



# Investigation the Influence of Risk Factors on the Occurrence of Road Accidents Using the Driver Performance Model

Maria Izabel Santos<sup>1</sup> · Paulo Tadeu M. S. Oliveira<sup>1</sup> · Ana Paula C. Larocca<sup>1</sup>

Received: 22 January 2020 / Accepted: 6 November 2021 / Published online: 10 January 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

Researchers are constantly seeking for a quick, inexpensive, and easy-to-use way to mitigate the severity of road accidents. This study proposes the use of a simulator driver model, i.e., a dynamic vehicle model presented in a fractional factorial experiment for the identification of active factors in road accidents. The investigation was based on an accident database of a real two-lane highway that composed the scenario simulation, and tests were conducted in a 10-km path with 39 tangent-curve configurations. Seven factors, namely curve radius, path profile, path condition, virtual driver skill, speed, period of the day, and carload were considered and three replicates were performed per treatment. Speed and friction coefficient were randomly defined for each treatment within the range established for each level. The first experimental design was a quarter fraction, and 42 accidents were observed in 96 events. Speed exerted the highest influence on their occurrence, followed by curve radius, period of the day, and some second-order interactions. The second part of the experimental design was based on the results of the first. A half fraction factorial design with five factors and 14 replicates per treatment were performed, and 96 accidents were observed in 224 events. Speed exerted the highest influence on their occurrence, followed by period of the day, curve radius, virtual driver skill, and second-order interactions. The study has contributed to the understanding of the way drivers drive vehicles on a road with a small *S*-curve and to assessment procedures for analyses of risk factors and improvements in road safety.

**Keywords** Road safety · Risk factors · Speed · Virtual simulation · Factorial design · Driver performance mode

## Introduction

Several studies have shown information about drivers' attitudes and behaviors is crucial for determining the likelihood of traffic accidents, since the knowledge of road traffic rules and risk factors influences collision involvement [1].

Improvements in driving simulators and their new applications have increased the use of simulations over the past few years, and several studies have focused on their evaluation and validation [2–15].

Regular investigations with volunteers driving simulators provide data sets of limited sizes, often due to seasickness and difficulties in recruiting participants and obtaining sufficient data for statistically sound results. The driver model

in a simulator enables users to assess a driver operating a vehicle within a road project and to identify whether the conditions in the tested design might result in loss of vehicle control (e.g., skidding or rollover).

The goal of the driver/vehicle module is to allow the user to assess the way a driver controls a vehicle within the context of a roadway design and to identify the existence of conditions that might lead to loss of vehicle control (e.g., skidding or rollover) within a given design. The driver performance model estimates a driver's speed and path along a highway with no traffic and records several measures, such as lateral deviation, acceleration, friction demand, and speed. Driver's performance is influenced by cues from the roadway/vehicle system, thus affecting vehicle performance [16–22].

Fractional factorial design (FFD) [22] is one of the advanced statistical methods that minimize the number of tests and maximize that of factors tested [23], and has been used in the traffic safety field, particularly in driving simulator environments.

✉ Ana Paula C. Larocca  
larocca.ana@usp.br

<sup>1</sup> Department of Transportation Engineering, São Carlos School of Engineering, University of São Paulo, São Carlos, São Paulo, Brazil

The replication of a simulation as many times as necessary promotes a better understanding of the impact of each factor considered in an event. Therefore, the designed experiments (design of experiments—DOE) conducted in this study analyzed the combination of the following seven potentially risk factors for the occurrence of traffic accidents [1, 24–29]: (1) curves of 130–230 m radius; (2) path profile, i.e., downhill and uphill; (3) driver's skill, i.e., novice and standard; (4) speed—low (50–70 km/h) and high (110–130 km/h); (5) period of the day, i.e., night and day; and (6) carload—one person (70 kg) and four people (280 kg).

The following three key ideas that are the basis for fractional factorial designs were applied: sparsity principle, according to which main factors and lower-order interactions are more likely to substantially affect responses; projection property, which is a sub-design obtained by the deletion of a complementary set of factors and experimentation; and possibility of combining factors and interactions in a fractional factorial design for estimating the influence of factors [2].

## Methods

### Manned Simulation

Risk factors were investigated by a fixed-base driving simulator that is part of the Sustainable Road Safety Project of the Department of Transportation Engineering from the São Carlos School of Engineering, University of São Paulo. It has a cockpit equipped with a car seat, steering wheel with force feedback, paddle shift, accelerator, brake, and clutch pedals. The cockpit station enabled height and distance adjustments between the seat and the pedals. The car seat was positioned in front of a flat screen of approximately  $1.40 \times 0.80$  m, onto which the virtual environment including a speedometer and lateral mirrors was projected at  $1920 \times 1080$  pixels resolution and 60 frames per second. The projected field of view was  $120^\circ$  (horizontal)  $\times$   $50^\circ$  (vertical). Engine sounds were simulated by speakers toward a more immersive environment. The processing system was integrated with three high-performance computers that worked simultaneously and in real time. Driving scenarios were simulated by ROD Scenario Editor, virtual test drive (VTD), and VI-CarRealTime software programs used in several other studies [11, 30–33].

The test route was a 10-km stretch of a Brazilian two-lane rural highway with 19 curves of varied curvature, deflection angle, direction, and grade. The radius of the curves ranges from 130 to 625 m, and the maximum grade and acclivity are 6% and 8%, respectively.

The maximum velocities allowed in the highway are 60 km/h for commercial vehicles and 80 km/h for light

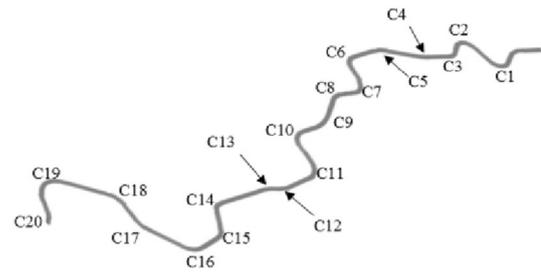


Fig. 1 Curves

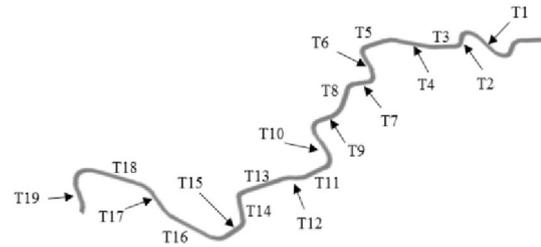


Fig. 2 Tangents

ones. The main probable cause of accidents is driver behavior—47% occur due to driver's errors, and 19% are caused by over-velocity. Passenger cars, commercial vehicles, and pick-ups have been responsible for more than 90% of the accidents.

