# Modelling overtaking strategy and lateral distance in car-to-cyclist overtaking on rural roads

## A driving simulator experiment

Farah, Haneen; Bianchi Piccinini, Giulio; Itoh, Makoto; Dozza, Marco

**DOI**
10.1016/j.trf.2019.04.026

**Publication date**
2019

**Document Version**
Accepted author manuscript

**Published in**
Transportation Research Part F: Traffic Psychology and Behaviour

**Citation (APA)**
Farah, H., Bianchi Piccinini, G., Itoh, M., & Dozza, M. (2019). Modelling overtaking strategy and lateral distance in car-to-cyclist overtaking on rural roads: A driving simulator experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *63*, 226-239. https://doi.org/10.1016/j.trf.2019.04.026

# Modelling Overtaking Strategy and Lateral Distance in Car-to-Cyclist Overtaking on Rural Roads: A Driving Simulator Experiment

**Haneen Farah[1,a], Giulio Bianchi Piccinini[b], Makoto Itoh[c], Marco Dozza[d]**

(a) Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands, Tel: +31 15 278 3401, Fax: +31 (0)15 278 7956, E-mail: h.farah@tudelft.nl

(b) Division of Vehicle Safety, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Lindholmspiren 3, 417 56 Göteborg, Sweden, Tel: +46 31 772 1421, E-mail: giulio.piccinini@chalmers.se

(c) Faculty of Engineering, Information and Systems, University of Tsukuba, Tennoudai, Tsukuba-shi, Ibaraki-ken, 305-8573, Japan, Tel: +81 29 853 5502, Fax: +81 29 853 5809, E-mail: itoh@risk.tsukuba.ac.jp

(d) Division of Vehicle Safety, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Lindholmspiren 3, 417 56 Göteborg, Sweden, Tel: +46 31 772 3621, E-mail: marco.dozza@chalmers.se



[1] Corresponding author

**Abstract**

The involvement of cyclists in road crashes has not been decreasing with the same magnitude as the involvement of other road users. In particular, the interactions between cyclists and motorized traffic can lead to high-severity crashes. To improve the safety of these interactions, a thorough understanding of road user behaviour is first needed. In this study, we focused on drivers overtaking cyclists on rural roads. The two main objectives of this study were to develop models that predicted: (a) drivers' decisions to perform either a flying or an accelerative overtaking manoeuvre in the presence of oncoming traffic, and (b) the lateral comfort distance that drivers maintain from cyclists during the overtaking.

A driving simulator study was designed to assess driver decision-making during the overtaking. The 37 drivers who participated in the study each performed seven overtaking manoeuvres with oncoming traffic. Out of the 259 overtaking manoeuvres, 168 were flying and 91 were accelerative. Binary logistic-regression models with mixed effects predicted the type of overtaking strategy (flying or accelerative). Driving speeds were found to significantly affect the strategy. The overall performance of the models predicting the strategy was 85-90%. Models were also developed for predicting the lateral comfort distance. The results show that the lateral comfort distance is mostly affected by the longitudinal distance between the subject vehicle and the oncoming vehicle, the longitudinal distance between the subject vehicle and the cyclist, and the presence of an oncoming vehicle—as well as by the drivers' characteristics (sensation seeking in flying overtaking manoeuvres and ordinary violations in accelerative manoeuvres). The root mean square error, which was used to assess the performance of the models, ranged from 0.56-0.62.

In conclusion, the models predicting the overtaking strategy performed reasonably well, while the models predicting lateral distance did not provide accurate predictions. The models predicting overtaking strategy may support 1) the development and evaluation of active safety systems, 2) the design of automated driving, and 3) policy making.

**Keywords:** Overtaking, Cyclists, Driver Behaviour, Driving Simulator, Active Safety Systems, Automated Driving

## 1. Introduction

Cycling is a sustainable and affordable transport mode which has major health, environmental, and economic benefits (Fishman, Schepers, & Kamphuis, 2015; Wegman, Zhang, & Dijkstra, 2012). On the other hand, cyclists are vulnerable road users, and their involvement in road crashes has not been decreasing with the same magnitude as other road users (Santacreu, 2018). In the Netherlands, cyclist fatalities constitute about 30% of the total number of road deaths (SWOV, 2017); in Sweden, cyclists are the most frequently injured road users (Trafikverket, 2014); in the US, 840 cyclists were killed in motor vehicle crashes in 2016, accounting for 2.2% of all traffic deaths that year (NHTSA, 2018); and in New Zealand, cyclists are ten times more likely to be involved in a serious or fatal crash per kilometre travelled than car drivers (Balanovic et al., 2016). These statistics raise tremendous concerns, especially as cycling's popularity has been increasing in many western countries in the last few years (Pucher & Buehler, 2017). The most serious type of crash for cyclists involves a collision with a vehicle (Wegman et al., 2012), including the scenario of a vehicle approaching a cyclist from behind and overtaking him/her (Feng, Bao, Hampshire, & Delp, 2018; Kay, Savolainen, Gates, & Datta, 2014). This type of interaction occurs mostly on rural roads, where the injuries can be more severe (Boufous, de Rome, Senserrick, & Ivers, 2012). Despite the higher severity of vehicle-cyclist crashes on rural roads, there have been very few studies, compared to those on urban roads (Dozza & Schwab, 2017; Llorca, Angel-Domenech, Agustin-Gomez, & Garcia, 2017). To mitigate or even prevent this type of crash via infrastructural, educational, or technological solutions, a thorough understanding of these interactions and of drivers' overtaking decisions processes is first needed.

In the literature, some studies have investigated drivers' cyclist-overtaking behaviour, with the focus on the lateral distance maintained from the cyclist during overtaking; however, none of these studies, to the best of our knowledge, has developed mathematical models for predicting either drivers' overtaking strategies or the lateral comfort distances. The word "comfort" is included in the term because we argue that the lateral distance that a driver keeps from a cyclist during the overtaking process is triggered by the comfort feeling of being not too close to the cyclist. The comfort distance was explained by Summala (Summala, 2007). The following paragraphs summarize the state of the art with respect to the behaviour of drivers' as they overtake cyclists.

One of the earliest studies to investigate the behaviours of drivers when overtaking cyclists was conducted by Walker (2007). The author collected data on the proximity of drivers to the cyclist during overtaking manoeuvres on various highways in the United Kingdom. Walker found that drivers increased their lateral distances during overtaking when the cyclist appeared from behind to be a long-haired female, and provided less space when he was wearing a helmet. In a follow-up study, Walker, Garrard, and Jowitt (2014) found that drivers do not adjust their lateral clearance when the cyclist's appearance is varied to indicate different types of cyclists with different skill and/or experience levels.

Shackel and Parkin (2014) instrumented a bike with an ultrasonic distance detector and forward- and side-facing cameras. It recorded the proximity and speed of motor traffic passing as it was ridden at one meter from the kerb. The results revealed that on narrower lanes with lower speed limits and no centre-line markings, the overtaking speeds were lower. Drivers also passed more slowly if they were driving a long vehicle or in a platoon, or when vehicles approaching from the opposite direction arrived at the passing point within five seconds.

Llorca et al. (2017) instrumented two bicycles with laser rangefinders, a GPS tracker, and three video cameras and rode them along seven rural road segments in Spain. The researchers collected data on the lateral clearance between the overtaking motor vehicle and the bicycle, the motor vehicle speed and type, its left lane occupation (i.e. ), the interaction with opposing traffic, and the cyclists' perceived safety. The analysis of the data revealed that the combined factors of lateral clearance, vehicle type, and vehicle speed were more significantly correlated with the cyclists' perceived risk than lateral clearance alone. In a finding similar to Walker (2007), heavy vehicles kept lower clearances to the cyclists than passenger cars. The authors also concluded that the current lateral distance standards (1.5 metres at 50 km/h in Spain) are not sufficient to guarantee safe overtaking manoeuvres, as they do not account for other factors such as the overtaking speed or the presence of heavy vehicles.

Bella and Silvestri (2017) also conducted a driving simulator experiment to investigate the influence of the road's cross-section configurations and its geometric elements on drivers' interactions with cyclists. They found that wider bicycle lanes ensured higher lateral clearance between the driver and the cyclist; on straight road sections, drivers maintained the same driving speeds whether there was a cyclist or not, and the lateral clearance was smaller than on curved road sections. Additionally, on left curves drivers tended to cut the curve when overtaking, exposing themselves to a higher risk of a collision with oncoming vehicles.

