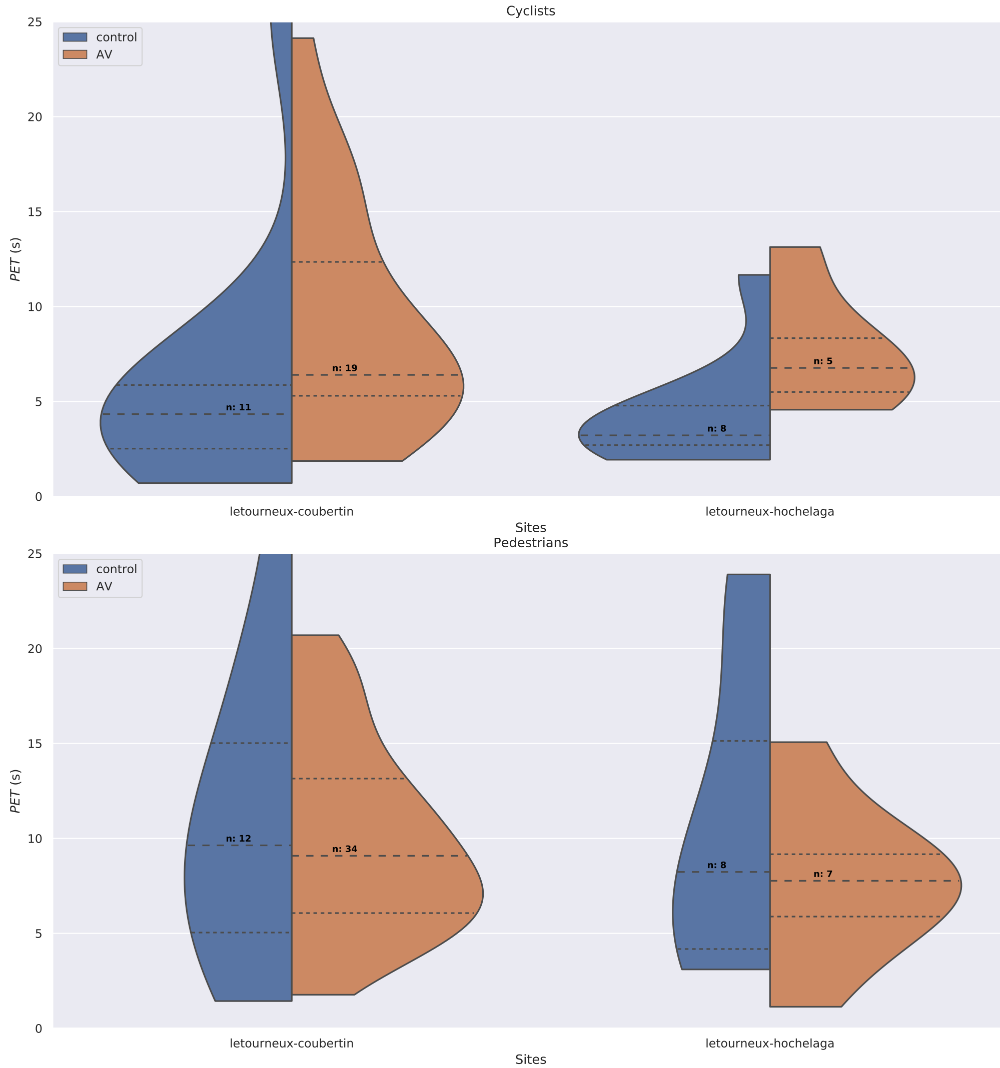

FIGURE 11 PET distributions for control vehicle and AV interactions involving cyclists (top) and pedestrians (bottom).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support of MITACS, Fondation CAA-Québec, Transport Québec and the Société de l'Assurance Automobile du Québec (SAAQ).

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: E. Beauchamp, M.-S. Cloutier and N. Saunier; analysis and interpretation of results: E. Beauchamp and N. Saunier; draft manuscript preparation: E. Beauchamp, M.-S. Cloutier and N. Saunier. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. World Health Organization, *Global status report on road safety 2018*. World Health Organization, 2018.