The virtual stretch was divided according to its geometric elements and then labeled in a numerical order for the identification of dangerous locations—the  $i$ -th curve is called  $C_i$  and the  $j$ -th tangent is  $T_j$  (Figs. 1, 2).

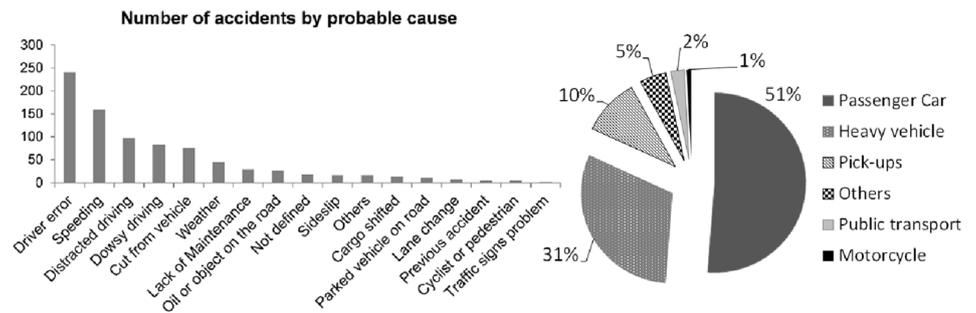
### Model Analysis

A bar-chart analysis revealed the main factors and levels that contribute to the occurrence of accidents (see Fig. 3).

According to Fig. 3,

- (i) Visibility is the distance at which an object is seen and can change in function of light and weather conditions; therefore, occurrences were classified according to period of the day, referred to as day (D) and night (N). 306 occurrences per day and 169 per night were registered.
- (ii) The road profile registered 138 occurrences in downhill stretch, 66 in level, and 271 in uphill stretch.
- (iii) The S-curves in the road are the causes of 87.5% of occurrences. Curves C6 and C14 are sharp and responsible for 50% of the accidents, and their small radius (130 m) does not comply with governmental recommendations [34]. 30 occurrences in the gentle

**Fig. 3** Causes of accidents per type of vehicle between January/2009 and December/2015



curve, 416 in a sharp curve, and 29 on the tangent were registered.

- (iv) Track condition field indicates whether the track was wet or dry at the time of the accident. Although most of the time the track remains dry, the number of accidents on a wet track is three times higher than that on a dry track due to their lower friction coefficient. 105 occurrences under dry conditions and 370 under wet ones were registered.
- (v) Weather conditions are classified into five categories, namely normal, cloudy, foggy, drizzling, and heavy rain, which are grouped into only two: (1) normal conditions, which include normal and cloudy ones, and (2) changed conditions, which encompass foggy, drizzling, and rainy situations. 135 occurrences under average conditions and 340 under “changed” ones were registered.
- (vi) Visibility is the most complex condition and depends on a victim's testimony and officer's analysis. 249 occurrences under good visibility and 224 under partial visibility were registered.

The geometry parameters for the choice of the most significant risk factors were (1) curve with 130 m – 230 m radius; (2) path profile: downhill and uphill; (3) driver' skill: novice and standard; (4) speed: low (50–70 km/h) and high (110–130 km/h); (5) period of the day: night and day; and (6) carload: 1 person (70 kg) and 4 people (280 kg).

**Simulations**

No volunteers were required for this study, since the driving task was performed by a virtual driver of a driving simulator software. VI-CarRealTime® has driver models with different skills, which were used in place of human drivers. Because the driver and their behavior were not the focus of the research, the virtual driver eliminated noise, such as driver distractions, previous learning about the scenery, and motion sickness.

The vehicle model used was a C3—PSA Peugeot Citroën, which encompasses four types of driving skills, namely

novice (inexperienced driver), standard (normal driver), professional (pilot), and robotic (no human skill) [35, 36].

As addressed elsewhere, this study is based on a road traffic accident database of a Federal Highway in Brazil whose maximum speeds allowed are 60 km/h for heavy vehicles and 80 km/h for light ones. The radius of the curves ranges from 130 to a 625 m, and the maximum grade and acclivity are 6% and 8%, respectively. The virtual road was created by VI-Road v17.0® and was based on a real Brazilian highway.

The smallest radius of the curve should be 230 m due to the 80 km/h speed guideline and the maximum elevation rate of the curves. Curves C6 and C14 show the highest rates of traffic accidents.

The average daily traffic (ADT) in the years 2011, 2012, and 2013 was 8845, 9271, and 9233 vehicles, respectively. According to the data available, heavy vehicles contribute to most traffic, and the average traffic speed is 69 km/h ± 13 km/h (average ± standard deviation), including all types of vehicles [37].

The road was modeled by VI-Road v17.0®; each road file has a specific configuration and path definition and includes all road parameters (x, y, and z coordinates, friction, and pavement irregularities).

**Outline of the Statistical Experiment**

**Experimental Design #01: Quarter Factorial Fractional Screening Design**

Due to the number of variables and toward guiding an initial investigation, a factorial design with two levels was used for screening purposes. The definition of both variables (henceforth called factors) and levels was based on road traffic accident databases (Table 1).

The definition of the curve radius (A) levels considered the lowest level at 130 m (the small radius in the baseline stretches of the Highway) and the highest one at 230 m (the minimum curve radius allowed by [34]).

The profile of the path (B) was established according to curve C14, which has an up-and-down slope of 6%, and the conditions of the path (C) were modeled by changing

**Table 1** DOE #01: factors and levels

Factors	Levels	
	(-)	(+)
A. Curve radius	130 m	230 m
B. Path profile	Downhill	Uphill
C. Path conditions	Wet	Dry
D. Driver’s skill	Novice	Standard
E. Speed	Low (50–70 km/h)	High (110–130 km/h)
F. Period of the day	Night	Day
G. Carload	1 person (70 kg)	4 people (280 kg)

**Table 2** Lateral acceleration variation in speeds at the same level

Level	V <sub>1</sub> (km/h)	V <sub>2</sub> (km/h)	Ratio ( $a_2/a_1$ )
(-)	50	70	1.96
(+)	110	130	1.40

**Table 3** Specifications of the  $2_{IV}^{7-2}$  fractional factorial design

Generator	Defining relation	Strings of aliased 2-factor interactions
$F=ABCD$ $G=ABCE$	$I=ABCDF=ABCEG=$ $=DEFG$	$DE+FG$ $DF+EG$ $DG+EF$

Source: Adapted from Box et al. ([19], p. 273)

the friction coefficient and in accordance with the literature. Toward variations, the friction coefficient ranged from 0.3 to 0.5 for a wet pavement, and from 0.7 to 0.9 for a dry [38–42].