Bianchi-Piccinini, Moretto, Zhou, and Itoh (2018) conducted a driving simulator study in Japan in which 36 Japanese drivers (21 males and 15 females) performed seven cyclist-overtaking manoeuvres on a rural road, with oncoming vehicles approaching at different nominal times to collision (TTC). The authors found a significant correlation between the overtaking strategy and the nominal TTC: as the TTC decreased, more drivers used the accelerative strategy, because they slowed down and waited for the oncoming vehicle to pass before accelerating to overtake the cyclist. The study also found that the minimum lateral safety margins were larger in accelerative manoeuvres than in flying manoeuvres, which are conducted without waiting for any oncoming traffic to clear the oncoming lane in the overtaking zone. Furthermore, during flying manoeuvres the drivers were also closer to the cyclist during the steering away and passing phases (see Dozza, Schindler, Bianchi-Piccinini, and Karlsson (2016) for precise definitions of the overtaking phases). These reduced safety margins during flying manoeuvres may be due to the possible risk of a collision with the oncoming traffic.

Evans, Pansch, and Singer-Berk (2018) instrumented a bike with 3FT radar and a GoPro video camera. They measured the passing vehicles' lateral distances and encroachments on the opposite lane on urban and suburban roads with different bicycling facilities. The results of the study show that the overall encroachment rate was low. The presence of a vehicle in the adjacent lane travelling in the same direction played the biggest role (except for facility type) in reducing the distance between a cyclist and a vehicle during a pass.

Feng et al. (2018) investigated which factors affect drivers' cyclist-overtaking manoeuvres using naturalistic driving data collected in Ann Arbor, Michigan. The results show that a substantial amount of overtaking involved drivers' crossing the solid centreline, although drivers rarely completely crossed into the other lane (with all the four wheels) even when there was no oncoming traffic. Dozza et al. (2016), who instrumented an electric bicycle with a LIDAR and two cameras, recorded 145 overtaking manoeuvres performed by car and truck drivers on public rural roads in Sweden. The authors identified four overtaking phases and quantified corresponding driver comfort zones accordingly. Unlike the study by Feng et al. (2018), who used a single measure to quantify lateral distance when overtaking, the comfort zone is a continuous measure that can be calculated as the minimum lateral distance between the cyclist and the vehicle at each time point during the different phases of an overtaking. Dozza et al. found that oncoming traffic had the most impact on the comfort zone, while neither vehicle speed, lane width, shoulder width, nor posted speed limit significantly affected the driver's comfort zone. When a vehicle is approaching, drivers drive significantly closer to the cyclist not only when passing, but also when approaching and steering away from the cyclist. Bianchi-Piccinini et al. (2018) and Kovaceva, Nero, Bärgman, and Dozza (2018) found similar results—from research with a driving simulator and the UDRIVE naturalistic dataset, respectively.

Finally, Abe, Sato, and Itoh (2018) were interested in understanding the factors that affect driver's trust in the behaviour of automated vehicles when passing a scooter or a bicycle. For that purpose, they first conducted a driving simulator experiment to investigate the overtaking behaviour of human drivers. The obtained data were used to parametrize the design of an automated vehicle in the driving simulator, and then varied to study the effects of different automated driving settings on drivers' trust. The results show that drivers trusted the system more when it applied driving speeds similar to the ones applied by human drivers, but maintained greater lateral distances and started the passing manoeuvres earlier than did human drivers.

## 2. Research Gaps and Research Questions

From the review of the state of the art, it is clear that there are still knowledge gaps to be filled in order to fully understand drivers' decisions to perform flying or accelerative overtaking manoeuvres, especially when there is oncoming traffic. Understanding the factors that affect this decision is important, as the two strategies differ significantly. Accelerative overtaking manoeuvres are safer than flying overtaking manoeuvres: the drivers often drive at lower speeds, have better control of the interaction with the oncoming vehicle, and leave larger clearances to the cyclist in all overtaking phases (Dozza et al., 2016). A major drawback to studying overtaking strategies is the lack of mathematical models that can predict drivers' overtaking strategies and lateral distances when overtaking cyclists.

The research questions investigated in this study are the following:

- Which factors significantly affect drivers' decisions regarding the overtaking strategy and the lateral comfort distance from the cyclist?
- How early (in terms of distance from the cyclist) and accurately can we predict a driver's overtaking strategy?
- Do the different factors impact the driver's lateral comfort distance from the cyclist similarly (direction and magnitude) in flying versus accelerative overtaking manoeuvres?

To answer these research questions quantitatively, the following two main objectives were defined: 1) to develop a model that can predict a driver's decision to perform a flying or accelerative overtaking manoeuvre when approaching a cyclist in the presence of oncoming traffic; and 2) to develop a model that can predict the lateral comfort distances that drivers maintain when approaching and overtaking a cyclist. The lateral comfort distance is defined in this study as the orthogonal component of the distance between the vehicle and the cyclist. The first objective is important for the design of active safety systems such as forward collision warning (FCW) and automated emergency braking (AEB). These systems typically act on a threshold of time-to-collision to the forward obstacle (Brannstrom, Coelingh, & Sjoberg, 2010) and are evaluated by Euro NCAP when overtaking a cyclist (EuroNCAP, 2017). They may prevent a rear-end crash with a cyclist when the driver has not decided which overtaking manoeuvre (flying or accelerative) to perform in a timely manner (an indication that the driver may not have seen the cyclist). The AEB and FCW systems may also support a driver who misjudges the kinematics of an oncoming vehicle and opts for a flying overtaking strategy, potentially leading to a head-on collision. In both cases, the systems need to understand the driver's intention to overtake in order to act appropriately.

The second objective, developing a predictive model for the lateral comfort distance, is important in order to understand the factors (subjective characteristics as well as surrounding environmental characteristics such as the presence of oncoming vehicle) can influence drivers' decision making. A better comprehension of those characteristics could provide valuable inputs for fine-tuning the behavioural models of automated vehicles, so they maintain safe driving behaviour when approaching and overtaking cyclists—while considering individual human differences and preferences. This consideration is important from the point of view of both the driver and the cyclist. For the overtaking driver there is a trade-off between the lateral comfort distances from the cyclist and from the oncoming traffic, whereas the cyclist makes no such trade-off, and is likely to experience a shorter lateral distance to the overtaking vehicle as riskier. It might be expected that drivers' trust and acceptance of automated vehicles would increase if automated vehicles are able to trade distances as humans do.

## 3. Research Method

### 3.1. Experiment and data collection

A driving simulator experiment was conducted at the University of Tsukuba in Japan. The fixed-base driving simulator was equipped with a steering wheel, an accelerator pedal, a brake pedal, and a gearshift. The driving scene was shown on five screens as presented in Figure 1.

The study involved 42 participants, but only the data for 37 participants could be used for the analyses. The other five were excluded due to simulation sickness, non-compliance with the experimental protocol, or missing data. The 37 participants (22 male and 15 female Japanese drivers) were 48.0 ± 19.1 years of age and had owned a driving licence for 23.5 ± 15.1 years. All the participants signed a consent form to confirm their participation in the study and permit the use of their collected data. In addition, they filled in a demographic questionnaire (the Driver Behaviour Questionnaire: DBQ) and the Arnett Inventory of Sensation Seeking (AISS) after the introductory session and after signing the consent form. We used the 28-item version of the DBQ which considered four behavioural factors: aggressive violations, ordinary violations, errors, and lapses (Lawton, Parker, Manstead, & Stradling, 1997). Two items (4 and 22) were removed because they dealt with roundabouts, which were still uncommon intersections in Japan at the time of the study. Therefore, the DBQ questionnaire administered to the participants contained just 26 items. The AISS includes 20 items, with ten items for each of the two subscales, intensity and novelty (Arnett, 1994). The AISS was chosen instead of the Sensation Seeking Scale Form V (Zuckerman, Eysenck, & Eysenck, 1978) because the latter includes numerous items on alcohol and drug use and sexual behaviour—which might have made the participants uncomfortable.

During the experiment, the participants first underwent a test trial followed by two experimental trials. The purpose of the test trial, *Trial route*, was to allow the participants to get comfortable in the driving simulator and get familiar with the scenery and the task (overtaking a cyclist). The experimental trials, *Route 1* and *Route 2*, were designed to evaluate the drivers' comfort zone boundaries and overtaking strategies without and with oncoming traffic, respectively (for more details about the experimental setup, please see Bianchi, Piccinini et al., 2018).

