UNIVERSITY OF SÃO PAULO
SÃO CARLOS SCHOOL OF ENGINEERING

MARIA IZABEL DOS SANTOS

Identifying active factors by a fractioned factorial experimental design
and simulation in road traffic accidents

São Carlos
2017

MARIA IZABEL DOS SANTOS

Identifying active factors by a fractioned factorial experimental design
and simulation in road traffic accidents

Revised version

(Original version is available at São Carlos School of Engineering)

A thesis submitted for the degree of
Master of Science to Department of
Transportation Engineering, São Carlos
School of Engineering, from University of
São Paulo.

Advisor: Professor Ana Paula Camargo
Larocca

São Carlos

MARIA IZABEL DOS SANTOS

Identificação de fatores ativos em acidentes rodoviários por experimento
fatorial fracionado e simulação

Versão Corrigida

(Versão original encontra-se na Escola de Engenharia de São Carlos)

Tese apresentada à Escola de
Engenharia de São Carlos da
Universidade de São Paulo, como
requisito para a obtenção do Título de
Mestre em Engenharia de Transportes.

Orientadora: Prof^a. Dr^a. Ana Paula
Camargo Larocca

São Carlos

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

S237i Santos, Maria Izabel dos
Identifying active factors by a fractioned
factorial experimental design and simulation in road
traffic accidents / Maria Izabel dos Santos;
orientadora Ana Paula Camargo Larocca. São Carlos,
2017.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia de Transportes e Área de Concentração em
Infraestrutura de Transportes -- Escola de Engenharia
de São Carlos da Universidade de São Paulo, 2017.

1. road safety. 2. simulation. 3. design of
experiments. 4. virtual driver. I. Título.

FOLHA DE JULGAMENTO

Candidata: **MARIA IZABEL DOS SANTOS**

Título da dissertação: " Identificação de fatores ativos em acidentes rodoviários por um experimento fatorial fracionado e simulação"

Data da defesa: 05/07/2017

Comissão Julgadora:

Resultado:

Prof. Titular **José Reynaldo Anselmo Setti**
(Presidente designado)
(Escola de Engenharia de São Carlos/EESC)

Aprovada

Profª. Titular **Linda Lee Ho**
(Escola Politécnica/EP – USP)

Aprovada

Prof. Dr. **Roberto Bortolussi**
(Centro Universitário FEI)

Aprovado

Coordenadora do Programa de Pós-Graduação em Engenharia de Transportes:

Profª. Associada **Ana Paula Camargo Larocca**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Luis Fernando Costa Alberto**

ACKNOWLEDGMENTS

I would like to thank my thesis advisor Professor Ana Paula Camargo Larocca who office was always open whenever I ran into a trouble spot or had a question about anything, and for the given opportunity. Also I would like to thank her for the confidence in my work, allowing it to be result of my efforts and learning.

I must express my gratitude to all employees of Transportation Engineering Department for the support and for provide the infrastructure needed. I would especially like to thank the faculty of the department for sharing their knowledge which was certainly valuable to the final result of this work. Thank you also to my colleagues for sharing knowledge and experiences, especially Paulo Tadeu Oliveira.

I would like to express the deepest appreciation to committee members Professor Linda Lee Ho, Professor Roberto Bortolussi and Professor Antonio Carlos Canale for their contribution to the research.

My acknowledgment for the support from Brazilian Federal Government through CAPES, for providing funding for this research and for VI-Grade for providing software licenses whenever needed and for PSA – Peugeot Citroën for providing vehicle data and model.

I would like to thank to my whole family: parents and sisters and Álvaro' family for constantly giving me strength to continue on. This accomplishment would not have been possible without them.

Finally, I would like to express my very profound gratitude to my husband, Álvaro, for providing me with unfailing support and continuous encouragement throughout the years of study and through the process of researching and writing this thesis; And to my beloved daughter Maria for being the reason of everything in my life.

EPIGRAPH

“Knowledge is power”

Sir Francis Bacon (Box *et al*, 2015)

ABSTRACT

SANTOS, M. I. **Identifying active factors by a fractioned factorial experimental design and simulation in road traffic accidents.** 2017. 143 f. Thesis (Master) – São Carlos School of Engineering, University of São Paulo, São Carlos, 2017.

Researchers around the world are constantly seeking for a quick, inexpensive and easy to use way to understand road traffic deaths. This study proposes the use of multibody (MBS) simulation, using a virtual driver, associated to fractional factorial experiments to identify active factors in road traffic accidents. The objectives of this work were to: (i) use DOE to show a more structured direction on the studies of road safety and (ii) investigate possible vehicle state variables to monitor vehicle dynamic stability. The first experiment was a quarter fraction It was designed based on an accident database of a Brazilian Federal Highway. Seven factors were considered (curve radius, path profile, path condition, virtual driver skill, speed, period of the day and car load) and 3 replicates were performed per treatment. Speed and friction coefficient were defined randomly for each treatment, within the defined range for each level. 42 accidents were observed in 96 events. Speed had shown the highest influence on the occurrence, followed by curve radius, period of the day and some second order interactions. The second experiment was based on the results of first one. A half fraction factorial design with five factors (curve radius, car load, virtual driver skill, period of the day and speed), with 14 replicates per treatment, was performed. Speed was defined randomly as per previous experiment. 96 accidents were observed in 224 events. Speed had the highest influence on the occurrence of accidents, followed by the period of the day, curve radius, virtual driver skill and second order interactions. Speed is also pointed by World Health Organization as one of the key factors for the occurrence of accidents. The study indicates that a well-designed experiment with a representative vehicle model can show a direction for further researches. At last, roll angle, yaw rate and displacement of the car on the road are variables suggested to be monitored in experiments using simulation to identify vehicle's instability

Keywords: Road safety, simulation, design of experiments, virtual driver.

RESUMO

SANTOS, M. I. **Identificação de fatores ativos em acidentes rodoviários por experimento fatorial fracionado e simulação**. 2017. 143 f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2017.

Pesquisadores do mundo estão constantemente buscando uma maneira rápida, barata e fácil de usar para entender acidentes de trânsito. O presente estudo propõe o uso de simulação, condutor virtual e experimentos fatoriais para a identificação de fatores ativos em acidentes rodoviários. Os objetivos deste trabalho foram: utilizar experimentos planejados, associado a simulação para obter uma direção para estudos futuros e investigar possíveis variáveis de estado do veículo a serem usadas para monitorar sua estabilidade dinâmica. Para tal, foi utilizado um modelo completo de veículo validado e dados reais de acidentes de um determinado trecho de rodovia brasileira. O primeiro experimento baseou-se em um banco de dados de acidentes de uma rodovia Federal brasileira. Optou-se por fracionar o experimento, utilizando um quarto de fração. Sete fatores foram considerados (raio da curva, perfil da pista, condição da pista, habilidade do condutor virtual, velocidade, período do dia e carga do carro) e foram realizadas três réplicas por tratamento. Velocidade e coeficiente de atrito foram utilizados como fontes de variação do experimento: para cada tratamento, e dentro do intervalo definido para cada nível, ambos foram definidos aleatoriamente. Em 54 dos 96 eventos foram observados acidentes. Velocidade, raio da curva, período do dia e algumas interações de segunda ordem foram os fatores com maior influência na ocorrência de acidentes. O segundo experimento utilizou como dado de entrada os resultados obtidos no experimento anterior. O experimento foi fracionado, meia fração, com cinco fatores (raio da curva, carga do carro, habilidade do motorista virtual, período do dia e velocidade). Foram realizadas 14 réplicas por tratamento, e a velocidade foi mantida como fonte de variação. Em 96 dos 224 eventos foram observados acidentes. Velocidade teve maior influência na ocorrência de acidentes, seguida por período do dia, raio da curva, habilidade do motorista virtual e interações de segunda ordem. A velocidade também é apontada pela Organização Mundial da Saúde como um dos fatores-chave para a ocorrência de acidentes. Isto indica que um experimento bem planejado, com um modelo de veículo representativo, pode apontar uma direção a ser seguida em pesquisas futuras. Por último é sugerido o

monitoramento do ângulo de rolagem (*roll angle*), da taxa de guinada (*yaw rate*), e do deslocamento lateral do carro na pista para identificar instabilidades no veículo quando são utilizadas simulações.