Speed (*E*) levels were established in function of the speeds allowed and measured on the stretch of the highway [11]. Lower levels can vary from 50 to 70 km/h and higher ones can range between 110 and 130 km/h. Table 2 shows the lateral acceleration variation between extremes for each level—lateral acceleration can influence the driver’s behavior and must display similar values between levels.

Speed and friction coefficient were used as a source of variation in the experiment. A value within the factor level’s range was set up randomly according to the level for each treatment. VI-Driver<sup>®</sup> established the driver’s skill, and carload was defined according to the car specification, obtained from the manufacturer (PSA Peugeot Citroën).

The  $2^{7-2}$  factorial design used was defined according to its resolution (IV), which enables the main effects to be completely clear from each other and second-order interactions to occur. Third- and higher-order interactions were assumed negligible. Table 3 shows the design specifications and its confounding matrix.

**Table 4** DOE #01: factors and levels

Factors	Levels	
	(-)	(+)
A. Curve radius	130 m	230 m
B. Carload	1 person (70 kg)	4 people (280 kg)
C. Driver’ skill	Novice	Standard
D. Period of the day	Night	Day
E. Speed	Low (60–80 km/h)	High (90–120 km/h)

**Table 5** DOE #02 variables response

Variable response	Measuring technique
$Y_1$ : number of accidents	Count (binary: 0=no occurrence; 1=accident)
$Y_2$ : Path distance	Linear measurement of distance traveled until the occurrence of an accident (continuous)

Treatments were replicated three times, and a completely random design was used to avoid systematic effects. Variable response  $Y_1$  is binary and indicates events with and without accidents—“0” indicates no occurrence of accident and “1” indicates occurrence. A car was under an unsafe condition, here denoted as an accident, when the wheel outside the curve was no longer in the trajectory, i.e., a 1-m lateral displacement was observed, indicating an accident.

DOE planning considered a  $2^{7-2}$  factorial design, whose definition was based on its resolution, which enables the main effects to be completely clear from each other and second-order interactions to occur. Third- and higher-order interactions were assumed negligible.

### Experimental Design #02: Half Factorial Fractional

The planning of the second DOE was based on the results of DEO #01 (see Table 4 for factors and levels), and the main objective was to quantify the factors’ effect on a road traffic accident.

Curve radius (*A*), carload (*B*), driver’s skill (*C*), and period of the day (*D*) levels were the same as those in DOE #01 (item 3.2.4.1). A downhill profile with a 6% slope and 0.8 friction coefficient (dry pavement) was used in all runs. Such two factors were withdrawn from the second DOE, since they exerted no influence on the response variable in DOE #01. Speed (*E*) levels were adjusted toward representing driving within speed limits (lowest level) and the upper tail of the measured speed distribution. Speed (*E*) was set randomly among runs for each treatment for simulating variations.

Table 5 shows the response variables chosen for the measurement and monitoring of DOE.  $Y_1$  is the same used in

**Table 6** Specifications of the  $2_V^{5-1}$  fractional factorial design. Source: Adapted from Box et al. ([19], p. 273)

Generator	Defining relation	Strings of aliased 2-factor interactions
$E=ABCD$	$I=ABCDE$	None

DOE #01, and  $Y_2$  is the linear measurement of the distance traveled until the end of the simulation, which differentiates the severity of each event in accidents.

The experiment was completely randomized and replicated 14 times for each treatment with a different value of speed ( $E$ ) randomly assigned, respecting the interval defined by the level.

DOE planning considered a  $2_V^{5-1}$  fractional factorial design.  $V$  was the resolution of the experiment, which enabled the main effects to be completely clear from each other and second-order interactions (also completely clear from each other) to occur. Table 6 shows the design specifications and its confounding matrix.

Lateral displacement indicated an unsafe condition, as defined in DOE #01.

## Results

This section discusses the results of the design of experiments (DOE).

### DOE #01

Out of the 96 events, 54 included no accidents and 42 involved them. *Multi vari chart* is a graphical representation of a factors' relationship regarding a response variable used in measurement system evaluations and examinations of reproducibility and repeatability, which help in analyses of interactions [38–42]. A systematic effect can be seen in factor **E** (speed), i.e., **no** occurrence is observed in events with **E** (− 1), whereas events with **E** (+ 1) denote occurrences. Three treatments, namely T20, T22, and T26 showed a change of behavior among the passages due to a variation in factors **C** (friction coefficient) and **E** (speed).

Microsoft Excel® estimated contrasts for main effects and second-order interactions (Table 7).  $Y_1$  was replaced by mean response  $y_1$ , which is the probability that  $Y_1 = 1$  for the given factors' levels.

Factors and interactions with contrasts equal to or higher than 0.05 (5%) were considered significant for  $y_1$ , and three-factor or higher-order interactions were negligible. **A** (curve radius), **D** (driver's skill), **E** (speed), **F** (period of the day), and **G** (carload) were important factors. **E** (speed) exerted the most significant effect on the response—seven

**Table 7** DOE #01: contrasts for main effects and second-order interactions for  $y_1$

Contrast summary			
Term	Contrast	Term	Contrast
$E$	0.4375	$AC$	0.0208
$A$	− 0.0625	$AD$	0.0208
$AE$	− 0.0625	$AG$	0.0208
$F$	− 0.0625	$BC$	0.0208
$AF$	− 0.0625	$BD$	0.0208
$DG$	− 0.0625	$BF$	0.0208
$EF$	− 0.0625	$BG$	0.0208
$B$	0.0208	$CD$	0.0208
$AB$	0.0208	$CE$	0.0208
$BE$	0.0208	$CF$	0.0208
$CG$	0.0208	$DE$	0.0208
$C$	0.0208	$DF$	0.0208
$D$	0.0208	$EG$	0.0208
$G$	0.0208	$FG$	0.0208

times greater than the other factors—and **B** (path condition) and **C** (friction coefficient) produced a marginal influence on  $y_1$ , and, therefore, were not considered in the further experiments.