**Fig. 1.** The driving simulator used in the study.

The trials were conducted on a two-lane rural road (one for each direction of travel) with no divider; Lanes were 3.2 m wide and the shoulders were 0.4 m wide. The data collected during *Route 2* were used for the models described in this paper. The road was about 16 km long and the trial included seven overtaking manoeuvres with oncoming traffic. The manoeuvres occurred on straight stretches of road with good visibility, in the presence of a dashed centre line. The order of the overtaking manoeuvres (which varied in terms of nominal TTC) was not randomised among the participants, thus ensuring that different participants experienced the same environmental conditions during each manoeuvre (e.g., the length of the straight stretch of road). The lack of randomization is not expected to be an issue, since the participants had already driven along the Trial route and Route 1; therefore, only marginal changes in participants' behaviours should have occurred. During the trial, the participants were requested to overtake cyclists as they would in real life and to keep the speed of the vehicle as close as possible to the speed limit of 70 km/h.

The cyclist to be overtaken by the participant was standing still until the subject's vehicle was 100 m away, at which point the cyclist started to move at a constant speed of 22 km/h, maintaining a constant distance of 0.3 m from the kerb of the road. The oncoming vehicle was standing still until the distance between the subject and oncoming vehicles reached a given distance (nominal TTC) which varied in different manoeuvres (Table 1). When the subject vehicle was 50 m away from the bicycle, the oncoming vehicle's speed varied to ensure the specific value for TTC in spite of changes in the subject vehicle speed.

**Table 1**
Distance and nominal Time-To-Collision (TTC) between subject and oncoming vehicles in the different overtaking manoeuvres

| **Overtaking number** | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
|---|---|---|---|---|---|---|---|
| **Distance between oncoming and subject vehicle [m]** | 500 | 350 | 480 | 450 | 520 | 400 | 380 |
| **Nominal TTC [s]** | 9.0 | 6.0 | 8.5 | 8.0 | 9.5 | 7.0 | 6.5 |

N.B.: the overtaking manoeuvres are reported in chronological order. During all manoeuvres, the cyclists were always in the same lane and travelling in the same direction as the subject vehicle.

In Japan, there is no official quantitative threshold for the minimal lateral safety margin to cyclists (JAF, 2017). However, some local governments, such as the Ehime Prefectural Government, recommend having at least 1.5 m lateral distance from the cyclist. With respect to the centre line in Japan, if the centre line is dashed and white, the vehicle is allowed to cross it.

3.2. Analysis technique

The analysis technique comprised three steps. First, descriptive analysis and statistic tests were conducted *(3.2.1),* followed by the development of the predictive models for flying or accelerative overtaking manoeuvres *(3.2.2)* and the development of predictive models for the lateral comfort distance *(3.2.3)*.

*3.2.1. Descriptive Analysis & Statistics*

First, drivers' speed profiles and their longitudinal distances from the cyclists for each overtaking manoeuvre were plotted against the cumulative distance driven along the road. Each manoeuvre was categorized as either flying or accelerative based on the minimum speed of the subject vehicle, from 100 m away from the cyclist until the driver reached the cyclist. If the minimum speed was less than 10 m/s, the overtaking was categorized as accelerative, otherwise it was categorized as flying. The threshold of 10 m/s was chosen after analysing the speed profiles of the subject vehicles when approaching and overtaking the cyclist during the manoeuvres (Figure 2): the whole set of overtaking manoeuvres could be divided in two clusters using a speed of 10 m/s.

*3.2.2. Predictive Continuous Model for Flying or Accelerative Overtaking Manoeuvres*

Each participant performed seven overtaking manoeuvres in the driving simulator; for each one there were multiple observations—such as clustered longitudinal data, as reported by (West, Welch, & Galecki, 2014). Because all observations from the same participant are by definition correlated, a Linear Mixed Model (McCulloch & Neuhaus, 2001) with both fixed and random effects was applied. Random effects allow the residuals associated with the longitudinal measures on the same unit of analysis to be correlated, thus taking into account the clustering effect. Linear mixed-effect models have been used in previous studies to model driver overtaking behaviour (Farah, 2013), driver speed behaviour on curves

(Farah, Daamen, & Hoogendoorn, 2018), and the modelling choices of control transitions in automated driving (Varotto, Farah, Toledo, van Arem, & Hoogendoorn, 2017). The models developed in this study were estimated using the 'lme4' package (Bates et al., 2014) and validated using the 'rpart' package (Therneau, Atkinson, Ripley, & Ripley, 2018) in the R statistical program (Team, 2013).

The first step in building the model was to define the earliest time point when we could predict whether a driver was going to perform a flying or accelerative manoeuvre. This decision point was defined as the time when 100 metres separated the cyclist from the subject vehicle approaching from behind. This distance is close to the limit of the typical detection range for many commercial radars used for active safety. At this distance, an active safety system may start predicting the type of overtaking manoeuvre which will be performed, and use this prediction to inform the decision-making and threat-assessment algorithms for FCW and AEB. Based on the defined decision point of 100 m, three models were estimated. The first model *(A1)* was based on observations when the subject vehicle was 80-100 m from the cyclist, the second *(A2)* when the subject vehicle was 50-70 m away, and the third *(A3)* when the subject vehicle was 20-40 m away from the cyclist.

Since the response variable can only take two possible values, flying (1) or accelerative (0), it is binary; a binary logistic regression model with mixed effects was found to be suitable in this case. A general specification of the model is presented in Eqs. (1) to (3):

$$y_{ni} = \begin{cases} 1 & if \quad \beta_o + \beta_j \cdot X_{ji} + b_{0n} + b_{0o} \geq 0 \\ 0 & else \end{cases} \tag{1}$$

$$\log\left(\frac{P(y_{ni}=1|X_{ji})}{1-P(yn_i=1|X_{ji})}\right) = \beta_o + \beta_j \cdot X_{ji} + b_{0n} + +b_{0o} \tag{2}$$

$$P(y_{ni} = 1 \middle| X_{ji}) = \frac{\exp(\beta_o+\beta_j \cdot X_{ji}+b_{0n}++b_{0o})}{1+\exp(\beta_o+\beta_j \cdot X_{ji}+b_{0n}++b_{0o})} \tag{3}$$

where: $y_{ni}$ is the response variable which takes a value of 1 for a flying manoeuvre or 0 for an accelerative one; $n$ and $i$ are the indices for the driver and the overtaking number, respectively; $\beta_0$ is the mean intercept; $\beta_j$ is the row vector of fixed-effect parameters corresponding to the column vector of the explanatory variables $X_j$; $b_{0n}$ and $b_{0o}$ are random-effect parameters for the intercepts, which are assumed to follow normal distributions with mean 0 and standard deviations of $\sigma_{b0n}$ and $\sigma_{b0o}$, respectively.

The three models *(A1, A2, A3)* were tested by examining different explanatory variables related to the relative distances and speeds of the cyclist and subject and oncoming vehicles, as well as drivers' characteristics. Initially, variables (such as the relative distances and speeds) were identified from the literature as the ones most likely to significantly affect the overtaking decision; further analysis was performed to determine whether additional explanatory variables improved the performance of the models.

### 3.2.3. *Predictive Continuous Model for the Lateral Comfort Distance*

The main purpose of this model is to predict the lateral comfort distances after the driver decides to overtake the cyclist and performs the overtaking manoeuvre. The model can be used by vehicle manufacturers to fine-tune the algorithms that govern the behaviour of automated vehicles, so that the vehicles adopt safe driving behaviour when approaching and overtaking cyclists while, at the same time, considering differences between humans as well as their preferences. Since the dataset comprises multiple overtaking manoeuvres belonging to the same participant, and multiple observations belonging to the same overtaking manoeuvre, the developed models should account for these correlations. This is clustered longitudinal data, meaning that the lateral comfort distance is measured continuously for each overtaking manoeuvre, with the different overtaking manoeuvres clustered for each driver. Therefore, the random effects in the model are associated with both the clusters (i.e. drivers), and the units of analysis within these clusters (i.e. overtaking). The formulation of this mixed model is presented in Eq. (4):

$$LCD_{not} = \beta_0 + \beta_j \cdot X_{jt} + \mu_{0n} + \mu_{0o} + \varepsilon_t \tag{4}$$

where: $LCD_{not}$ is the lateral comfort distance for driver $n$, overtaking number $o$, and observation $t$ (dependent variable); $\beta_0$ is the average lateral comfort distance for the population; $\beta_j$ is the row vector of fixed-effect parameters corresponding to column vector of the explanatory variables $X_{jt}$; $X_{jt}$ is the column vector of explanatory variables of observation $t$; $\mu_{0n}$ is the driver-specific residual (effect of clustering observations at the driver level), $\mu_{0n} \sim N(0, \sigma_{0n})$; $\mu_{0o}$ is the overtaking-specific residual (effect of clustering overtaking-level observations), $\mu_{0o} \sim N(0, \sigma_{0o})$; and $\varepsilon_t$ is the observation-specific error term, $\varepsilon_t \sim N(0, R_i)$, where $R_i$ is the covariance matrix. Different covariance structures for the residuals were examined and the estimation results were compared. The models were estimated using the R statistical program using the 'nlme' package (Pinheiro et al., 2017), and the 'lme' function, which allowed us to define the variance-covariance structure of the residuals. The Restricted Maximum Likelihood (REML) estimation method (which is often preferred to ML estimation) was applied, because it produces unbiased estimates of covariance parameters by taking into account the loss of degrees of freedom that results from estimating the fixed effects in β (West et al., 2014). Different models' specifications were tested by examining different explanatory variables, adopting a strategy similar to the one used for predicting the overtaking manoeuvre type (*Section 3.2.2*).