Palavras-chave: Segurança viária, simulação, experimentos planejados, condutor virtual.

LIST OF FIGURES

Figure 1 - Population, road traffic death and registered vehicles by country income.	26
Figure 2 - Deaths by user category and trends in road traffic deaths in Brazil.	28
Figure 3 – Panoramic view of baseline Federal Highway.	37
Figure 4 – Indication of curves with smaller radius than the allowed radius.	38
Figure 5 – Distribution of vehicles for the given stretch.	38
Figure 6 – Number of accidents per year in the stretch of the Highway.	39
Figure 7 – Number of accidents by probable cause in Jan/2009 and Dec/2015.	40
Figure 8 – Number of accidents per type of vehicle in Jan/2009 and Dec/2015.	40
Figure 9 – Curves identifications.	41
Figure 10 – Tangents identifications.	41
Figure 11 – Accidents distribution per geometric element (Jan/2009 - Dec/2015). ...	42
Figure 12 – Police records' probable cause distribution (Jan/2009 – Dec/2015).	42
Figure 13 – Analysis of occurrences per period: day (D) and night (N).	43
Figure 14 – Analysis of occurrences per track profile.	44
Figure 15 – Analysis of occurrences per lane route.	44
Figure 16 – Analysis of occurrences per track condition.	45
Figure 17 – Analysis of occurrences per weather condition.	45
Figure 18 – Analysis of occurrences per visibility condition.	46
Figure 19 – SAE vehicle axis system.	46
Figure 20 – Examples of different roll angles.	48
Figure 21 – Examples of different yaw rate.	48
Figure 22 – Datamodel representing MBS.	49
Figure 23 – Rigid parts of vehicle model.	50
Figure 24 – Example of suspension modeling.	51
Figure 25 – Curve and indication of MF-Tire parameters.	52
Figure 26 – Bicycle model.	54
Figure 27 – Effect of preview time in vehicle trajectory.	55
Figure 28 – Human driver cognitive model.	55
Figure 29 – Road and path example: front iso (top) and lateral views (bottom).	56
Figure 30 – Definition of an unsafe condition.	62
Figure 31 – DOE #01 Multi Vari Chart for Y_1	67
Figure 32 – DOE #02 Multi Vari Chart for Y_1	69
Figure 33 – Lateral path deviation <i>versus</i> travelled distance for runs with $Y_1 = 1$	71
Figure 34 – Multi vari chart of Y_2	72
Figure 35 – Summary of roll angle behavior observed in DOE #02.	74
Figure 36: Roll angle variation due to normal forces variation when off track.	75
Figure 37 – Summary of yaw rate behavior observed in DOE #2.	76
Figure 38 – Example of overlapping curves of DOE #02.	78
Figure 39 – DOE #01 and #02 actual by predicted plot comparison.	79
Figure 40 – DOE #01 and DOE #02 residual plot.	79
Figure 41 – DOE #02 - Residual and predicted plot for Y_2	80

Figure 42 – Example of tire normal forces plots and its relation with roll angle.	81
Figure 43 – Tire normal force: identification of the point where tires loose contact with ground.	82
Figure 44 – Influence of speed in roll angle and yaw rate.	82
Figure 45: Roll angle and Yaw Rate relation.	83
Figure 46: Driver demands.	84
Figure 47: Path lateral deviation for different drivers skill.	84

LIST OF TABLES

Table 1 – MF-Tire coefficients calculation.	52
Table 2 – C3 model technical specifications.....	53
Table 3 – DOE #01: factors and levels.	60
Table 4 – Lateral acceleration variation of speeds at the same level.	61
Table 5 – Specifications of the 2_{IV}^{7-2} fractional factorial design.	62
Table 6 – DOE #01: factors and levels.	63
Table 7 – DOE #02 variables response.....	63
Table 8 – Specification of the 2_V^{5-1} fractional factorial design.....	64
Table 9 – Analysis procedure for a 2^k design.	65
Table 10 – DOE #01: main effects and second order interactions contrasts for y_1	66
Table 11 – DOE #02 Main effects and second order interactions contrasts for y_1	70
Table 12 – Calculated contrasts of Y_2	71
Table 13 – Passages from DOE #02 grouped by speed.	78

LISTA DE ABREVIATURAS E SIGLAS

ADT	Average Daily Traffic
BRS	Body Reference System
CRT	CarRealTime®
DNER	Departamento Nacional de Estradas de Rodagem
DOE	Design of Experiments
DOF	Degree of Freedom
FHWA	Federal Highway Administration
FRD	Factor Relationship Diagram
F-Tire	Magic Formula Tire Model
GRS	Global Reference System
HCM	Highway Capacity Mannual
MBS	Multibody System
OFAT	One Factor At Time
PSA	Peugeot Citroën Group
PT	Preview Time
WHO	World Health Organization

CONTENTS

1	INTRODUCTION	25
1.1	Justification	27
1.2	Research hypotheses	30
1.3	Research objectives.....	30
1.3.1	Research secondary objectives	30
1.4	Research scope.....	31
2	REVIEW OF LITERATURE	33
3	METHODOLOGY	37
3.1	The Highway.....	37
3.1.1	Geometric design of the Highway.....	37
3.1.2	Road traffic accident data	39
3.2	Basics of vehicle dynamics	46
3.2.1	Roll angle	47
3.2.2	Yaw rate.....	48
3.3	Multibody system methodology	49
3.3.1	Multibody system modeling.....	49
3.3.2	Vehicle model.....	50
3.3.3	Virtual driver	53
3.3.4	Road models and events	56
3.4	Factorial design: basic definitions and principles.....	56
3.5	Experiment outline.....	60
3.5.1	Design of the experiment #01 – Quarter fractional factorial screening design.....	60
3.5.2	Design of the experiment #02 – Half fractional factorial design	62
4	RESULTS	65
4.1	DOE #01.....	65
4.1.1	Y ₁ : Frequency of accidents.....	65
4.2	DOE #02.....	68
4.2.1	Y ₁ : Frequency of accidents	68
4.2.2	Y ₂ : Path distance.....	70
4.3	Vehicle state variables.....	73
4.3.1	Roll angle	73
4.4	Yaw rate.....	75
5	DISCUSSION	77

5.1	Designed experiments and simulation.....	77
5.2	Response variable	80
5.3	Vehicle state variables.....	81
5.4	Virtual driver	83
6	CONCLUSION.....	85
6.1	Recommendation for further studies	87

1 INTRODUCTION

The present work is a set of researches performed in a virtual environment that evaluate road safety in Brazil. It also deals with interaction of drivers with the highways and their environments. This study makes an assessment of some factors and their relation with the occurrence of accidents. Since 1960 (Allen *et al.*, 2011), simulations have been used for vehicle dynamics studies throughout the world. With the advances of virtual environments, the application of simulations has increased. For example, physicians used simulation in the evaluation of Alzheimer's patients (Frittelli *et al.*, 2016).

In 2015, World Health Organization (WHO) listed ten facts about road safety around the world. One of the facts is that by controlling the vehicle speed, the severity of injuries and deaths can be reduced. *Speeding* is listed as one of the key risk factors for road safety injuries by the WHO. Drink-driving, negligence in use of motorcycle helmets, seat belts and child restraints complete the list of risk factors (WHO, 2015).

Low and middle income¹ countries have 90% of road traffic deaths. Brazil is the fourth ranked in the number of deaths reaching to approximately 44 thousand. Middle-income countries hold 53% of the registered motorized vehicles are responsible for 74% of deaths in road traffic worldwide (Figure 1) (WHO, 2013).

WHO, in its annual report on road safety (WHO, 2015), says that strengthening road safety legislation reduces road traffic crashes, injuries and deaths and improves driver behavior. Seventeen countries have changed laws on risk factors as per WHO. APPENDIX A shows the "Best Practices" as per WHO and its orientations.