### DOE #02

The variable used in DOE#02, i.e.,  $Y_1$  (frequency of accidents), was the same as that of DOE #01. Accidents were observed in 96 ( $Y_1 = "1"$ ) out of 224 runs, and 128 runs involved no accident ( $Y_1 = "0"$ ). Six treatments showed a change of behavior among passages, indicating the level used [especially factor **E** (speed)] had been more properly adjusted in comparison to DOE #01. The graphical analysis could not reveal whether a main effect or second-order interaction had substantially affected the response, thus indicating a more balanced experiment regarding level choice.

Factors and interactions with contrast equal to or higher than 0.05 (5%) were considered relevant. Three factors and higher interactions were negligible, and **A** (curve radius), **B** (carload), **C** (driver skill), **D** (period of the day), **E** (speed), **AB**, **BC**, and **DE** exerted substantial effects.

The linear model that used factors **A**, **B**, **C**, **D**, **E** and interactions **AB**, **BC**, and **DE** is described as

$$y = 0.4285 + (-0.1428) \cdot A + 0.0267 \cdot B + (-0.0714) \cdot C + (-0.2232) \cdot D + 0.2678 \cdot E + (-0.0625) \cdot AB + (-0.1339) \cdot BC + (-0.0804) \cdot DE$$

where 0.4285 is the  $y_1$  overall mean,  $y$  is the predicted value, and  $A, B, C, D, E, AB, BC, DE$  are replaced by codes (+ 1) and (− 1) per treatment combination.

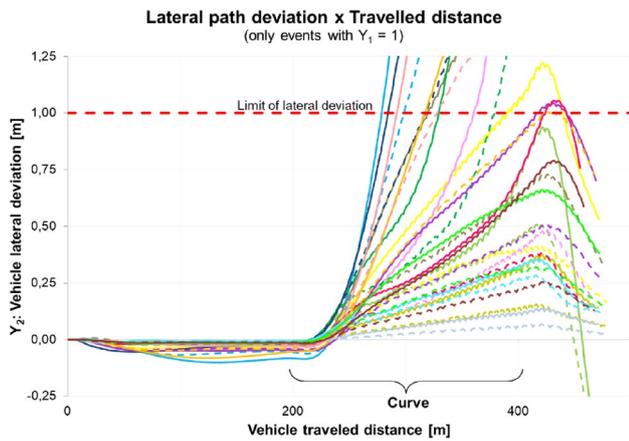


Fig. 4 Lateral path deviation versus traveled distance for runs with  $Y_1 = 1$

$Y_2$  (path distance) is a numerical continuous variable that indicates the severity of an event according to the distance traveled until the accident—the sooner the occurrence of the accident, the more severe the event.

Figure 4 displays the worst and the best runs for each treatment of lateral path deviation versus traveled distance for events with accidents. The dotted horizontal line indicates maximum lateral displacement for a safe condition—data were discarded after the line had been crossed. Solid lines are the shortest distances traveled in each treatment, and dotted lines are the longest ones. The color scale indicates different treatments.

Table 8 shows contrasts calculated for  $Y_2$ —those equal to or higher than 10 were considered relevant. **E** (speed), **D** (period), interaction **ED**, **A** (curve radius), interaction **CB**, and **C** (driver skill), in this order, were the highest.

Table 8 Contrasts calculated for  $Y_2$

Contrasts		
Term	Contrast	Aliases
E	-51,2634	D*A*C*B
D	42,2902	E*A*C*B
A	23,8080	E*D*C*B
C	13,8795	E*D*A*B
B	-9,6652	E*D*A*C
E*D	25,3348	A*C*B
E*A	9,5670	D*C*B
D*A	2,5491	E*C*B
E*C	3,4420	D*A*B
D*C	-1,7902	E*A*B
A*C	-5,2187	E*D*B
E*B	-5,5134	D*A*C
D*B	0,7187	E*A*C
A*B	6,8616	E*D*C
C*B	15,0045	E*D*A

The multi vari chart (Fig. 5) shows the traveled distance, which is the point of the stretch where the accident occurred for events with 1.0 m lateral displacement. Red and yellow lines indicate, respectively, the beginning and the end of the curve. 86 out of 96 events with accidents occurred on the curve, and the other 10 occurred until 25 m after the end of the curve. Treatments T1, T6, T7, and T13 showed the shortest traveled distance, hence, higher severity. The shortest the

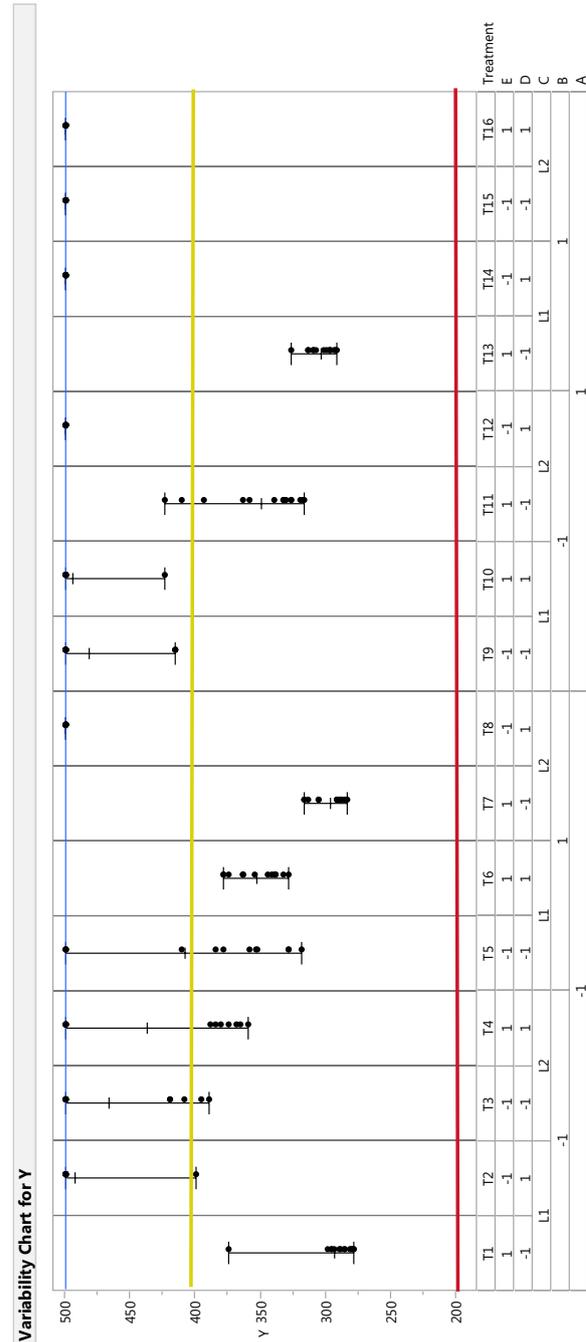


Fig. 5 Multi vari chart of  $Y_2$

distance, the sooner the accident and the more dangerous the combination of factors. Speed (*E*) displayed the highest level for all four treatments.