## 4. Results

In this section, we first present descriptive statistics *(Section 4.1)* of the dataset, followed by the results of the predictive models for flying or accelerative overtaking manoeuvres *(Section 4.2)* and the predictive models for the lateral comfort distance *(Section 4.3)*.

### 4.1. Descriptive statistics

The dataset resulted in 259 overtaking manoeuvres, 168 categorized as flying and 91 as accelerative. Figure 2 presents drivers' speed profiles and their longitudinal distances from the cyclists for each overtaking manoeuvre against the cumulative distance along the road. The TTC for each overtaking manoeuvre is plotted at the top of each sub-figure.

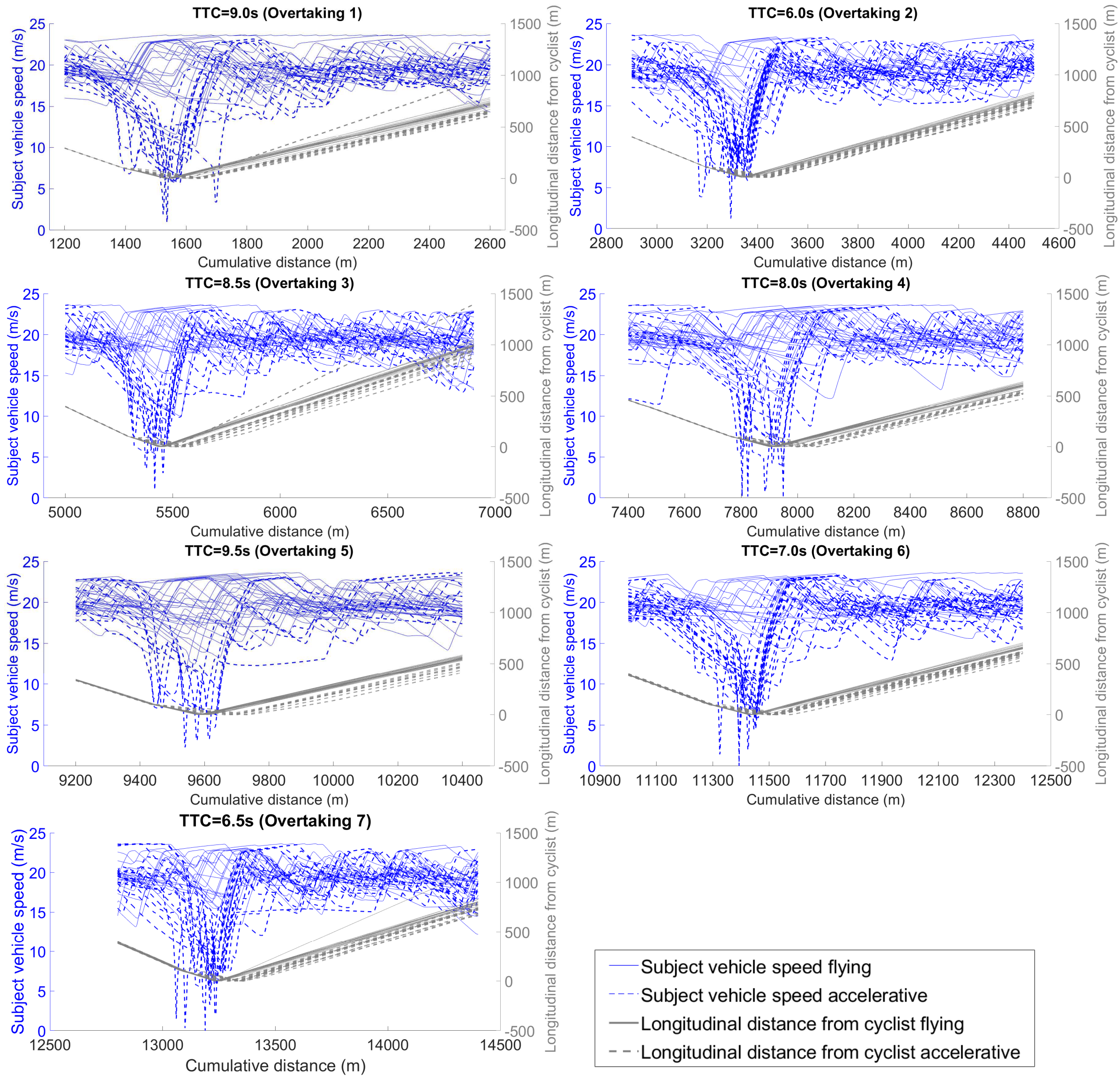


**Fig. 2.** Speed profiles and the longitudinal distance from the cyclist as a function of the cumulative distance along the road for each overtaking manoeuvre.

The speed profiles clearly distinguish two patterns, one including overtaking manoeuvres with a speed drop of ~15 m/s (blue dashed lines), and the other one including overtaking manoeuvres with a smaller speed drop of ~5 m/s (blue solid lines). By definition (Dozza et al., 2015), accelerative manoeuvres imply a significant reduction of speed before the overtaking: for this reason, the overtaking manoeuvres in the first identified pattern (with a speed drop of ~15 m/s) were categorized as accelerative and the overtaking manoeuvres in the second identified pattern (with a speed drop of ~5 m/s) were categorized as flying. Figure 2 shows these two clear repetitive patterns: 1) the speed profiles of the flying and accelerative manoeuvres, and 2) the longitudinal distance between the cyclist and the vehicle plotted against the cumulative distance. Drivers performing flying manoeuvres approached and passed the cyclist faster because their relative speeds were higher than those performing accelerative manoeuvres. (In Figure 2, the reader may note that the grey solid lines reach a distance of zero earlier than the grey dashed lines).

Figure 3 presents drivers' lateral comfort distances and their longitudinal distances from the cyclists for each overtaking manoeuvre against the cumulative distance along the road. As in Figure 2, the TTC for each overtaking manoeuvre is plotted at the top of each sub-figure.