In urban areas, if a vehicle hits a pedestrian, travelling at the speeds of up to 50km/h reduces the chances of death to 20%, as compared to 60% death chances, if the vehicle is moving at 80km/h (WHO, 2015). HCM (AASHTO, 2000), chapter 22, suggests that adjustments should be made in free flow speed depending on climatic conditions. The Federal Highway Administration (FHWA) has a program dedicated to study how weather events impact roads (FHWA, 2016). During the study, 22% of the

---

¹ WHO uses gross national income to categorize into classes: low-income = U\$1045 or less; middle-income = U\$1046 a U\$12745; high-income = U\$12746 or more.

registered occurrences happened due to climatic events (rain, foggy, snow, lateral wind) and the majority of them took place due to wet road. Same situation has been observed around the world (Maze *et al.*, 2006).

Balci (1994) defines simulation as a modeling process of a system with a problem and the study using this model where, the objective is to find a solution by performing virtual experiments. Hence, correct modeling and problem formulation are important to achieve meaningful and representative results. Solving and formulating problems assertively is a challenge for any engineering area. In a large-scale production, swift solutions are vital for a profitable business. The quality of a product is reflection of its productive excellence.

Figure 1 - Population, road traffic death and registered vehicles by country income.

Source: WHO (2015).

Monitor quality by productive samples is common. However, depending on the strategy of collecting samples it is possible that data quality do not show the real picture. In the 80's, several quality-related initiatives were developed and introduced in manufacturing environments (Quality Circles, Zero Defects, Total Quality Management) to make American products competitive with Japanese products induced into American market after the Treaty of the Americas (Raisinghani *et al.*, 2013). At the same time, Motorola realized that poor quality entailed high costs that made its products uncompetitive. To work on this issue, Motorola had set a new aggressive target acceptable for defects as 3.4 parts per million. It was the beginning of Six Sigma Methodology® (6σ).

Harry and Schroeder (2000, p. vii) defined 6σ as “a highly disciplined process that helps a company focus on developing and delivering near-perfect products and services (...) with the ultimate goal of high levels of customer satisfaction”. General Electric, Honda, Bombardier, Polaroid, Hitachi, Sony, Whirlpool Co, among others, had adopted 6σ to increase market share and profit margin, and reduce costs (Harry and Schoeder, 2000). The main application of 6σ is to monitor, control and adjust production in order to maintain quality levels. However, Genichi Taguchi (from Taguchi Methodology) already argued that quality should be considered since design phase, that is, it should be designed and not only monitored (Raisinghani *et al.*, 2013). Sequentially, Taguchi’s approach was put into practice. Nowadays, healthcare, financial, engineering & construction as well as research and development sectors are examples of applications of 6σ principles (Kwak and Anbari, 2006).

Since 2011, the world is living a *Decade of Action for Road Safety*. The goal is to reduce the number of deaths and injuries by half which are occurring due to road traffic accidents. Despite the progress occurred in some countries, a task force will be needed to meet the target. Brazil has the sixth registered motorized vehicle fleet (WHO, 2015). United States of America (USA) lies in the first rank, with three times of Brazilian fleet. China and India are responsible for almost half million of deaths due to road traffic accidents.

1.1 Justification

This research was motivated by the commitment made by the WHO to reduce deaths and injuries in traffic accidents. The aim is also to reduce the deficiency of using virtual simulations for road safety between Brazil and north hemisphere countries. Reliable data about road traffic crashes are needed to correctly identify place and risk and the severity of factors that generate such risks. By evaluation of these data, one can effectively plan and monitor the safety of a road (Harvey *et al.*, 2010). Most countries are able to collect data, but few of them can collect good quality data. That is a serious problem (Evgenikos *et al.*, 2010).

Brazil is ranked in third place among America region countries regards to road traffic accidents (WHO, 2015). In Figure 2 is shown the number of deaths by road user category and the trends in road traffic deaths in Brazil. Drivers and passengers of 4-wheeled cars are responsible for 23% of deaths (pizza chart) and the trend graphic shows no perspective of reaching the goal, as established by ONU.

Figure 2 - Deaths by user category and trends in road traffic deaths in Brazil.

Source: WHO (2015).

One way to study risk factors and its influences in road traffic accidents is using virtual simulations. Once a controlled environment is achieved, data collection process is completely reliable as compared to the field data. Another advantage in using simulations is the capacity to record data of all simulated events, allowing researchers to access those anytime.

The ability to replicate a simulation as many time as needed, allows researchers to better understand the impact of each factor considered in the event. For instance, it is possible to have exactly the same traffic situation, climatic conditions and road conditions for several types of drivers in a virtual environment. Even it is possible to have a comparison of the behavior of the same driver, which can help to understand how behavior of the driver can be influenced by the environmental changes. There are researches using simulations to understand the influence of drunk-driving in driver's behavior. Brazil is far behind the usage of simulation to support road safety studies, when compared to north hemisphere countries.

Studies that use new applications for methodologies consolidated in other related areas can direct researchers to new horizons. Having this in mind, the current

work proposes the study of factors considered as risk factors for occurrence of traffic accidents, aided by designed experiments (Design of Experiments – DOE). Besides the original application in quality issues, DOEs have been used successfully to support engineers in product development (Bayle *et al.*, 2001). Tuning up an experiment is a task that demands a well-defined objective and measurable response variables. Researchers are then able to quantify categorical variables and to measure their effects on response variable.

One way to study risk factors and its influences in road traffic accidents is using virtual simulations. Once a controlled environment is achieved, data collection process is completely reliable as compared to the field data. Another advantage in using simulations is the capacity to record data of all simulated events, allowing researchers to access those anytime.

The ability to replicate a simulation as many time as needed, allows researchers to better understand the impact of each factor considered in the event. For instance, it is possible to have exactly the same traffic situation, climatic conditions and road conditions for several types of drivers in a virtual environment. Even it is possible to have a comparison of the behavior of the same driver, which can help to understand how driver's behavior can be influenced by environmental changes. There are researches using simulations to understand the influence of drunk-driving in driver's behavior. Brazil is far behind the usage of simulation to support road safety studies, when compared to north hemisphere countries.

Studies that use new applications for methodologies consolidated in other related areas can direct researchers to new horizons. Having this in mind, the current work proposes the study of factors considered as risk factors for occurrence of traffic accidents, aided by designed experiments (Design of Experiments – DOE). Besides the original application in quality issues, DOEs have been used successfully to support engineers in product development (Bayle *et al.*, 2001). Tuning up an experiment is a task that demands a well-defined objective and measurable response variables. Researchers are then able to quantify categorical variables and to measure their effects on response variable.

1.2 Research hypotheses

In this study two hypotheses were tested on the use of simulators to study risk factors of the occurrence of traffic accidents, as follows:

- a.) Virtual environment can reproduce same results as observed in field regard risk factors for occurrence of road traffic accidents;
- b.) Virtual driver can replace volunteers in experiments where driver's behavior is not the focus of the study, but need to account driver's skill.

The first hypothesis concerns about the validity of the vehicle model and others parameters considered in the study. It also concerns about a correct selection of software to avoid simplifications that can interfere on vehicle dynamics response, and so, in results. The second one makes reference to the capacity of having a mathematical formulation that can model the different skills among drivers (novice, pilot, standard).

1.3 Research objectives

The current study makes an assessment of risk factors for occurrence of traffic accidents, based on real taken from a Brazilian Federal Highway. The study is aided by the use of simulation and designed experiments. The main objective is to identify active factors for the occurrence or non-occurrence of accidents. The scope is restricted to a compact vehicle and a specific road configuration.

1.3.1 Research secondary objectives

- a) New application for DOE. It is expected that, from this application on, a more structured direction on studies of road traffic accidents;
- b) Investigate possible state variables that can be monitored in a vehicle model, in order to define the imminence of the occurrence of an accident. It would eliminate the subjective analysis presented in lots of studies;
- c) Propose the usage of virtual drivers on road accident studies;

- d) Encourage the use of simulations for road safety studies;

1.4 Research scope

Input data regarding road traffic accidents used as initial information in this work refer to occurrences between Jan/2009 and Dec/2015. They are from a ten kilometer stretch of a Brazilian Federal Highway with high accidentality rates. All data were supplied by the concessionaire that manages the stretch.

Disturbances such like traffic and pavement defects are not considered. Drivers' interaction and drivers' behaviors are not part of the scope of the project. Only dynamic behavior of a compact car is considered. The result might not be the same for others categories.