The prediction expression for  $Y_2$ , calculated by the linear regression method, is

$$y_2 = \begin{cases} 430.00 - 51.26E + 42.29D + 23.81A - 13.88 - 9.67B + 25.33ED - 15.00BC, & \text{for } C = -1 \\ 430.00 - 51.26E + 42.29D + 23.81A + 13.88 - 9.67B + 25.33ED + 15.00BC, & \text{for } C = +1 \end{cases}$$

where 430,00 is the  $Y_2$  overall mean,  $y_2$  is the predicted value, and factors *A*, *B*, *D*, *E*, *BC*, and *ED* are replaced by codes (+ 1) and (− 1) per treatment combination. *C* is a categorical factor, and a different expression must be used for each of its codes.

### Vehicle state variables

The variables analyzed for the monitoring of vehicle’s control and handling stability in a steady-state cornering include sideslip, roll and pitch angles, yaw rate, slip angle of the wheels, and lateral velocity. Roll angle and yaw rate were chosen because they are easy to understand and correlated to the other variables.

Figure 6 shows a summary of the roll angles’ behaviors observed in DOE #2. The red line (1) denotes an event with no accident, and the other curves represent different accident situations before which a hop oscillation occurs at different frequencies and amplitudes, pointing to the severity of the event. The blue line (2) denotes an event in which the driver was able to make a curve, but with wheels (front and rear) off the track, i.e., in a larger radius. The green line (3) indicates a similar situation, but with wheels inside and outside the curve. The pink curve (4) is a rollover in which the tire/ground kept contact with the wheels outside the curve. Finally, the orange line (5) represents a complete rollover.

A roll angle can also show when wheels outside the curve are out of track. The highlighted area in Fig 7 is a

disturbance in the system not observed in other curves. A disturbance is the lowest friction coefficient of “shoulder” that results in different reaction forces in the tire, hence, a roll angle variation.

The yaw rate expresses the speed at which the yaw angle

changes in time—the faster the changes, the more likely the loss of vehicle’s control and handling. Therefore, analyses of the yaw rate require high frequencies and amplitudes. When the yaw rate reaches zero, the vehicle is no longer spinning, and when in cornering, it loses trajectory.

The analysis of the yaw rate curve is similar to that of the roll angle. An abrupt and sudden variation in the curve indicates wheels outside the curve have lost contact with the track. Figure 8 shows examples of yaw rate behaviors observed in DOE #02. The orange line (1) denotes a typical event with no accident, and red (2), blue (3), green (4), and pink (5) curves represent events with accidents. Arrows indicate the moment at which the wheel outside the curve left the track.

Figure 9 displays an influence of speed on both roll angle and yaw rate. Speed can amplify roll angle and yaw rate behavior, and is, therefore, considered one of the most significant factors for traffic accidents.

### Discussion and conclusion

Working with two-level factorial designs requires a correct choice of levels. In DOE #01, speed (*E*) was responsible for almost half of  $y_1$  value, and the adjustments in levels in DOE #02 led to a better understanding of the way factors and the influence of interactions on the response variable.