2. Transport Canada, *National Collision Database (motor vehicles)*, 2018.
3. Godsmark, P., B. Kirk, V. Gill, and B. Flemming, *Automated vehicles: The coming of the next disruptive technology*. The Conference Board of Canada, 2015.
4. Erke, A., Effects of electronic stability control (ESC) on accidents: A review of empirical evidence. *Accident Analysis & Prevention*, Vol. 40, No. 1, 2008, pp. 167–173.
5. Favarò, F., S. Eurich, and N. Nader, Autonomous vehicles' disengagements: Trends, triggers, and regulatory limitations. *Accident Analysis & Prevention*, Vol. 110, 2018, pp. 136–148.
6. Banerjee, S. S., S. Jha, J. Cyriac, Z. T. Kalbarczyk, and R. K. Iyer, Hands off the wheel in autonomous vehicles?: A systems perspective on over a million miles of field data. In *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, IEEE, 2018, pp. 586–597.
7. Dixit, V. V., S. Chand, and D. J. Nair, Autonomous vehicles: disengagements, accidents and reaction times. *PLoS one*, Vol. 11, No. 12, 2016, p. e0168054.
8. Laureshyn, A., C. Johnsson, T. De Ceunynck, Å. Svensson, M. de Goede, N. Saunier, P. Włodarek, R. van der Horst, and S. Daniels, *Review of current study methods for VRU safety. Appendix 6–Scoping review: surrogate measures of safety in site-based road traffic observations: Deliverable 2.1–part 4*. In-Depth understanding of accident causation for Vulnerable road users (InDeV) Project, 2016.
9. Hayward, J. C., *Near misses as a measure of safety at urban intersections*. Master's thesis, Pennsylvania State University University Park, 1971.
10. Johnsson, C., A. Laureshyn, and T. De Ceunynck, In search of surrogate safety indicators for vulnerable road users: a review of surrogate safety indicators. *Transport reviews*, Vol. 38, No. 6, 2018, pp. 765–785.
11. Jackson, S., L. F. Miranda-Moreno, P. St-Aubin, and N. Saunier, Flexible, mobile video camera system and open source video analysis software for road safety and behavioral analysis. *Transportation research record*, Vol. 2365, No. 1, 2013, pp. 90–98.
12. Polders, E. and T. Brijs (eds.) *How to analyse accident causation? A handbook with focus on vulnerable road users*. Hasselt University, 2018.
13. *Road Safety Manual*. World Road Association (PIARC), 3rd ed., 2019.
14. Saunier, N. and T. Sayed, A feature-based tracking algorithm for vehicles in intersections. In *Third Canadian Conference on Computer and Robot Vision*, IEEE, Québec, 2006.
15. Saunier, N., T. Sayed, and K. Ismail, Large-scale automated analysis of vehicle interactions and collisions. *Transportation Research Record*, Vol. 2147, No. 1, 2010, pp. 42–50.
16. St-Aubin, P., L. Miranda-Moreno, and N. Saunier, An automated surrogate safety analysis at protected highway ramps using cross-sectional and before–after video data. *Transportation Research Part C: Emerging Technologies*, Vol. 36, 2013, pp. 284–295.
17. Tarko, A., *Measuring Road Safety with Surrogate Events*. Elsevier, 2019.
18. Svensson, A., *A Method for Analyzing the Traffic Process in a Safety Perspective*. Ph.D. thesis, University of Lund, 1998, bulletin 166.
19. Songchitruksa, P. and A. P. Tarko, The extreme value theory approach to safety estimation. *Accident Analysis & Prevention*, Vol. 38, No. 4, 2006, pp. 811 – 822.
20. Emanuele Sacchi, T. S. and P. deLeur, A comparison of collision-based and conflict-based safety evaluations: The case of right-turn smart channels. *Accident Analysis & Prevention*, Vol. 59, 2013, pp. 260 – 266.

21. El-Basyouny, K. and T. Sayed, Safety performance functions using traffic conflicts. *Safety Science*, Vol. 51, No. 1, 2013, pp. 160 – 164.
22. Sacchi, E. and T. Sayed, Conflict-Based Safety Performance Functions for Predicting Traffic Collisions by Type. *Transportation Research Record*, Vol. 2583, No. 1, 2016, pp. 50–55.
23. Peesapati, L. N., M. P. Hunter, and M. O. Rodgers, Evaluation of postencroachment time as surrogate for opposing left-turn crashes. *Transportation research record*, Vol. 2386, No. 1, 2013, pp. 42–51.
24. Luettel, T., M. Himmelsbach, and H.-J. Wuensche, Autonomous ground vehicles—Concepts and a path to the future. *Proceedings of the IEEE*, Vol. 100, No. Special Centennial Issue, 2012, pp. 1831–1839.
25. Koopman, P. and M. Wagner, Autonomous vehicle safety: An interdisciplinary challenge. *IEEE Intelligent Transportation Systems Magazine*, Vol. 9, No. 1, 2017, pp. 90–96.
26. SAE international, *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Standard J3016, SAE international, 2018, retrieved from `https://www.sae.org/standards/content/j3016_201806/`.
27. Shalev-Shwartz, S., S. Shammah, and A. Shashua, On a formal model of safe and scalable autonomous vehicles. *ARXIV*, 2017.
28. Department of Motor Vehicles, *Autonomous Vehicles Testing With A Driver*. `https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/testing-autonomous-vehicles-with-a-driver/`, 2020, online; accessed 13 June 2020.
29. Department of Motor Vehicles, *Autonomous Vehicle Collision Reports*. `https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-collision-reports/`, 2020, online; accessed 13 June 2020.
30. Transports Canada, *Canada's Safety Framework for Automated and Connected Vehicles*. Transports Canada, 2019, retrieved from `https://www.tc.gc.ca/en/services/road/documents/tc_safety_framework_for_acv-s.pdf`.
31. Transport Canada, *Testing Highly Automated Vehicles in Canada*. `https://www.tc.gc.ca/en/services/road/safety-standards-vehicles-tires-child-car-seats/testing-highly-automated-vehicles-canada.html#_TOC_250004`, 2018, online; accessed 15 June 2020.
32. Gouvernement du Québec, *Highway Safety Code*. `http://legisquebec.gouv.qc.ca/en/ShowDoc/cs/C-24.2`, 2018, online; accessed 15 June 2020.
33. Tesla, *Tesla Vehicle Safety Report*. `https://www.tesla.com/fr_CA/VehicleSafetyReport?redirect=no`, 2020, online; accessed 29 June 2020.
34. Favarò, F. M., N. Nader, S. O. Eurich, M. Tripp, and N. Varadaraju, Examining accident reports involving autonomous vehicles in California. *PLoS one*, Vol. 12, No. 9, 2017, p. e0184952.
35. Papadoulis, A., M. Quddus, and M. Imprialou, Evaluating the safety impact of connected and autonomous vehicles on motorways. *Accident Analysis & Prevention*, Vol. 124, 2019, pp. 12–22.
36. Rahman, M. S., M. Abdel-Aty, J. Lee, and M. H. Rahman, Safety benefits of arterials' crash risk under connected and automated vehicles. *Transportation Research Part C: Emerging Technologies*, Vol. 100, 2019, pp. 354–371.