No volunteers where used. The driving task was performed by virtual driver. VI-CarRealTime® (CRT) has drivers' models with different skills and they were used instead of human drivers. Once the driver and its behavior were not the focus, virtual driver can eliminate noises like driver's distractions, learning and sickness.

2 REVIEW OF LITERATURE

Computational simulations started in automotive industries due to the need of increase profitability. One way to increase profits was by reducing design and tests costs (Allen *et al.*, 2011). Quality improvement and saving time on design products also had good impact on profitability (Raisinghani *et al.*, 2013). The constant need of automotive industry in getting more and more reliable results on simulations was the key to develop software able to reproduce high-fidelity limit maneuvers and handling (Allen *et al.*, 2011).

Naturalistic studies and car accidents are one-time events. Meantime, simulations can be replicated as often as needed. It results in a more precise assessment of driver's behavior (Boyle and Lee, 2010).

While using simulation, there are mainly three mistakes that might happen. The first one is the Model Builder's Risk, when researcher doesn't believe in the model when in fact it is sufficiently representative. The second is the Model User's risk, which is the opposite of first one: believe in results when model is in fact not sufficiently representative. The third one is deal with the wrong problem (Balci, 1994).

With the advanced of driving simulator and the new applications, simulations have been widely used in the last years. Once driving simulators are relatively easy to use, lots of studies focused on evaluation and validation of driving simulators are available in literature (Blana, 1996; Kemeny and Panerai, 2003; Lee *et al.*, 2003; Shechtman, 2010; Underwood *et al.*, 2011; Ronen and Yair, 2013). The main areas that use simulation are: civil engineering, mechanical engineering and human area.

In Mechanical Engineering simulations are widely used to concepts definitions and product design. Racing teams (F1, Stock Car, Rally Dakar) use simulation for suspension and handling tuning and to training their pilots. Volvo, BMW, Ferrari, Porsche are examples of vehicle assemblers that have been using this technology for years. Nowadays, they have shifted for a new level of simulation: the dynamic driving simulators. The dynamic simulators are able to reproduce the same sensations as real driving.

Simulation has wide use for civil Engineers. Stine *et al.* (2010) analyzed the influence of the transversal design of a highway using CarSim®. Horst and Ridder

(2008) analyzed highway infrastructure using a driving simulator. Speed and vehicle lateral position on road was the focus of this study as the key factors for car crashes.

In Healthcare sector, Fisher *et al.* (2002) used a fixed-based driving simulator to compare the risk awareness training on younger drivers. The younger drivers trained on a personal computer had an anticipatory risk awareness then untrained drivers. Chan *et al.* (2010) evaluated how secondary tasks can distract experience and novice drivers while driving, using a driving simulator. The main objective was to analyze the driver's ability of maintenance focus on the main task (driving). The evaluation of drivers' ability was made by monitoring eye track of each volunteer while driving. At the end, Chan *et al.* (2010) could clearly observe different behaviors.

All experiments have in common the way they were prepared and executed. The factors evaluated were introduced and varied in isolation, one factor at a time (OFAT). This approach does not allow the analysis of the interaction among the factors. In that way, researcher loses the capacity of learning and reduces its space of inference. None of studies show clearly what are the researcher's expectation regarding the effects of each factor. Finally, the factors don't have a clearly defined level or range. There are no techniques to filter or deal with noises. It is not possible to say if the noise was controlled, neglected or just ignored.

When dealing with several factors, to conduct a factorial design might be best approach (Montgomery, 2012). A factorial design allows creating and observing significant events instead of wait it to happen. But, at the same time, this kind of experiment has a limited usage due to theorems and statistical proofs that surround them. When well designed, factorial designs can reduce time and cost. It is usually applied on quality issues, and so it is easy to find success stories for such application. Recently, design of experiments (DOE) is being applied for product development, what allows quality excellence to be work during development stage (Bayle *et al.*, 2001; Koch *et al.*, 2004). When dealing with something new, researchers need to understand how factors influence the response of interest. In order to avoid waste of valuable resources by using OFAT approaches, screening experiments can be used (Montgomery, 2012).

The usage of DOE with simulation is an uncommon application. Restrictions as the absence of noises when in virtual environments, demands more caution in data collection, execution and analysis. Edara and Shih (2004) conducted four studies to optimize suspension performance. The first study analyzes rear

suspension bushings that must bear lateral and longitudinal loads. The objective was finding bushing rates that complies vehicle dynamics requirements. To define the best concept among all possible combinations, they used a designed experiment. Next step was another DOE in order to define bushing stiffness. The whole study was made using simulation.

In another example, Bayle *et al.* (2001) needed to understand two patterns of behavior observed in brake designs used to stop rotating parts, within a certain time, in washing machines: the common cause variation (Deming, 2000) where the time to brake change considerably between washing machines, and the special cause variation, where washing machines didn't stop. After three DOE it was possible to solve both problems, with braking times within requirements and a reduction in time variation between machines.

Allen, Rosenthal and Cook (2011) define vehicle dynamics as a vehicle response to an input command or external disturbance. Usually vehicle dynamic is divided into three parts: longitudinal, lateral and vertical. Forces and movements imposed on the vehicle through tire/ground contact, gravity and aerodynamics determine vehicle dynamic behavior. Thus, the right definition of which approach will be used to modeling the system and the hypothesis to describe motions are essentials (Gillespie, 1992).

The Multibody System (MBS) Methodology will be used for the modeling of the vehicle. It is based on classic mechanical. Costa Neto (1992) defines MBS as "a mechanical system with several degrees of freedom". Differential and algebraic equations make the formulation of rigid body movements and the desired constraints and imposes to the system/movement respectively. There are several software that help in the study of vehicle dynamics (MSC - Adams®, Siemens – Virtual.Lab®, Altair – MotionSolve / Hyperworks®, Dassault Systèmes – SimPack®). The choice was the customized software, VI-CarRealTime® (CRT) from VI-grade, to modeling and analyses vehicle dynamics in real time.

3 METHODOLOGY

3.1 The Highway

3.1.1 Geometric design of the Highway

This study is based on a road traffic accident database of a Federal Highway in Brazil. The Highway links São Paulo with Curitiba (southern direction). It is 404 kilometers long; however, accident data collected from only ten kilometers is used for research. Red line as shown in Figure 3 depicts a panoramic view of the stretch of interest. It links the city of Cajati (São Paulo State) and São Paulo-Paraná States border. It is mountainous and has a winding lane. The southern direction is mainly uphill with 3 lanes without shoulder. The maximum speed allowed is 60 km/h for heavy vehicles and 80 km/h for light vehicles. Curves vary from 130m to 625m radius and the maximum grade and acclivity are 6% and 8% respectively (Torres, 2015).

Figure 3 – Panoramic view of baseline Federal Highway.

Source: Google Earth®.

This road is classified as a rural highway Class I-A (DNER, 1999). Considering the speed guideline of 80 km/h and the maximum elevation rate of the curves; the smallest radius of the curve should be 230m. However, four curves have radius smaller than the minimum allowed radius as stated in speed guideline (Figure 4). Curves named as C6 and C14 have the highest rates of traffic accidents. Main parameters of each curve are in APPENDIX B.

Figure 4 – Indication of curves with smaller radius than the allowed radius.

The average daily traffic (ADT) in the years of 2011, 2012 and 2013 was 8845, 9271 and 9233 vehicles respectively in this stretch. The distribution of the type of vehicles is shown in Figure 5. According to the data available, heavy vehicles are the main part of the traffic and the average traffic speed is 69 km/h $\pm$ 13 km/h (Average $\pm$ Standard Deviation), including all type of vehicles (Rangel, 2015).

Figure 5 – Distribution of vehicles for the given stretch.

3.1.2 Road traffic accident data

From January/2009 to December/2015, 862 accidents occurred in this stretch of the Highway (Figure 6). It is observed that the number of accidents is reduced by 80% since 2011, which was the peak year for accidents. This decline is due to the improvement in road signs, speed limit, facilities and awareness campaigns. It is noticed that in 697 occurrences (from the total of 862), there were property damages whereas; 151 times accidents resulted in having victims including 3 fatalities.