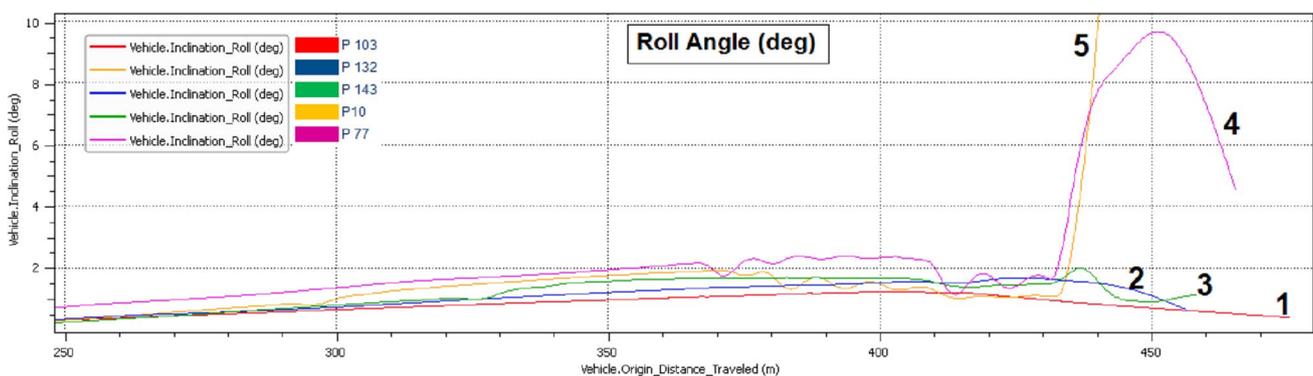


Fig. 6 Summary of the roll angle behaviors observed in DOE #02

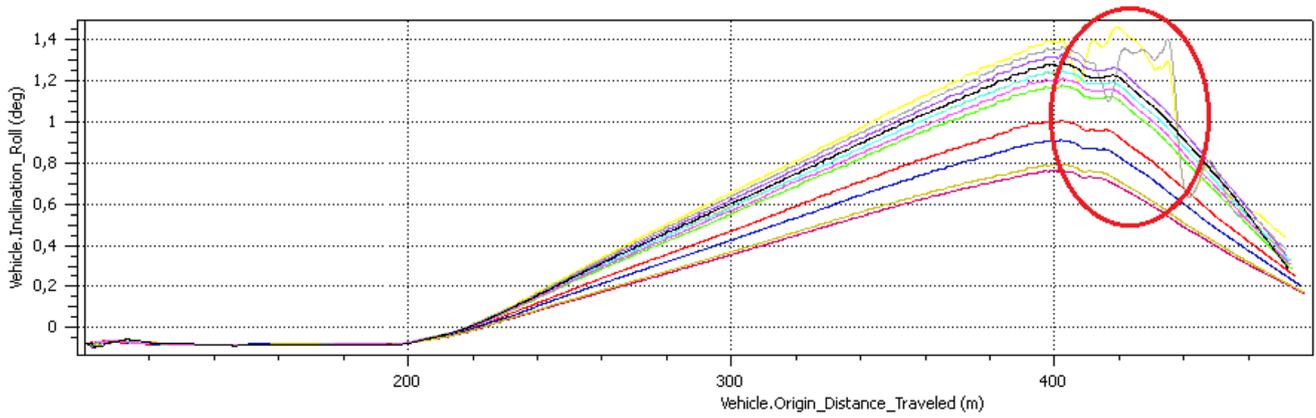


Fig. 7 Roll angle variation due to a variation in normal forces off the path

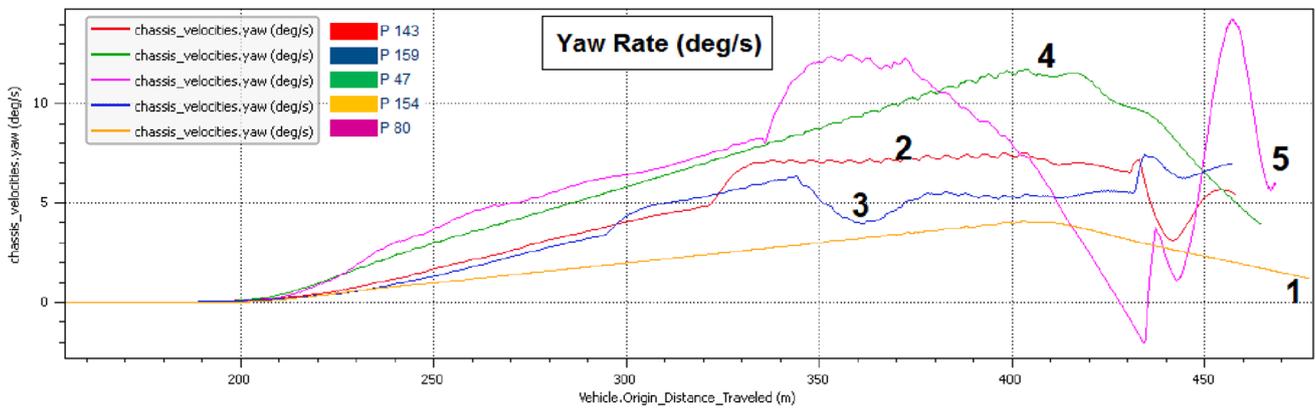
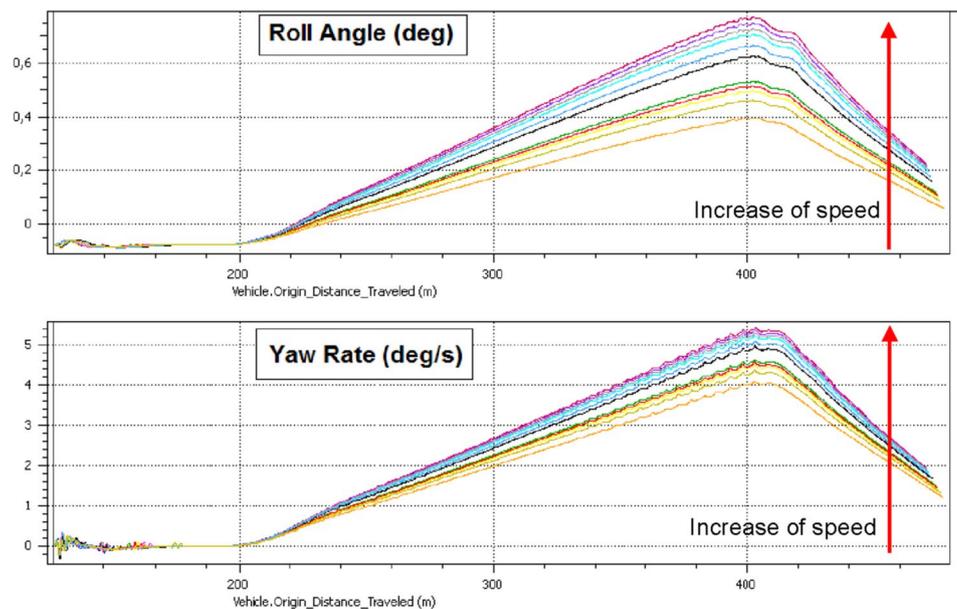
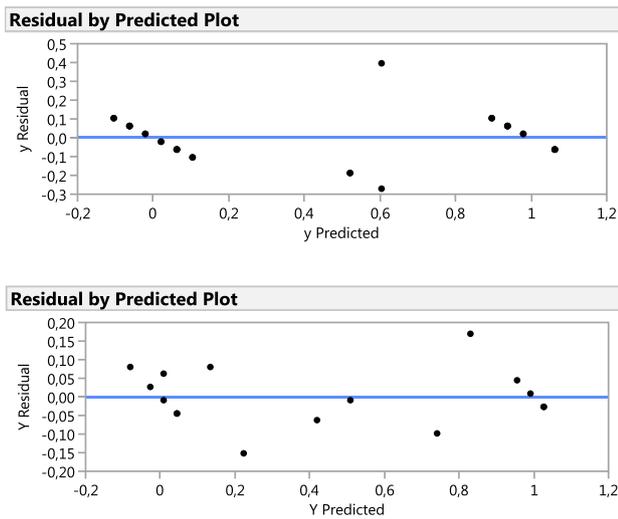
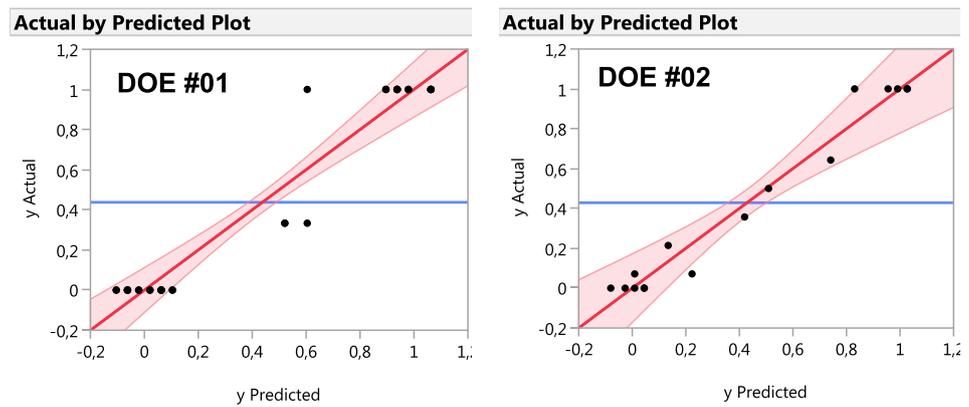


Fig. 8 Summary of yaw rate behaviors observed in DOE #2

Fig. 9 Influence of speed on roll angle and yaw rate



**Fig. 10** Comparison of DOE #01 and #02



**Fig. 11** DOE #01 and DOE #02 residual plot

Simulations are a fully controlled environment with no hassles between events. Replication and fully randomized designs also promote study replications in a virtual environment. Since velocity ( $E$ ) was randomly defined within each level, events with the same velocity values occurred in the same treatment.