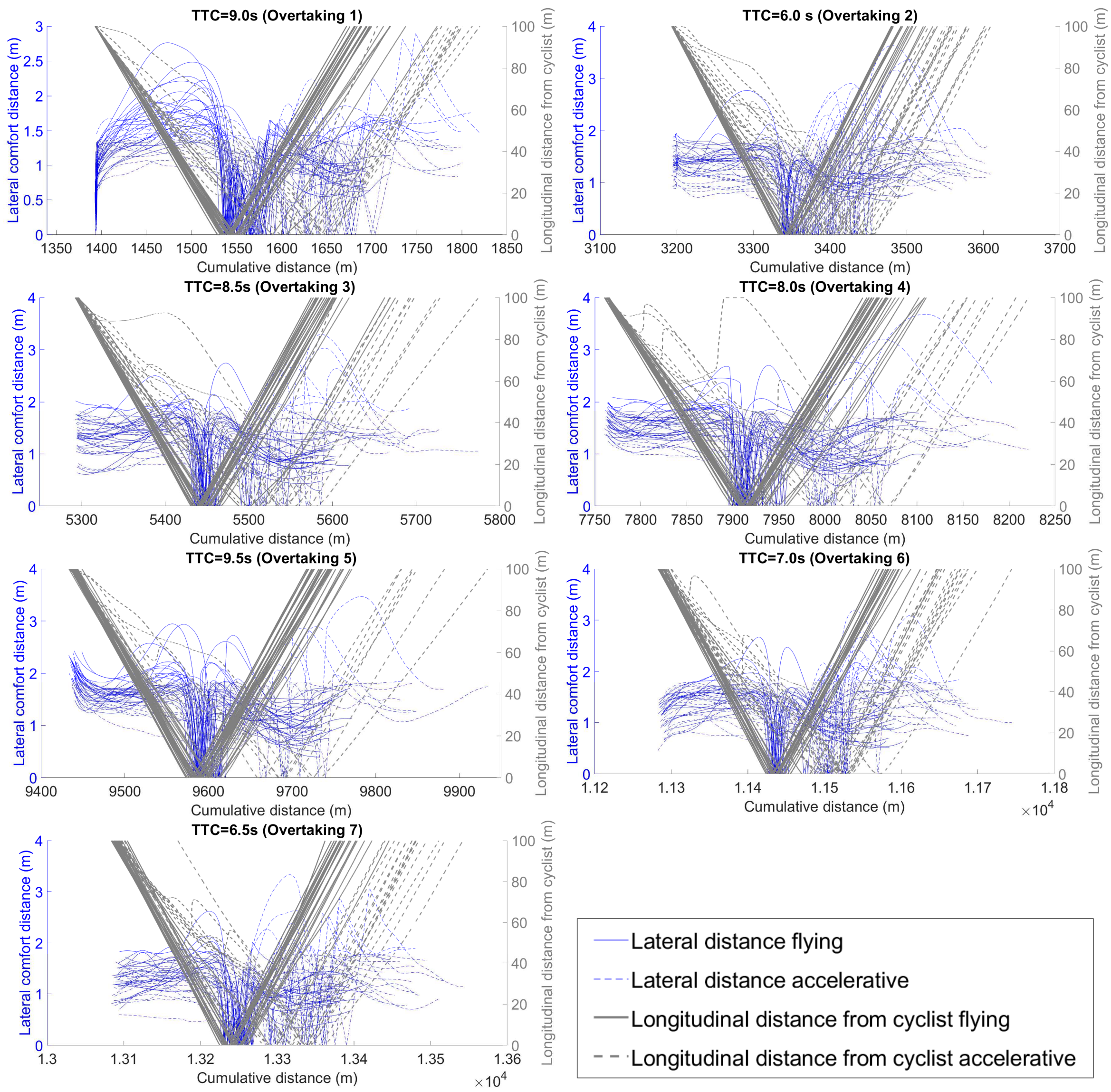


**Fig. 3.** The lateral distance and the longitudinal distance from the cyclist as a function of the cumulative distance along the road for each overtaking manoeuvre.

Figure 3 shows a clear distinctive pattern of the lateral comfort distance that drivers maintain during flying (blue continuous lines) and accelerative (blue dashed lines) overtaking manoeuvres. Drivers kept shorter lateral comfort distances once they steered away from the cyclists when they performed flying manoeuvres. Further, drivers passed the cyclists earlier (i.e. drivers initiated the overtaking when farther from the cyclist) in flying manoeuvres because driving speeds were higher.

Table 2 presents descriptive statistics of the different measured or calculated variables per overtaking type (flying vs. accelerative), taking into account all the observations from the time when the longitudinal distance between the subject vehicle and the cyclist was 100 metres until the subject vehicle reached the cyclist.

As can be seen from the results in Table 2, the mean longitudinal distance between the subject vehicle and the oncoming vehicle is, as expected, higher for flying overtaking manoeuvres. This difference is due to the fact that, in accelerative overtaking manoeuvres, the drivers in most cases waited for the oncoming vehicle to pass before overtaking, leading to what we assigned as negative values of the distances between the subject and oncoming vehicles. In contrast, in flying manoeuvres, the driver overtook the cyclist before the oncoming car arrived, so the distance between the two vehicles was positive for the entire manoeuvre. In addition, the mean lateral distance from the cyclist was higher for flying overtaking manoeuvres because drivers who performed a flying manoeuvre steered away from the cyclist earlier. Drivers who performed flying manoeuvres drove on average faster than drivers who performed accelerative manoeuvres (as seen in Figure 2).

**Table 2**

Descriptive statistics of driver behaviour characteristics in flying versus accelerative overtaking manoeuvres (in the last 100 metres of the subject vehicle's approach to the cyclist).

| **Factor** | | **Flying Overtaking** | | **Accelerative Overtaking** | |
|---|---|---|---|---|---|
| | | Mean | Std. | Mean | Std. |
| Longitudinal distance between Subject and Oncoming Vehicles (m)* | LongDisSO | 217.77 | 153.92 | 193.53 | 162.17 |
| Lateral distance between Subject and Oncoming vehicles (m) | LatDisSO | -0.31 | 0.86 | -1.15 | 1.93 |
| Longitudinal distance between Subject vehicle and Cyclist (m) | LongDisSC | 49.43 | 29.24 | 50.86 | 26.66 |
| Lateral distance between Subject vehicle and Cyclist (m) | LatDisSC | 1.49 | 0.38 | 1.20 | 0.42 |
| Relative speed between Subject vehicle and Oncoming vehicle (m/s) | RelSpeedSO | 28.67 | 11.19 | 26.12 | 11.46 |
| Relative speed between Subject vehicle and Cyclist (m/s) | RelSpeedSC | 12.26 | 2.99 | 4.77 | 3.90 |
| Time-to-Collision between subject vehicle and oncoming vehicle (s) | TTCSO | 7.85 | 3.69 | 7.87 | 3.89 |
| Subject vehicle speed (m/s) | SubjectVehSpeed | 18.37 | 2.99 | 10.88 | 3.90 |

* the distance is considered positive when the oncoming traffic is ahead of the subject vehicle and negative when the oncoming traffic is behind the subject vehicle.

## 4.2. Estimation and result validation for the overtaking type predictive model

### *4.2.1. Model estimation*

Prior to model development and estimation, correlation analysis was run to identify any high correlations between the explanatory variables. Significant, but relatively low, correlations were found between AISS score and age ($r = -0.370$; $p = 0.024$), between ordinary violations factor of the DBQ and gender ($r_s = -0.511$; $p=0.001$), and between subject vehicle speed and lateral distance between the subject vehicle and the cyclist ($r = 0.411$; $p < 0.001$). The correlation between the type of overtaking (dependent variable) and the driving speed is significant ($r_s = 0.753$; $p < 0.001$), indicating that the driving speed is a strong predictor of the strategy of overtaking. The correlations among the variables presented in Table 2 guided the creation of different models, the results of which are presented in Table 3.

**Table 3**
Results of the binary logistic regression models (*A1-A3*) for the decision to perform flying or accelerative overtaking (*reference category: accelerative overtaking manoeuvre*), considering random effects while capturing the correlations through the driver-specific error term.

| | ***Model A1 (80-100 m)*** | | | ***Model A2 (50-70 m)*** | | | ***Model A3 (20-40 m)*** | | |
|---|---|---|---|---|---|---|---|---|---|
| | *Coeff.* | *Std. Error* | *Z value*[1] | *Coeff.* | *Std. Error* | *Z value*[1] | *Coeff.* | *Std. Error* | *Z value*[1] |
| ***Fixed Effects*** | | | | | | | | | |
| Intercept ($\beta_o$) | -198.89 | 14.63 | -13.60*** | -167.94 | 16.59 | -10.12*** | -321.12 | 7.66 | -41.90*** |
| Subject Vehicle Speed ($\beta_1$) | 13.55 | 0.94 | 14.28*** | 10.45 | 1.01 | 10.29*** | 20.80 | 0.52 | 39.75*** |
| ***Random Effects*** | | | | | | | | | |
| $\sigma_{b_{on}}$ | | | 103.34 | | | 50.77 | | | 208.1 |
| $\sigma_{b_{oo}}$ | | | 78.81 | | | 52.92 | | | 160.4 |
| ***Model Performance*** | | | | | | | | | |
| Log Likelihood | | | -101.0 | | | -76.9 | | | -67.9 |
| AIC[2] | | | 210.1 | | | 161.8 | | | 143.7 |
| BIC[3] | | | 241.1 | | | 193.5 | | | 175.9 |

[1]Significance codes: *** indicates p<0.001, ** indicates p<0.01, * indicates p<0.05
[2]Akaike information criterion
[3] Bayesian information criterion

The results in Table 3 show that drivers with higher driving speeds are more likely to perform a flying manoeuvre than an accelerative manoeuvre. In addition, model A3 has the lowest AIC value, which indicates, unsurprisingly, that the type of overtaking manoeuvre is best predicted by the driving speed when the subject vehicle is only 20 to 40 metres from the cyclist. We also estimated models that account for additional explanatory variables, such as whether there is an oncoming vehicle or not (dummy variable), the TTC between the subject and the oncoming vehicle, the longitudinal distance from the cyclist, and drivers' characteristics (different factors of the DBQ and AISS scores). However, these additional variables did not improve the models' performances to an extent that it would justify their inclusion in the final selected models. In addition, for reasons of parsimony, since these models may run in real time in active safety systems, simpler predictive models are preferred. Interestingly, as the grouping in Figure 2 illustrates, drivers actually seem to decide on their overtaking strategy when they are more than 100 metres away from the cyclist; therefore, what the model captures is the outcome of the drivers' decision-making process. In other words, the driver may have evaluated the other variables (TTC between the subject and the oncoming vehicles and longitudinal distance) earlier in order to decide which overtaking strategy to use, and thus whether to reduce speed.