Figure 6 – Number of accidents per year in the stretch of the Highway.

Studies reveals that the main probable cause of accidents is driver's behavior: 47% of accidents occur due to the *performance errors*, followed by *speeding* (19%), as shown in Figure 7. *Performance error* is any error caused by driver's ability or misjudgment of his mental/physical condition (drowsy, distraction, etc.). Almost half of the accidents could be avoided, only by driving carefully. It is important to remind that the declared cause of accidents might not be the real one, as it is defined by the investigating officer, responsible for data collection.

Figure 7 – Number of accidents by probable cause in Jan/2009 and Dec/2015.

Three types of vehicles are responsible for more than 90% of the accidents: passenger cars (51%), heavy vehicles (31%) and pick-ups (10%), as shown in Figure 8. The number of accidents for each vehicle type was taken into account for distribution.

Figure 8 – Number of accidents per type of vehicle in Jan/2009 and Dec/2015.

To identify the risky locations, the road span was divided according to its geometric element and then labeled in numerical order. The i -th curve was named as C_i , as also the j -th tangent was named as T_j (Figure 9 and Figure 10).

Figure 9 – Curves identifications.

Figure 10 – Tangents identifications.

APPENDIX C shows the information available in police/concessionaire database, regarding road traffic accidents on this road span. In this work, an assessment has been done for possible variables which could be used as factors by analyzing the database. Henceforth, this research will consider database analysis of passenger car accidents only.

Figure 11 indicates the number of traffic accidents by geometric elements for the time period of Jan/2009 and Dec/2015. Curve C6 has the highest accident rate, and also a small than allowed radius. It is a downhill curve with a 275m long tangent preceding it. Tangents are almost insignificant in terms of accidents, as compared to

curves; however, T11 and T18 are the worst points. T13 is the tangent for curve C14, this segment has the third highest accident rate.

Figure 11 – Accidents distribution per geometric element (Jan/2009 - Dec/2015).

Accident records have a specific field that describes accident's immediate consequence. One third of the occurrences had off track as a consequence. It is observed that there are no serious injuries reported, once this road span has shoulder and it is duplicated too. Figure 12 shows the distribution type of accidents in Jan/2009 and Dec/2015.

Figure 12 – Police records' probable cause distribution (Jan/2009 – Dec/2015).

An analysis of fields available in database selected those that could be considered in a simulation. A bar-chart analysis was used to understand the main factors and levels that contribute to the occurrence of accidents on the stretch of the Highway.

Visibility is the distance at which an object can be seen. It can change according to light and weather conditions. The light at night and day is different, affecting visibility. Hence, occurrences were classified according to *Period of the day*,

referred as day (D) and night (N). The time of the accident was compared to sunset or sunrise time of the same day and then classified. Sunset and sunrise time were determined using information from Astronomical Applications Dept., U. S. Naval Observatory.

If there are more vehicles during day time, it is expected to have more accidents in the same period, as shown in Figure 13. Unfortunately, there were no data available regarding the traffic volumes during day and night. This information would allow making an analysis with respect to the specific traffic volume. It would show that there are more accidents during night as compared to the day time, with regard to the average volume of each period.

Figure 13 – Analysis of occurrences per period: day (D) and night (N).

Figure 14 illustrates the number of accidents by *Track Profile*. There is a contradiction in database on track profile field. For example, kilometer 514 + 900m (curve C14), three different profiles were assigned: uphill in November 2010; in level in August 2011; and downhill in October 2012. Although the stretch of road is uphill, this specific curve is downhill. Hence, there is a disagreement between the records and the real world. In the records, the class is attributed by the officer, what may cause the divergence of information.

Figure 14 – Analysis of occurrences per track profile.

Lane Route accidents distribution is shown in Figure 15. As expected, sharp curve is responsible for 87,5% of occurrences. Curves C6 and C14 have sharp curves and 50% of accidents happen there. Both of them has small radius (130 m) and don't comply with Government recommendations (DNER, 1999).

Figure 15 – Analysis of occurrences per lane route.

Track Condition field indicates whether the track was wet or dry at the time of the accident. It is obvious that most of the time the track remains dry, still the accidents occurred on wet track are three times higher than dry track accidents (Figure 16), due to lower friction coefficient.

Figure 16 – Analysis of occurrences per track condition.

Weather Conditions were originally classified into five categories: normal condition, cloudy, foggy, drizzling and heavy rain conditions. These categories were then grouped, resulting in only two: (i) normal conditions which include normal and cloudy conditions and (ii) changed conditions which include foggy, drizzling and rainy situations. Conditions that change track status somehow, were put together. Analysis of occurrences due to weather change is shown in Figure 17. The difference between *Track Condition* and *Weather Condition* is that the last one may include visibility changes (fog and rain conditions).

Figure 17 – Analysis of occurrences per weather condition.

The last analysis was *Visibility Condition*. This field is the most complex one, because depends on victim's testimony and/or officer's analysis. Neither good nor new information could be revealed with visibility condition datum (Figure 18).

Figure 18 – Analysis of occurrences per visibility condition.

3.2 Basics of vehicle dynamics

Vehicle dynamics is the study related to the movement of the vehicle in response to loads (forces and movements) in result of driver's commands and the environment (Costa Neto, 2005). The vehicle movements are defined with reference to a fixed orthogonal coordinate system of the vehicle using the right-hand rule originating from the center of gravity and that travels with the vehicle (Figure 19).

Figure 19 – SAE vehicle axis system.

Source: Gillespie (1992).

The study of the dynamics of vehicles is divided into three main areas: Longitudinal, Lateral and Vertical dynamics. The first area is associated with longitudinal movements and rotations around its lateral axis in response to torques applied on its wheels; the second is related to lateral translations and angular

velocities around vehicle's longitudinal and vertical axis in response to steering wheel inputs; and the third deals with translations on vehicle's vertical axis and angular velocities around its longitudinal and lateral axis due to ground irregularities (Costa Neto, 2005).

Vehicle motion is usually described by the velocities (longitudinal, lateral, vertical, roll, pitch and yaw) with respect to the vehicle's fixed coordinate system (local reference system that travels with vehicle and has origin in vehicle's center of gravity), where the velocities are referenced to the earth fixed coordinate system (Gillespie, 1992).

The vertical behavior of the vehicle is defined basically by the suspension type and tuning. Vertical (z-axis) and wheel/suspension displacements, pitch and roll angles are quantities used to analyze vehicle's stability with regard to vertical dynamics. The lateral behavior of the vehicle is defined by suspension and steering geometry subsystem. Lateral displacements, yaw and roll are outputs used to analyze vehicle's stability with regard to lateral dynamics. Roll angle and yaw rate analysis are essential for automotive safety, as they are associated to the stability of the vehicle moving in a curve trajectory. (Gillespie, 1992; Costa Neto, 2005).

3.2.1 Roll angle

Roll angle is the angle between the vehicle's lateral axis (y-axis) and the ground plane (rotation around x-axis) and indicates when a vehicle is tilting to left or right in a turn. In other words, it can indicate when a vehicle rollovers. It can be easily measured, and is largely used by engineers to analyze vehicle safety and stability during virtual development and/or suspension tuning phase. This state variable is useful to indicate if there is an unstable situation and it can be used to classify the severity of the occurrence. The last application is not part of the scope of the study and will not be discussed.

Gillespie (1992) defines rollover as "any maneuver in which the vehicle rotates 90 degrees or more about its longitudinal axis such that the body makes contact with the ground". Roll angle can help identify rollover tendency in a maneuver.

Figure 20 shows examples of different roll angles for the same situation: a vehicle moving on a tangent followed by a sharp curve (constant radius), at constant speed. Curve #1 is the vehicle at normal condition of operation. Curve #2 represents

undesirable conditions of roll angle. The first oscillation observed after 10 seconds (small circle) is a hop, where vehicle is “tilting” fast and indicates a possible loss of control. The second condition is the sudden variation of roll angle observed after 13 seconds, which means that vehicle did rollover.

Figure 20 – Examples of different roll angles.

3.2.2 Yaw rate

Yaw rate is the vehicle’s angular velocity around its vertical axis (z-axis) and indicates how fast the vehicle is spinning. Yaw rate (and yaw angle) is related to the controllability of the vehicle. It is used for handling analysis in vehicle’s development. It shows an unstable situation and can be used to classify the severity of the events.