DOE #01 was a screening design, and the results should be used with caution and only for guiding other experiments. An increase in replications may increase the occurrence of accidents, thus biasing the results. Figure 10 shows a comparison between the models adjusted for both DOEs. Despite a narrower confidence interval, the adjusted DOE #01 curve has proven bad. The new speed level used in DOE #02 minimized the influences on  $Y_1$ , enabling the identification of other effects.

Figure 11 shows the influence of number of replicates on the results. DOE #02 appears to be more randomized than DOE #01 plot, thus leading to the worst fitted curve and a biased experiment with systematic effects.

This study assessed the risk factors associated with the occurrence of traffic accidents. It was developed in a virtual environment and based on a traffic accident database from the Brazilian Federal Highway.

One of its contributions is the identification of the factors that most affect traffic accidents, based on simulation results from virtual pilots. The first experiment screened and identified the direction of the next experiment. Three main effects, namely speed, curve radius, and period of the day, and four second-order interactions influenced the occurrence of accidents. Such main effects were considered in the second experiment, which also included driver skill and carload, because of their interaction and despite their low significance. Speed, period of the day, curve radius, and driver skill, in this order, were the factors of relative influence on the occurrence of accidents; for the linear probability model, a fifth main effect (carload) was considered due to its interaction with curve radius and driver skill.

Another significant contribution is the design of an experiment that evaluates the influence of one factor at a time and its interactions with the response variable. Fractional factorial experiments increased the number of factors considered and the number of repetitions performed.

Although it was not the focus of this research, a third contribution is related to the use of virtual drivers for road safety studies, since no mathematical formulation has been able to accurately describe human behavior, which requires naturalistic research.

The virtual driver provided distinctions in the level of driving skills. According to the project, this tool can be useful for screening, since it reduces time and the requirement of human resources during the early stages of research. Moreover, it can test experiments prior to their startup as a kick-off tool, as well as response variables, levels of factors, geometric design variations, and other factors.

Our results are valid only for the vehicle model used in this study; however, similar vehicles of the same category from other companies are expected to show similar

performance. Some road boundary conditions that can affect the results are slopes, bank considerations, curves of radii below 130 m, lane width, shoulders, and friction coefficient.

Future research will involve the replication of the experiment with a larger and diverse group of volunteers toward a better understanding of the risk factors for the occurrence of road accidents.

**Acknowledgements** The authors acknowledge the University of São Paulo for the support provided to the research activities, São Paulo Research Foundation (FAPESP) for the financial support, Grants 2013/25034-5; 2016/13736-3; CNPq Grants 307772/2019, 307772/2019-5; CAPES Grant 001; MULTITTECH Engineering and VI-grade. The authors also thank all the volunteers who participate in the research and the employees of the Technical Section of Informatics (*Seção Técnica de Informática—STI*) of the São Carlos School of Engineering.

**Funding** Funding was provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (001), FAPESP (2013/25034-5), Conselho Nacional de Desenvolvimento Científico e Tecnológico (303955/2016-3), FAPESP (2016/13736-3), Conselho Nacional de Desenvolvimento Científico e Tecnológico (307772/2019), Nacional de Desenvolvimento Científico e Tecnológico (143124/2018-8).