37. Morando, M. M., Q. Tian, L. T. Truong, and H. L. Vu, Studying the safety impact of autonomous vehicles using simulation-based surrogate safety measures. *Journal of Advanced Transportation*, Vol. 2018, 2018.
38. Sinha, A., S. Chand, K. P. Wijayaratna, N. Virdi, and V. Dixit, Crash Severity and Rate Evaluation of Conventional Vehicles in Mixed Fleets with Connected and Automated Vehicles. *Procedia Computer Science*, Vol. 170, 2020, pp. 688–695.
39. Rahman, M. S. and M. Abdel-Aty, Longitudinal safety evaluation of connected vehicles' platooning on expressways. *Accident Analysis & Prevention*, Vol. 117, 2018, pp. 381–391.
40. Rehrl, K. and C. Zankl, Digibus©: results from the first self-driving shuttle trial on a public road in Austria. *European Transport Research Review*, Vol. 10, No. 2, 2018, p. 51.
41. Straub, E. R. and K. E. Schaefer, It takes two to tango: Automated vehicles and human beings do the dance of driving–four social considerations for policy. *Transportation research part A: policy and practice*, Vol. 122, 2019, pp. 173–183.
42. Morse, P., P. St-Aubin, L. Miranda-Moreno, and N. Saunier, Transferability study of video tracking optimization for traffic data collection and analysis. In *Transportation Research Board Annual Meeting Compendium of Papers*, 2016, pp. 16–6810.
43. Bernardin, K. and R. Stiefelhagen, Evaluating multiple object tracking performance: the CLEAR MOT metrics. *EURASIP Journal on Image and Video Processing*, Vol. 2008, 2008, pp. 1–10.
44. Abramson, M., C. Audet, G. Couture, J. Dennis, Jr., S. Le Digabel, and C. Tribes, *The NOMAD project*. Software available at `https://www.gerad.ca/nomad/`, 2011.
45. Le Digabel, S., Algorithm 909: NOMAD: Nonlinear Optimization with the MADS algorithm. *ACM Transactions on Mathematical Software*, Vol. 37, No. 4, 2011, pp. 1–15.
46. Mohamed, M. G. and N. Saunier, Behavior Analysis Using a Multilevel Motion Pattern Learning Framework. *Transportation Research Record*, Vol. 2528, No. 1, 2015, pp. 116–127.
47. Niaki, M. S. N., N. Saunier, and L. F. Miranda-Moreno, Is that move safe? Case study of cyclist movements at intersections with cycling discontinuities. *Accident Analysis & Prevention*, Vol. 131, 2019, pp. 239–247.
48. Saunier, N., T. Sayed, and C. Lim, Probabilistic collision prediction for vision-based automated road safety analysis. In *2007 IEEE Intelligent Transportation Systems Conference*, IEEE, 2007, pp. 872–878.
49. St-Aubin, P., N. Saunier, and L. F. Miranda-Moreno, Comparison of Various Objectively Defined Surrogate Safety Analysis Methods. In *Transportation Research Board Annual Meeting Compendium of Papers*, Washington, D.C., 2015, 15-4629.
50. Litman, T., *Autonomous vehicle implementation predictions: Implications for transport planning*. Victoria Transport Policy Institute, 2020.