An example of different yaw rate is shown in Figure 21, for a vehicle moving on a tangent followed by a sharp curve (constant radius), at constant speed. Curve #1 is the normal condition and Curve #2 represents an undesirable condition. The small oscillation after 11 seconds (small circle) indicates that the driver is losing control (handling) of the vehicle. The sudden variation of yaw rate after 13 seconds indicates that the vehicle is no longer in control and that it is spinning around its vertical axis.

Figure 21 – Examples of different yaw rate.

3.3 Multibody system methodology

3.3.1 Multibody system modeling

A system is an amount of parts or components in an imaginary frontier conveniently chosen by an analyst. In engineering, the term dynamic is related to the time. In dynamic, time functions variables are studied (Felício, 2007). Multibody systems (MBS) are interconnected rigid or flexible mechanical systems composed of parts with large rotational and translational displacement with each other. The parts are connected by force elements (such as spring dampers) and by kinematic constraints (such as joints), within constraints' conditions. MBS can be represented by a datamodel as shown in Figure 22 (Costa Neto, 2015).

Figure 22 – Datamodel representing MBS.

Source: Adapted from Costa Neto (1992).

According to Costa Neto (2005), some important concepts in MBS are as follows:

- Body: part (flexible or rigid) of a mechanical system;
- Vectors: used to define movements of points and bodies. They have magnitude and direction;
- Referential system: defines a foundation for calculation of magnitudes of motion of a mechanical system. They can be: global or inertial referential and local referential;

- d.) Positioning and orientation methods: for positioning, Cartesian coordinates can be used whereas, for orientation, orientation angles or Euler parameters can be used;
- e.) Link: connection of bodies or body-ground;
- f.) Degree of Freedom (DOF): It indicates how the mechanical system can move. It depends on constraints (type of joints and its alignments). To calculate the number of DOFs of a system, Grueblers equation is used:

$$\text{DOF} = 6 * (\text{moving\_parts}) - (\text{degrees\_of\_constraints})$$

- g.) Inertial properties: each rigid body must have: mass and center of mass location; moments and products of inertia defined in relation to an established reference; or moments and principle directions of inertia.

Dynamic behavior of the system is described using the equations of motions, derived from Newton-Euler and Lagrange's equations. Newton's second law is the basis for constraints' equations.

3.3.2 Vehicle model

According to Felício (2007), modeling is the mathematical equation process and model is the set of equations. In this research, vehicle was modeled using VI-CarRealTime®.

Figure 23 – Rigid parts of vehicle model.

Source: Adapted from VI-Grade (2015a).

Current work uses a simplified mathematical model of a four-wheeled vehicle, with five rigid parts: one sprung mass and four unsprung masses as presented in Figure 23. Vehicle model is divided in eight subsystems: front suspension, rear suspension, steering, powertrain, front wheel and tires; rear wheel and tires; brakes and auxiliary subsystem; and body. Besides the rigid parts, there are no extra parts. Suspension and steering are described in tables with their properties (component data, kinematics and compliance) as shown in Figure 24. Brakes and powertrain are described by algebraic or differential equations. This vehicle model predicts longitudinal, vertical and lateral dynamic behavior accurately (VI-Grade, 2015a).

Figure 24 – Example of suspension modeling.

Source: VI-Grade (2015a).

Tire properties are determinants for vehicle dynamic behavior. There are several tire mathematical models, each one for a specific purpose. They can be divided as per approach adopted to develop tire model, examples are: experimental data only, similarity method, simple physical model and complex physical model. A semi-empirical formula called *Magic Formula Tire Model* (MF-Tire) is widely used to calculate steady-state tire characteristics for vehicle dynamics purpose. The general form of MF-Tire is (Pacejka, 2005):

$$y = D \sin[C \arctan \{Bx - E(B - \arctan(Bx))\}]$$

where coefficients are described in Table 1 and Figure 25.

To produce curves with similar characteristics from measured curves, the Magic Formula $y(x)$ needs a horizontal (S_H) and a vertical (S_V) shift, offsetting the original curve with respect to the origin (Figure 25) and arising a new set of coordinates $Y(X)$ (Pacejka, 2005), where:

$$Y(X) = y(x) + S_V$$

$$x = X + S_H$$

Table 1 – MF-Tire coefficients calculation.

Description	Identification	Equation
Stiffness factor	B	$B = \frac{C_{F\alpha}}{CD}$
Shape factor	C	Estimated or determined by regression techniques
Peak value	D	$D = \mu F_z$
Curvature factor	E	Estimated or determined by regression techniques
Cornering stiffness	$C_{F\alpha}$	$C_{F\alpha} = c_1 \sin \left\{ 2 \arctan \left(\frac{F_z}{c_2} \right) \right\}$
Parameters	c_1 c_2	Estimated or determined by regression techniques.
Friction coefficient	μ	Estimated or determined by regression techniques.
Vertical load	F_z	--

Source: Adapted from Pacejka (2005).

Figure 25 – Curve and indication of MF-Tire parameters.

Source: Pacejka (2005).

Table 2 shows main technical specifications of vehicle model used.

Table 2 – C3 model technical specifications.

Description	Value
Length	3941 mm.
Width	1728 mm.
Height	1538 mm.
Wheelbase	2460 mm.
Front track	1465 mm.
Rear track	1470 mm.
Ixx	400000000 kg-mm ²
Iyy	1400000000 kg-mm ²
Izz	1700000000 kg-mm ²
Kerb weight	1048 kg
Max. weight	1500 kg
Seats	5
Position of engine	Front, transversely
Engine displacement	999 cm ³
Steering type	Steering rack, with electric steering (power steering)
Drive wheel	Front wheel drive
Suspension	Front: Independent, Spring McPherson, with stabilizer. Rear: Semi-independent, spring, with stabilizer.
Tire size	195/60 R15

Source: PSA Peugeot Citroën.

3.3.3 Virtual driver

Bicycle model as shown in Figure 26, is used as foundation for a model based predictive controller. It captures the dynamic effects needed to an efficient and simple driver model (VI-Grade, 2015b). This model represents the two front car wheels by only one. Rear wheels have same assumption (Gillespie, 1992).

Figure 26 – Bicycle model.

Source: Adapted from Gillespie (1992).

For cornering, VI-Driver model calculates the required action at each moment of time. The differential flatness principle is used to define a connecting contour based on the target curve. Vehicle speed (V), slip angle (α), preview time (PT , a user defined variable) and preview distance (D , computed as $V \cdot PT$) are used to compute the required steering angle (δ) to compensate trajectory errors (VI-Grade, 2015b).

Torques needed to longitudinal dynamics computation (brake or accelerate) are based on a feedforward/feedback scheme. Target speed profile from stationary prediction is also used to compute torque. A controller based on vehicle motion acts on throttle, braking and steering to reduce tracking errors to acceptable limits. Lateral and longitudinal controller algorithms are separate; although one has influence on the dynamic behavior of the other (VI-Grade, 2015b).

Preview time (PT) is used to calculate connecting contour. By changing PT control action stability will also change. When PT is increased, vehicle trajectory presents a bigger corner cutting (Figure 27).

Figure 27 – Effect of preview time in vehicle trajectory.

Source: Adapted from VI-Grade (2015b)

This research uses Human Driver model to drive vehicle in a real driving comparable way. It collects information from the vehicle through perception layer. It defines control action from logical layer and operate vehicle from actuation layer. Figure 28 shows internal structure of the human driver cognitive model. The limitations of this model are: inputs come only from vehicle model; as well as tactical and strategic levels are not implemented in this model. There are four types of driving skills available in this model: novice (inexperienced driver), standard (normal driver), professional (pilot) and robotic (no human skill) (VI-Grade, 2015b).

Figure 28 – Human driver cognitive model.

Source: VI-Grade (2015b).

3.3.4 Road models and events

Road was modeled using VI-Road v17.0®. It is similar to a CAD model, but with some particular parameters, such as friction definition. Each road file has its specific configuration and path definition, with all road parameters (x, y and z coordinate, friction, irregularities, and others). The path file is the desirable driver path, related to a road. An example of road used is show in Figure 29.