## References

- AASHTO (2010) The highway safety manual, American Association of State Highway Transportation Professionals. AASHTO, Washington, DC
- Van Der Horst R, De Ridder S (2007) Influence of roadside infrastructure on driving behavior: driving simulator study. *Transp Res Rec* 2018(1):36–44. <https://doi.org/10.3141/2018-06>
- Lana E, Golias J (2002) Differences between vehicle lateral displacement on the road and in a fixed-base simulator. *Hum Factors* 44(2):303–313. <https://doi.org/10.1518/0018720024497899>
- Easa S, Ganguly C (2005) Modeling driver visual demand on complex horizontal alignments. *J Transp Eng* 131(8):583–590
- Bella F (2005) Validation of a driving simulator for work zone design. *Transp Res Rec* 1937(1):136–144. <https://doi.org/10.1177/0361198105193700119>
- Torres ALM (2015) Análise de consistência de traçado de uma rodovia de múltiplas faixas. Doctoral dissertation, Universidade de São Paulo. <https://doi.org/10.11606/D.3.2016.tde-21062016-153408>
- Bella F (2009) Can driving simulators contribute to solving critical issues in geometric design? *Transp Res Rec* 2138(1):120–126. <https://doi.org/10.3141/2138-16>
- Domenichini L, Branzi V, Meocci M (2018) Virtual testing of speed reduction schemes on urban collector roads. *Accid Anal Prev* 110:38–51. <https://doi.org/10.1016/j.aap.2017.09.020>
- Ronen A, Yair N (2013) The adaptation period to a driving simulator. *Transp Res F Traffic Psychol Behav* 18:94–106. <https://doi.org/10.1016/j.trf.2012.12.007>
- Bham GH, Leu MC, Vallati M, Mathur DR (2014) Driving simulator validation of driver behavior with limited safe vantage points for data collection in work zones. *J Saf* 49:53.e1
- Larocca APC, Ribeiro RL, da Cruz Figueira A, Lulio LC, Rangel MAC (2018) Analysis of perception of vertical signaling of highways by drivers in a simulated driving environment. *Transp Res F Traffic Psychol Behav* 58:471–487. <https://doi.org/10.1016/j.trf.2018.06.034>
- Balci O (1994) Validation, verification, and testing techniques throughout the life cycle of a simulation study. *Ann Oper Res* 53(1):121–173. <https://doi.org/10.1007/BF02136828>
- Khademi N, Choupani AA (2018) Investigating the road safety management capacity: toward a lead agency reform. *IATSS Res* 42(3):105–120. <https://doi.org/10.1016/j.iatssr.2017.08.00>
- Dai R, Lu Y, Ding C, Lu G, Wang Y (2019) A simulation-based approach to investigate the driver route choice behavior under the connected vehicle environment. *Transp Res F Traffic Psychol Behav* 65:548–563. <https://doi.org/10.1016/j.trf.2018.04.008>
- Kuniyoshi JRG, Costa AT, Figueira AC, Kabbach FI Jr, Larocca APC (2021) Driver's visual perception as a function of age Using a driving simulator to explore driver's eye movements in vertical signs. *Transp Res Interdiscip Perspect*. <https://doi.org/10.1016/j.trip.2021.100460>
- Marzougui D, Mohan P, Kan CD, Opiela K (2007) Performance evaluation of low-tension three-strand cable median barriers. *Transp Res Rec* 2025(1):34–44. <https://doi.org/10.3141/2025-03>
- Stine JS, Hamblin BC, Brennan SN, Donnell ET (2010) Analyzing the influence of median cross-section design on highway safety using vehicle dynamics simulations. *Accid Anal Prev* 42(6):1769–1777. <https://doi.org/10.1016/j.aap.2010.04.018>
- De Ridder S, Van der Horst R, Naing CL, Thomson R, Fagerlind H, Lanner G, Wink W (2006) Roadside infrastructure for safer European roads: D04 Envelope of vehicle and driver response prior to collisions. Loughborough University. Report. <https://hdl.handle.net/2134/2203>
- Box GE, Hunter JS, Hunter WG (2005) Statistics for experimenters: design, innovation, and discovery, vol 2. Wiley-Interscience, New York
- Rangel MAC (2015) Análise da percepção da sinalização vertical por parte do condutor, utilizando ambientes simulados de direção: um estudo de caso na rodovia BR-116. Doctoral dissertation, Universidade de São Paulo. <https://doi.org/10.11606/D.18.2015.tde-29072015-101734>
- Levison WH, Campbell JL, Kludt K, Bittner AC, Potts IB, Harwood DW, Schreiner CS (2007) Development of a driver vehicle module (DVM) for the interactive highway safety design model (IHSDM) (no. FHWA-HRT-08-019). United States. Federal Highway Administration. Office of Research and Technology Services
- Dozza M, Boda CN, Jaber L, Thalya P, Lubbe N (2020) How do drivers negotiate intersections with pedestrians? The importance of pedestrian time-to-arrival and visibility. *Accid Anal Prev* 141:105524. <https://doi.org/10.1016/j.aap.2020.105524>
- Machado-León JL, Oña J, Oña R, Ebohi L, Mazzulla G (2016) Socio-economic and driving experience factors affecting drivers' perceptions of traffic crash risk. *Transp Res F Traffic Psychol Behav* 37(2016):41–51. <https://doi.org/10.1016/j.trf.2015.11.010>
- Canale AC (1989) Automobilística: dinâmica e desempenho. Érica, São Paulo, pp 57–77
- Plöchl M, Edelmann J (2007) Driver models in automobile dynamics application. *Veh Syst Dyn* 45(7–8):699–741
- Jürgensohn T (1997) Hybride Fahrermodelle. Pro-Universitäts-Verlag, Würzburg
- Macadam CC (2003) Understanding and modeling the human driver. *Veh Syst Dyn* 40(1–3):101–134
- Hoffmann ER (1976) Human control of road vehicles. *Veh Syst Dyn* 5(1–2):105–126
- Reid LD (1983) A survey of recent driver steering behavior models suited to accident studies. *Accid Anal Prev* 15(1):23–40. [https://doi.org/10.1016/0001-4575\(83\)90004-0](https://doi.org/10.1016/0001-4575(83)90004-0)

30. Figueira AC, Larocca APC (2020) Analysis of the factors influencing overtaking in two-lane highways: a driving simulator study. *Transp Res F Traffic Psychol Behav* 69:38–48. <https://doi.org/10.1016/j.trf.2020.01.006>
31. Vieira FS, Larocca APC (2017) Drivers' speed profile at curves under distraction task. *Transp Res F Traffic Psychol Behav* 44:12–19. <https://doi.org/10.1016/j.trf.2016.10.018>
32. Santos MI, Meira PT, de Oliveira S, Ribeiro RL, Larocca APC, Junior FIK (2017) Conceito, configuração e aplicação de um simulador de direção no Brasil-Estudo de caso. *Transportes* 25(2):1–14. <https://doi.org/10.14295/transportes.v25i2.1174>
33. Figueira AC, Larocca APC (2020) Proposal of a driver profile classification in relation to risk level in overtaking maneuvers. *Transp Res F Traffic Psychol Behav* 74:375–385. <https://doi.org/10.1016/j.trf.2020.08.012>
34. DNER (1999) Manual de projeto geométrico de rodovias rurais. IPR, Rio de Janeiro
35. VI-GRADE. VI—CarRealTime 17.0: Documentation—VI—CarRealTime framework—build mode. Germany: VI-grade engineering software and services 2015.
36. Pacejka H (2005) *Tire and vehicle dynamics*. Elsevier, Amsterdam
37. Irmscher M, Jürgensohn T, Willumeit HP (1999) Driver models in vehicle development. *Veh Syst Dyn* 33(sup1):83–93. <https://doi.org/10.1080/00423114.1999.12063072>
38. Bittner AC Jr, Simsek O, Levison WH, Campbell JL (2002) On-road versus simulator data in driver model development driver performance model experience. *Transp Res Rec* 1803(1):38–44. <https://doi.org/10.3141/1803-06>
39. Hamblin BC (2007) Vehicle rollover prediction from low order dynamic models or varying terrain encountered during median traversal. Doctoral dissertation, Pennsylvania State University
40. Edara R, Shih S (2004) Effective use of multibody dynamics simulation in vehicle suspension system development (no. 2004-01-1547). SAE technical paper. <https://doi.org/10.4271/2004-01-1547>
41. Cheng C-S (2006) Projection properties of factorial designs for factor screening. In: Dean A., Lewis S (eds) *Screening*. Springer, New York, NY. [https://doi.org/10.1007/0-387-28014-6\\_7](https://doi.org/10.1007/0-387-28014-6_7)
42. Cavazzuti M (2012) *Optimization methods: from theory to design scientific and technological aspects in mechanics*. Springer Science & Business Media, Berlin

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.