#### *4.2.2. Model validation*

The results of the k-fold cross-validation with k=5 for the three models *(A1, A2, A3)* are presented in Table 4. To perform the cross-validation, the 'caret' library (Kuhn, Wing, & Weston, 2015) in R was used with GroupKFold (which ensures that the same group is not represented in both testing and training sets). For more information on the 'caret' library, please refer to (Kuhn, 2015). Based on the k-fold cross-validation, model *A3* considering the data in the range of 20-40 metres performs better than the other two models *A2, A1* considering the data in the ranges of 50-70 and 80-100 metres, respectively. This conclusion is in accordance with the conclusion reached based on the AIC values presented in Table 3.

**Table 4**

Results of the k-fold cross-validation for the two sets of models (k=5)

| Model | 1 | 2 | 3 | 4 | 5 | Average |
|---|---|---|---|---|---|---|
| ***A1*** *(80-100 m)* | 0.864 | 0.879 | 0.811 | 0.851 | 0.827 | **0.847** |
| ***A2*** *(50-70 m)* | 0.826 | 0.865 | 0.902 | 0.883 | 0.899 | **0.875** |
| ***A3*** *(20-40 m)* | 0.851 | 0.809 | 0.991 | 0.882 | 0.973 | **0.901** |

### 4.3. Estimation and validation results of the lateral comfort distance predictive model

#### *4.3.1. Model estimation*

Because the results showed a distinctive difference in the lateral comfort distance for flying versus accelerative manoeuvres (see Figure 3), separate models were developed to estimate the distance for the two manoeuvres. We decided to develop and estimate the models along a distance of 25 metres before and after the cyclist, since the findings by Dozza et al. (2016) indicate that the driver begins to steer away from the cyclist when the longitudinal distance is 16 metres for flying overtaking manoeuvres and 11 metres for accelerative manoeuvres.

Table 5 presents the results of the linear mixed-effect models estimated for flying and accelerative manoeuvres separately as well as the results of the baseline models, which include only an intercept, without any explanatory variables. We compared the models based on the simple rules of thumb defined by Burnham and Anderson (2004), which indicate that when $\Delta_i > 10$ the $i_{th}$ model is not supported, $\Delta_i$ is defined as: $\Delta_i = AIC_i - AIC_{min}$, where $AIC_i$ is the $AIC$ of the $i_{th}$ model, and $AIC_{min}$ is the lowest $AIC$ one obtains among the set of models examined (i.e., the preferred model). The results indicate that the estimated models ($AIC_M$, $BIC_M$) perform better than the baseline models ($AIC_B$, $BIC_B$). We also tested models that include neither an AISS score nor the score on the ordinary violations factor of the DBQ; however, they did not perform better than the ones presented in Table 5.

**Table 5**

Results of the mixed linear models for predicting the lateral comfort distance of the subject vehicle from the cyclist when the longitudinal distance is less than 25 metres, for both flying and accelerative overtaking.

| | ***Flying Overtaking (FO1)*** | | | ***Accelerative Overtaking (AO1)*** | | |
|---|---|---|---|---|---|---|
| | *Coeff.* | *Std. Error* | *t value*[1] | *Coeff.* | *Std. Error* | *t value*[1] |
| ***Fixed Effects*** | | | | | | |
| Intercept ($\beta_o$) | 2.81 | 0.30 | 9.26*** | 1.93 | 0.24 | 8.04*** |
| LongDisSO ($\beta_1$), km. | 2.86 | 0.12 | 23.76*** | -0.28 | 0.04 | -7.02*** |
| LongDisSC ($\beta_2$), km. | 4.48 | 0.44 | 10.18*** | -6.29 | 0.19 | -31.38*** |
| OncomingVeh ($\beta_3$), (1=yes; 0=no) | -1.69 | 0.08 | -19.45*** | 0.21 | 0.02 | 6.71*** |
| AISS ($\beta_4$) | -0.72 | 0.12 | -5.70*** | | | |
| OncomingVeh:AISS ($\beta_5$) | 0.61 | 0.03 | 17.17*** | | | |
| DBQ Ordinary Violation ($\beta_6$) | | | | -0.29 | 0.12 | -2.44** |
| OncomingVeh: DBQ Ordinary Violation ($\beta_7$) | | | | -0.09 | 0.01 | -7.58*** |
| ***Covariance Parameters*** | | | | | | |
| $\sigma_{on}$ (drivers) | | | 0.16 | | | 0.21 |
| $\sigma_{0o}$ (observations) | | | 0.19 | | | 0.29 |
| $\sigma_{\varepsilon}$ (residual) | | | 0.49 | | | 0.53 |
| ***Model Performance*** | | | | | | |
| Log Likelihood$_M$ | | | -64547.7 | | | -63800.1 |
| $AIC_M$ | | | 129113.5 | | | 127618.3 |
| $BIC_M$ | | | 129198.2 | | | 127701.9 |
| ***Baseline Model Performance*** | | | | | | |
| Log Likelihood$_B$ | | | -71010.7 | | | -64644.2 |
| $AIC_B$ | | | 142029.5 | | | 129296.5 |
| $BIC_B$ | | | 142067.2 | | | 129333.7 |

[1]Significance codes: *** indicates p<.001, ** indicates p<0.01, * indicates p<0.05

The results indicate that, during flying overtaking manoeuvres, when the longitudinal distance to the oncoming vehicle is larger, drivers also keep a larger lateral distance from the cyclist while overtaking.

This is expected, as the risk of a head-on collision decreases with increasing longitudinal distance to the oncoming vehicle. However, for accelerative overtaking manoeuvres, the opposite effect was found: the larger the distance from the oncoming vehicle the smaller the lateral distance from the cyclist. This is possibly because in the range of 25 metres from the cyclists in accelerative overtaking, the oncoming vehicle is already behind the subject vehicle, so the distance is recorded as negative. Notice that the coefficient (-0.28) is much smaller than the one for flying overtaking (2.86). Similarly, in flying overtaking manoeuvres, the larger the longitudinal distance between the subject vehicle and the cyclist, the larger the lateral comfort distance—while in accelerative manoeuvres the lateral comfort distance is smaller. This difference between accelerative and flying overtaking manoeuvres can be explained as follows: in the former, while drivers follow the cyclist, they keep a small lateral distance to let the oncoming traffic pass. Once they begin overtaking, the lateral distance increases. For flying overtaking, it is the opposite. Drivers have a large lateral distance to the cyclist before overtaking, to prepare for the passing phase, but the lateral distance gets reduced when they overtake the cyclist, due to the oncoming traffic. The results in Table 5 also show that, when there is an oncoming vehicle that has not yet passed the subject vehicle, drivers who are performing a flying overtaking maintain a smaller lateral distance than those performing an accelerative overtaking. This is probably because of the low number of observations where there was an oncoming vehicle in the range of distance starting at 25 metres from the cyclist. Scores on the AISS were found to be significant for flying manoeuvres: drivers who had higher scores in AISS kept a shorter lateral comfort distance from the cyclist when overtaking. Furthermore, they were willing to get closer to the oncoming vehicle, indicating the essence of sensation seeking. The scores for the ordinary violations DBQ factor were significantly correlated with the accelerative manoeuvres: drivers with higher scores kept a smaller lateral comfort distance from the cyclists. Additionally, there is an interaction effect with the presence of an oncoming vehicle, indicating that these drivers get even closer to the cyclist when an oncoming vehicle is approaching. None of the other factors of the DBQ were found to be statistically significant.

#### *4.3.2. Model validation*

The results of the k-fold cross-validation using the *Root Mean Squared Error (RMSE)* as an accuracy measure for the models are presented in Table 6.