Figure 29 – Road and path example: front iso (top) and lateral views (bottom).

3.4 Factorial design: basic definitions and principles

Statistical methods are used to analyze collected data. Graphical methods, confidence interval estimation, empirical models and residual analysis are important tools in data analysis and interpretation (Montgomery, 2012).

According to Antony et al² (2003, apud Kwak e Anbari, 2006, p. 709) the fundamental principle of 6 σ is to take an organization to an improved level of sigma capability through the rigorous application of statistical tools and techniques (Kwak and Anbari, 2006). The same principle can be applied on researches, with regard to saving time, money and improving knowledge, instead of enhancing capability.

In their book, Box, Hunter and Hunter (2005) mentioned that “[...] statistical methods and particularly experimental designing, catalyzes scientific method greatly

² ANTONY, J.; ESCAMILLA, J.; CAINE, P. (2003) *Lean sigma*. Manufacturing Engineer, v. 82, n. 2, p. 40-42.

and increases the research efficiency”. Although, most of the experiments present in literature are not designed, the most common is to find studies that analyze one factor at time. In this case, it can only acquire the effect of a single factor at the defined condition of other factors. The combined influence of factors is not analyzed, even having perception that some interactions are important without a doubt. Moreover, Box *et al.* (2005) declare that “it is very important to know which variables do what to which responses”.

It is well recognized that the most common experiments are the ones with factorial designs of two levels. This type of design fits into a sequential strategy (the DMAIC – define, measure, analyze, improve, control) which is an essential feature of scientific method (Box *et al.*, 2005). Montgomery (2012), when referring to factorial design, states that it “means that in each complete trial or replicate of the experiment all possible combinations of the levels of the factors are investigated”. Some advantages in the usage of factorial design are as following: it requires few runs; arithmetic, common sense and graphic analysis are the tools needed for the interpretation of observations; when using quantitative variables, it is possible to get a good direction for further experiments and; finally the design can be fractioned while looking for the most important variables in a large number of factor (Box *et al.*, 2005).

Usually, for a two-level factorial design, it is used coded designs variables, “+” and “–”, instead of the original units of the design factors. It is used to make easier results’ interpretation. This practice is preferable in most all situations (Montgomery, 2012).

The four basic steps to have a successful factorial experiment are:

- a.) Define factors;
- b.) Choose levels;
- c.) Choose design;
- d.) Choose number of runs.

Studies in early stages may need to identify among many factors, those with large effects on the response. Experiments with screening purposes use fractional factorial design. On this design, only a fraction of the factorial experiment is run, and yet, information on the main effects and lower-order interaction are obtained. Three key ideas are the basis for fractional factorial designs (Montgomery, 2012):

- a.) The sparsity principle: main effects and lower-order interactions are more likely to have large effect on response (Montgomery, 2012);
- b.) The projection property: a subdesign can be obtained by deleting complementary set of factors (Cheng, 2006);
- c.) Sequential experimentation: it is possible to combine factors and interactions in a fractional factorial design in order to estimate factors effects (Montgomery, 2012).

The notation used to denote a fractional factorial design is:

$$2_R^{k-p}$$

what means a two level experiment with k -factors, with p -fractions using 2^{k-p} runs and of resolution R . Or in other words, it is a $\left(\frac{1}{2}\right)^p$ fraction of a 2^k design (Box *et al.*, 2005). The resolution is a criteria used to select the best design. Montgomery (2012) defines resolution as: "A design is of resolution R if no p -factor effect is aliased with another effect containing less than $R-p$ factors". It also indicates the amount of interactions that the design is able to estimate. A Roman numeral subscript notation is used to describe the resolution of a fractional factorial design (Box *et al.*, 2005). The higher resolution, the less confounding within factors, and so, more main effects to be analyzed.

The most convenient form to identify the generating relation is use of the identifier I. When employing fractional factorial designs, the generating relation for the sign must be carefully selected in order to avoid confounding significant effects (whether they are main or interaction effects). "The effect of a factor is defined to be the change in response produced by a change in the level of the factor" (Montgomery, 2012). The main effect is the primary factor effect and can be determined, for each one, as under:

$$\text{Main effect} = \bar{y}_+ - \bar{y}_-$$

where, $\bar{y}_+$ and $\bar{y}_-$ are the average response of each factor to the plus and minus level. The interaction effect, similarly, is the effect of the interactions, and can be calculated like the main effect (Box *et al.*, 2005).

Considering an experiment with four factors (**A**, **B**, **C** and **D**), designed as half fraction and that have the generation relation for the design **D = ABC**. The notation used to describe the designed experiment is 2^{4-1} , and the identity is **I = ABDC**.

The resolution of the example is determined by the number of letters contained in **I**. It also means that, in the given example, there are main effects aliased with three-factor interactions and two-factor interactions aliased to one another (Box *et al.*, 2005): 2_{IV}^{4-1} .

To spread the effect of nuisance variables across the design, randomization technique is commonly used (Box *et al.*, 2005). Furthermore, Box *et al.* (2005) state that “to obtain an estimate of error that can be used to estimate the standard error of a particular effect, each experimental run must be genuinely replicated”. A complete random design is the best strategy to spread effect of nuisance and to estimate the standard error. The standard error of each effect must be taken into account to determine which effects have more chance to be real, considering that each effect is probably due to chance variation. If the effect is 2 or 3 times its standard error, then it is probably a real effect.

To interpret DOE results, following schemes and charts can be used: plots of main effects for mean and for standard deviation; Pareto chart of means for main factor effects and higher order interactions; or Pareto chart on the standard deviation of factors and interactions. Box *et al.* (2005) also provide with a list of designs for two-level fractional factorial (APPENDIX F).

Regression models are useful representations to estimate responses, based on experiment's results, when one or more factors are quantitative. Considering a two-factorial design, the regression model could be as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$$

where, β_0 is the average of all responses, β_1 and β_2 are the one half of the estimated main effects, β_{12} is one half of the estimated interaction effect and x 's are variables that represent factors and its interactions, defined on the coded scale from -1 to +1 (Montgomery, 2012).

3.5 Experiment outline

3.5.1 Design of the experiment #01 – Quarter fractional factorial screening design

In order to save time, due to the number of variables and also to give the direction, a factorial design with two levels was used for screening purpose. Henceforth, variables will be called as factors. The factors and levels were defined on the basis of road traffic accident database and literature and are shown in Table 3.

Table 3 – DOE #01: factors and levels.

FACTORS	LEVELS	
	(-)	(+)
A. Curve radius	130 m	230 m
B. Path profile	Downhill	Uphill
C. Path conditions	Wet	Dry
	(0,3 ~ 0,5)	(0,7 ~ 0,9)
D. Driver' skill	Novice	Standard
E. Speed	Low	High
	(50 km/h ~ 70 km/h)	(110 km/h ~ 130 km/h)
F. Period of the day	Night	Day
G. Car load	1 person	4 person
	(70 kg)	(280 kg)

Curve radius (A) levels were defined considering the lower level at 130 m (the small radius in the baseline stretch of the Highway), and the highest level at 230 m (the minimum curve radius allowed according to DNER (1999)).

Path profile (B) was defined according to curve C14, which has up and downhill slope of 6%. In addition, *Path conditions (C)* was modeled by changing friction coefficient, based on literature. In order to have variations, friction coefficient vary from 0.3 ~ 0.5 for a wet pavement and from 0.7 ~ 0.9 for a dry pavement (Canale, 1993; Hall et al., 2009).

Speed (E) levels were defined based on the allowed and measured speeds on the stretch of the Highway (Rangel, 2015; Torres, 2015). Lower level can vary from 50 km/h ~ 70 km/h and higher level can vary from 110 km/h ~ 130 km/h. Lateral acceleration variation between extremes for each level is shown in Table 4. Lateral acceleration may influence the driver's behavior and must have similar values between levels. Although acceleration variation is higher at level (+), it was considered satisfactorily close.

Table 4 – Lateral acceleration variation of speeds at the same level.