**Table 6**

Results of the *Root Mean Squared Error (RMSE)* in predicting the lateral comfort distance (m.), used in the k-fold cross-validation for the two sets of models (k=5)

| **Model** | **1** | **2** | **3** | **4** | **5** | **Average** |
|---|---|---|---|---|---|---|
| ***FO1*** | 0.50 | 0.49 | 0.57 | 0.60 | 0.62 | **0.56** |
| ***AO1*** | 0.78 | 0.59 | 0.57 | 0.48 | 0.69 | **0.62** |

The validation results indicate that the models cannot accurately predict the lateral comfort distance, although the explanatory variables are statistically significant. This means that there are other explanatory variables beyond the ones considered so far which affect the lateral comfort distance. Further research is needed to identify them.

## 5. Discussion & Conclusions

In this study, we had two main objectives: first to predict the type of overtaking manoeuvre (flying or accelerative) that a driver will perform when approaching a cyclist, and to predict the comfort lateral distance that a driver maintains from the cyclist during the overtaking manoeuvre.

Binary logistic regression models were developed to predict which manoeuvre a driver will perform when approaching a cyclist in the presence of an oncoming vehicle. In all models, the subject vehicle speed has been shown to be a good indicator of the driver's choice, in line with the findings of Bianchi-Piccinini et al. (2018) and Dozza et al. (2016), showing that the subject vehicle speed is different in the two manoeuvres. The suggested explanation for this difference is that drivers adapt their speed once they have decided which overtaking manoeuvre to perform. Our results suggest that this decision is made when drivers are further than 100 metres away from the cyclist. This is a somewhat unexpected result, since it proves that the overtaking strategy is decided on quite early (about 5 s before reaching the cyclist), leaving enough time for both intervention and warning systems to help the driver. Our models account for the correlations among the observations for the same overtaking and the same driver by including an overtaking-specific error term and a driver-specific error term, respectively. The results show that it is very important to take individual variability into account when predicting which overtaking strategy a driver may opt for. The main difference between the three models is the distance from the cyclist (80–100 m, 50–70 m, and 20–40 m). The models' estimation results indicate that the overtaking strategy is best explained by the driving speed when the subject vehicle is 20 to 40 metres from the cyclist (model A3). Moreover, when these models were validated on a new dataset, model A3 performed better than models A1 and A2.

The developed predictive models for the lateral comfort distance showed that the following four factors significantly affect the lateral comfort distance of the subject vehicle when it is -25 to 25 metres away from the cyclist longitudinally: 1) the longitudinal distance between the subject vehicle and the oncoming vehicle, 2) the longitudinal distance between the subject vehicle and the cyclist, 3) the presence of an oncoming vehicle, and 4) the drivers' characteristics. The extent of the impact these variables have on the lateral comfort distance also depends on whether the overtaking manoeuvre is flying or accelerative. Furthermore, higher scores on the AISS and Ordinary Violations DBQ factor significantly decrease the lateral comfort distance. These results highlight the importance of accounting for these variables when developing active safety systems and automated driving. However, it should be noted that the validation results indicate that the models cannot accurately predict the lateral comfort distance, even though the explanatory variables that were investigated are statistically significant. This lack of predictive power indicates that there might be other explanatory variables in addition to those that were considered which also affect the lateral comfort distance and therefore further research is needed on this topic.

### 5.1. Research Methodology

Driving simulators may not always provide ecologically valid results (Boda et al., 2018); nevertheless, they have been informing the design of active safety systems (e.g. for system acceptance (Lubbe & Davidsson, 2015)) and their evaluation (e.g. helping define EuroNCAP scenarios) for several years. Although they provide an artificial environment, driving simulators are still the best place for drivers to experience critical situations without severe ethical and safety concerns. The use of a driving simulator for studying drivers' overtaking strategy meant that the oncoming vehicle speed and TTC could be accurately measured, thus overcoming a limitation of previous studies that used instrumented bicycles (Dozza et al., 2016; Evans et al., 2018). On the other hand, in driving simulator studies it is not possible to investigate the safety perceptions of cyclists while being overtaken by vehicles (as was done, for example, in the study by Llorca et al. (2017) using instrumented bicycles).

Further advances in simulation technology could perhaps link a driving simulator with a cycling simulator in order to investigate these interactions. Furthermore, future naturalistic studies with enhanced sensor technology could measure the distances to the surrounding vehicles from both the cyclist's and driver's perspectives, providing additional data to test the present models. Overall, research of the interactions between drivers and cyclists will benefit from hybrid research approaches combining data collected from different research methods and using improved technology.

### 5.2. Implications for Policy Making, Active Safety, and Automated Driving

In this study, drivers characterized by higher AISS scores maintained shorter lateral comfort distances from the cyclist during flying overtaking manoeuvres. Previous research found that lateral comfort distance (and, in general, all measures of comfort distance from cyclists in all overtaking phases) are reduced during flying overtaking manoeuvres, often below the legal minimum (Bianchi-Piccinini et al., 2018; Dozza et al., 2016; Kovaceva et al., 2018). This dangerous practice calls for those responsible for enforcing regulations and providing driver training programs to educate drivers regarding the implications of their lateral passing distance on cyclists' safety. Furthermore, road authorities can increase drivers' awareness of cyclists and the minimum lateral clearance prescribed by the law by posting additional information on the road, perhaps in the form of warning signs (Dozza et al., 2016).

Our results may also contribute to the development of active safety systems, such as FCW and AEB, by helping to determine thresholds for warnings and interventions that are within driver comfort boundaries (and therefore more likely to be acceptable). Both FCW and AEB may warn a driver approaching a cyclist from behind, or even initiate braking if the driver does not initiate an overtaking manoeuvre in time. For example, the fact that 90% of drivers would have initiated a flying manoeuvre by a certain time (and that drivers preferring an accelerative manoeuvre would have already slowed down by then) may justify a warning from a FCW system at that time. This study confirmed that the start of the overtaking manoeuvre depends on the strategy chosen by the driver, and that this decision is made early enough that warnings and intervention systems can be effective. Threat assessment algorithms for FCW and AEB could, by applying models similar to the ones presented in this paper, predict the driver's overtaking strategy and include this information in the system's decision-making algorithms. Of course, field tests should verify the ecological validity of the models presented in this paper because the models were built on data collected in a virtual environment.

Because our models detect when the decision to perform an overtaking flying manoeuvre is made, they can also improve the threat assessment for potential head-on collisions within the passing phase (Brannstrom et al., 2010). In fact, although this finding needs to be replicated in real-world traffic, it indicates the presence of a large time window within which the driver would be likely to accept a warning. Therefore, FCW (or a mild AEB) may prevent a driver from performing a dangerous overtaking manoeuvre a few seconds before the passing phase—when such an intervention would be useful, acceptable, and safe. This scenario could be included in the Euro NCAP protocol to assess AEB as a driver overtakes a cyclist; today, the protocol only focusses on preventing rear-end collisions with cyclists (EuroNCAP, 2017).

All the models described in this study may improve the design of automated vehicles by guiding them to overtake a cyclist as a human driver would do, without compromising safety. These models may also help an automated vehicle avoid surprising its driver-passenger by tailoring the overtaking manoeuvre to the driver-passenger's individual characteristics (from AISS and DBQ measurements). However, it is worth keeping in mind that the safety of all road users should be prioritized during the development of automated driving. In fact, automated vehicles have the capacity to be safer than human drivers and increase cyclist comfort by adapting the lateral clearance and approaching distance to the cyclist's perceived safety.

### 5.3. Limitations and Future Work

Despite the promising results, this study has some limitations that should be considered in future research. The data in this study was obtained from a driving simulator experiment in Japan, and therefore the results should be validated using naturalistic data from the field and drivers from other countries. Furthermore, testing the ability of the developed models to predict the type of overtaking manoeuvre and the lateral comfort distance based on other datasets would be a stronger validation than the cross-validation analysis conducted in this study. Another limitation of the current study is linked to the realism of the driving environment with respect to the cyclist: a) the cyclist was standing still until the subject's vehicle was within 100 m of the bicycle, b) the cyclist was riding with constant speed and lateral position, and c) the cyclist's appearance was the same in all overtaking manoeuvres. Future driving simulator studies could introduce some variability into the cyclist's behaviour and appearance, in order to examine its impact on drivers' decisions about the type of overtaking manoeuvre and the lateral comfort distance that drivers maintain while overtaking. Connected driving and riding simulators may make it possible for the virtual environment to capture both driver behaviour and the cyclist's perception of that behaviour, during overtaking manoeuvres in different conditions. It has been shown that drivers who do not cycle may have more negative attitudes towards cyclists than drivers who do (Fruhen & Flin, 2015). Future studies should also account for the drivers' cycling experience, since it might influence the type of overtaking manoeuvre and the comfort distance maintained from the cyclist. In this study we have classified the overtaking manoeuvres as either flying or accelerative based on the minimum speed of the subject vehicle, from 100 m away from the cyclist until the driver reached the cyclist using a cut-off threshold of 10 m/s. This could have led to misclassification, and therefore, future research should further investigate the validity of this threshold. Finally, the cycling facility type also plays an important role, as has been shown by Bella and Silvestri (2017); therefore, future studies should also take road design into account when investigating drivers' cyclist-overtaking strategies.

**ACKNOWLEDGMENT**

The authors thank the ADS project sponsored by the Area of Advance Transport at Chalmers. Dr. Haneen Farah thanks the Aspasia Grant Faculty 2016 for funding her research visit to SAFER Vehicle and Traffic Safety Centre at Chalmers. We would like to thank Kristina Mayberry for language revision.

**REFERENCES**


Abe, G., Sato, K., & Itoh, M. (2018). Driver Trust in Automated Driving Systems: the case of overtaking and passing. *IEEE Transactions on Human-Machine Systems, 48*(1), 85-94.

Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and individual differences, 16*, 289-289.

Balanovic, J., Davison, A., Thomas, J., Bowie, C., Frith, B., Lusby, M., . . . Robertson, C. (2016). Investigating the feasibility of trialling a Minimum Overtaking Gap law for motorists overtaking cyclists in New Zealand.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., . . . Grothendieck, G. (2014). Package 'lme4'. *R foundation for statistical computing, Vienna, 12*.

Bella, F., & Silvestri, M. (2017). Interaction driver–bicyclist on rural roads: Effects of cross-sections and road geometric elements. *Accident Analysis & Prevention, 102*, 191-201.

Bianchi-Piccinini, G., Moretto, C., Zhou, H., & Itoh, M. (2018). Influence of oncoming traffic on drivers' decision to overtake cyclists. *Transportation Research Part F: Traffic Psychology and Behaviour, Accepted for publication*.

Boda, C.-N., Dozza, M., Bohman, K., Thalya, P., Larsson, A., & Lubbe, N. (2018). Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare? *Accident Analysis & Prevention, 111*, 238-250.

Boufous, S., de Rome, L., Senserrick, T., & Ivers, R. (2012). Risk factors for severe injury in cyclists involved in traffic crashes in Victoria, Australia. *Accident Analysis & Prevention, 49*, 404-409.

Brannstrom, M., Coelingh, E., & Sjoberg, J. (2010). Model-based threat assessment for avoiding arbitrary vehicle collisions. *IEEE Transactions on Intelligent Transportation Systems, 11*(3), 658-669.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research, 33*(2), 261-304.

Dozza, M., Schindler, R., Bianchi-Piccinini, G., & Karlsson, J. (2016). How do drivers overtake cyclists? *Accident Analysis & Prevention, 88*, 29-36.

Dozza, M., & Schwab, A. (2017). Safety Science Special Issue on Cycling Safety. *Safety science, 92*, 262-263.

EuroNCAP. (2017). Euro NCAP assessment protocol AEB VRU test protocol v2. 0. *Assess. Protoc.*

Evans, I., Pansch, J., & Singer-Berk, L. (2018). Factors Affecting Vehicle Passing Distance and Encroachments While Overtaking Cyclists. *Institute of Transportation Engineers. ITE Journal, 88*(5), 40-45.

Farah, H. (2013). Modeling drivers' passing duration and distance in a virtual environment. *IATSS research, 37*(1), 61-67.

Farah, H., Daamen, W., & Hoogendoorn, S. (2018). How do drivers negotiate horizontal ramp curves in system interchanges in the Netherlands? *Safety science*.

Feng, F., Bao, S., Hampshire, R. C., & Delp, M. (2018). Drivers overtaking bicyclists—An examination using naturalistic driving data. *Accident Analysis & Prevention, 115*, 98-109.

Fishman, E., Schepers, P., & Kamphuis, C. B. M. (2015). Dutch cycling: quantifying the health and related economic benefits. *American journal of public health, 105*(8), e13-e15.

Fruhen, L. S., & Flin, R. (2015). Car driver attitudes, perceptions of social norms and aggressive driving behaviour towards cyclists. *Accident Analysis & Prevention, 83*, 162-170.

JAF. (2017). *Rules of the road*. Retrieved from

Kay, J. J., Savolainen, P. T., Gates, T. J., & Datta, T. K. (2014). Driver behavior during bicycle passing maneuvers in response to a Share the Road sign treatment. *Accident Analysis & Prevention, 70*, 92-99.

Kovaceva, J., Nero, G., Bärgman, J., & Dozza, M. (2018). Drivers overtaking cyclists in the real-world: Evidence from a naturalistic driving study. *Safety science*.

Kuhn, M. (2015). A Short Introduction to the caret Package. *R Found Stat Comput*, 1-10.

Kuhn, M., Wing, J., & Weston, S. (2015). Package 'caret'. *Classification and regression training*.

Lawton, R., Parker, D., Manstead, A. S., & Stradling, S. G. (1997). The role of affect in predicting social behaviors: The case of road traffic violations. *Journal of applied social psychology, 27*(14), 1258-1276.

Llorca, C., Angel-Domenech, A., Agustin-Gomez, F., & Garcia, A. (2017). Motor vehicles overtaking cyclists on two-lane rural roads: Analysis on speed and lateral clearance. *Safety science, 92*, 302-310.

Lubbe, N., & Davidsson, J. (2015). Drivers' comfort boundaries in pedestrian crossings: A study in driver braking characteristics as a function of pedestrian walking speed. *Safety science, 75*, 100-106.

McCulloch, C. E., & Neuhaus, J. M. (2001). *Generalized linear mixed models*: Wiley Online Library.

NHTSA. (2018). Bicycle Safety [Press release]. Retrieved from https://www.nhtsa.gov/road-safety/bicycle-safety#topic-resources

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package 'nlme'. *Linear and Nonlinear Mixed Effects Models, version*, 3-1.

Pucher, J., & Buehler, R. (2017). Cycling towards a more sustainable transport future. *Transport Reviews, 37*(6), 689-694.

Santacreu, A. (2018). Cycling safety: summary and conclusions of the ITF Roundtable on Cycling Safety, 29-30 January 2018, Paris.

Shackel, S. C., & Parkin, J. (2014). Influence of road markings, lane widths and driver behaviour on proximity and speed of vehicles overtaking cyclists. *Accident Analysis & Prevention, 73*, 100-108.

Summala, H. (2007). Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing. In P. C. Cacciabue (Ed.), *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems* (pp. 189-207). London: Springer London.

SWOV. (2017, June 2017). Cyclists. SWOV Fact sheet.

Team, R. C. (2013). R: A language and environment for statistical computing.

Therneau, T., Atkinson, B., Ripley, B., & Ripley, M. B. (2018). Package 'rpart'. *Available online: cran. ma. ic. ac. uk/web/packages/rpart/rpart. pdf (accessed on 20 April 2016).*

Trafikverket. (2014). *Säkrare cykling - gemensam strategi för år 2014-2020* Retrieved from

Varotto, S. F., Farah, H., Toledo, T., van Arem, B., & Hoogendoorn, S. P. (2017). Resuming Manual Control or Not? Modeling Choices of Control Transitions in Full-Range Adaptive Cruise Control. *Transportation Research Record: Journal of the Transportation Research Board*(2622), 38-47.

Walker, I. (2007). Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. *Accident Analysis & Prevention, 39*(2), 417-425.

Walker, I., Garrard, I., & Jowitt, F. (2014). The influence of a bicycle commuter's appearance on drivers' overtaking proximities: an on-road test of bicyclist stereotypes, high-visibility clothing and safety aids in the United Kingdom. *Accident Analysis & Prevention*, *64*, 69-77.

Wegman, F., Zhang, F., & Dijkstra, A. (2012). How to make more cycling good for road safety? *Accident Analysis & Prevention*, *44*(1), 19-29.

West, B. T., Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: a practical guide using statistical software*: Chapman and Hall/CRC.

Zuckerman, M., Eysenck, S. B., & Eysenck, H. J. (1978). Sensation seeking in England and America: cross-cultural, age, and sex comparisons. *Journal of consulting and clinical psychology*, *46*(1), 139.