Level	V ₁ (km/h)	V ₂ (km/h)	Ratio (a ₂ /a ₁)
(-)	50	70	1,96
(+)	110	130	1,40

Speed and friction coefficient were used as source of variation of the experiment. A value within the factor level's range was set up randomly according to their level for each treatment.

Driver' skill was defined with VI-Driver® and *Car load* was defined according to the car specification, obtained with the manufacturer (PSA Peugeot Citröen). Each passenger weight 70 kg, without baggage.

Period of the day was modeled considering how far the driver can see at night and in daylight; and how much the reaction time is available for each period. CRT has the preview time parameter, which is "how far the controller looks ahead (..) for planning control action" (VI-Grade, 2015a). By increasing PT, driver will have a lower response on steering and a reduced accuracy of path tracking (VI-Grade, 2015a). By default the minimum allowed for PT parameter is 1,0s. It was defined that at highest level it would be 1,0s and at lowest level it would be 2,5s.

DOE planning form is presented in APPENDIX D (Moen *et al.*, 2012). A 2^{7-2} factorial design was used. The design was defined based on its resolution: IV. This resolution allows main effects to be completely clear from each other and from second order interactions. Third and higher order interaction was assumed to be negligible. The design specification and its confounding matrix are shown in Table 5. Generators and defining relations were determined as per Box *et al.* (2005) recommendations.

Table 5 – Specifications of the 2_{IV}^{7-2} fractional factorial design.

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

Source: Adapted from Box *et al.* (2005, p.273).

Treatments were replicated three times, and a complete random design was used avoid systematic effects. The factor relationship diagram (FRD) is available in APPENDIX E.

Variable response Y1 is binary and simply indicates which event occurred an accident and which not. For a first screening, it will give enough information. “0” indicates no occurrence of accident and “1” indicates accident.

A car was considered in an unsafe condition, here denoted as accident, when the wheel outside of the curve is no longer in the path, that is, when lateral displacement of one meter is observed (Figure 29).

Figure 30 – Definition of an unsafe condition.

3.5.2 Design of the experiment #02 – Half fractional factorial design

The second DOE was planned based on the results observed on DEO #01. Factors and levels are displayed in Table 6. The main objective was to quantify the factors' effect on road traffic accident.

Table 6 – DOE #01: factors and levels.

FACTORS	LEVELS	
	(-)	(+)
A. Curve radius	130 m	230 m
B. Car load	1 person (70 kg)	4 person (280 kg)
C. Driver' skill	Novice	Standard
D. Period of the day	Night	Day
E. Speed	Low (60 km/h ~ 80 km/h)	High (90 km/h ~ 120 km/h)

Curve radius (A), *Car load (B)*, *Driver' skill (C)* and *Period of the day (D)* levels are the same as described in DOE #01 (item 3.4.2). A downhill profile, with slope of 6% and friction coefficient of 0,8 (dry pavement) was used in all runs. These two factors were withdrawn from the second DOE once they show no influence on variable response in DOE #01. *Speed (E)* levels were adjusted to represent a driving within speed limits (lowest level) and the upper tail of the measured speed distribution (Rangel, 2015). *Speed (E)* was set randomly among runs for each treatment to simulate variation.

Table 6 shows the response variables chosen to measure and monitor the DOE. Variable Y_1 is the same used on DOE #01. Variable Y_2 is the linear measurement of the distance travelled until the end of simulation. It allows differentiating the severity of each event, when accidents are observed.

Table 7 – 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 travelled until accident (continuous)

The experiment was replicated 14 times for each treatment. For each replication of the same treatment a different value of *Speed (E)* was randomly assigned, respecting the interval defined by the level. The experiment was completely randomized.

DOE planning form is presented in APPENDIX G (Moen *et al.*, 2012). A 2_V^{5-1} fractional factorial design was used. The experiment is of resolution V, allowing main effects to be completely clear from each other and from second order interactions. Second order interactions are also completely clear from each other. The design specification and its confounding matrix are shown in Table 8. The FRD is available in APPENDIX H.

Lateral displacement was used to indicate an unsafe condition, just like defined on DOE #01.

Table 8 – Specification of the 2_V^{5-1} fractional factorial design.

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

Source: Adapted from Box *et al.* (2005, p.273).

4 RESULTS

Montgomery (2012) presented a general approach to the statistical analysis of the 2^k design. The analysis procedure is summarized in Table 8. Jump v.13® and Microsoft Excel 2010® were employed to analyze the results.

Results will be presented by the response variable, as described on Table 6

Table 9 – Analysis procedure for a 2^k design.

Analysis procedure

---

1. Estimate factor effects;
  2. Form initial model:
    - a. If the design is replicated, fit the full model;
    - b. If there is no replication, form the model using a normal probability plot of the effects;
  3. Perform statistical testing;
  4. Refine model;
  5. Analyze residuals;
  6. Interpret results.
-

Source: Adapted from Box *et al.* (2005).

4.1 DOE #01

4.1.1 Y1: Frequency of accidents

In the total of events (96 events), no accidents were observed in 54 cases and in 42 events there were accidents.

The *multi vari chart* is a graphical representation of factors relationship with regards to one response variable. This analysis is used for measurement system evaluation and examines reproducibility and repeatability. They can help analyze interaction (SAS Institute Inc., 2017). Figure 31 shows the chart for Y₁. A systematic effect can be seen in factor **E** (speed): in every event with **E** (-1) no occurrence is

observed and all events with **E** (+1) there are occurrences. Three treatments had change of behavior among passages: T20, T22 and T26. It can be explained by the variation in factors **C** (friction coefficient) and **E** (speed).

Contrasts for main effects and second order interactions (Table 10) were estimated. Microsoft Excel® was employed to the estimations. Y_1 was substitute by the mean response y_1 , which is the probability that $Y_1 = 1$ for the given factors' levels.

Factors and interactions with contrasts equal or higher than 0,05 (5%) were considered significance for y_1 . All three-factor or higher order interactions were considered negligible. **A** (curve radius), **D** (driver' skill), **E** (speed), **F** (period of the day) and **G** (car load) were important factors. **E** (speed) had the larger effect on response. It effects was seven times greater than others. Factors **B** (path condition) and **C** (friction coefficient) have too small influence on y_1 , and can be discarded from further experiments.

Prediction values (y) for all 128 combinations of factors and the confidence interval for each combination are presented in APPENDIX F. Negative predicted values are zero.

Table 10 – DOE #01: main effects and second order interactions contrasts 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

Figure 31 – DOE #01 Multi Vari Chart for Y_1 .

4.2 DOE #02

4.2.1 Y_1 : Frequency of accidents

This variable is the same as described previously on DOE #01. In 224 runs, accidents were observed in 96 ($Y_1 = "1"$) and in 128 runs no accident were observed ($Y_1 = "0"$). There were six treatments with change of behavior among passages. It indicates that the level used had a more proper adjustment compared to DOE #01, especially factor **E** (speed). Figure 32 shows the multi vari chart. It is not possible to determine by graphical analysis if there is any main effect or second order interaction with large effect on response. It might indicate a more balanced experiment in terms of level choice.

Y_1 was substituted by mean response y_1 and contrasts were calculated (Table 11). Factors and interactions with contrast equal or higher than 0,05 (5%) were considered as relevant. Three factor and higher interactions were considered negligible. **A** (curve radius), **B** (car load), **C** (driver skill), **D** (period of the day), **E** (speed), **AB**, **BC**, and **DE** have the large effects.

The linear model using factors **A**, **B**, **C**, **D**, **E** and the interactions **AB**, **BC**, and **DE**, is:

$$y = 0,4285 + (-0,1428) \cdot \mathbf{A} + 0,0267 \cdot \mathbf{B} + (-0,0714) \cdot \mathbf{C} + (-0,2232) \cdot \mathbf{D} + 0,2678 \cdot \mathbf{E} + (-0,0625) \cdot \mathbf{AB} + (-0,1339) \cdot \mathbf{BC} + (-0,0804) \cdot \mathbf{DE}$$

where 0,4285 is y_1 overall mean, y is the predicted value, factors' coefficients are the calculated contrasts and factors **A**, **B**, **C**, **D**, **E**, **AB**, **BC**, and **DE** shall be replaced by the code (+1) and (-1) as per treatment combination.

Prediction values (y) for the 32 combinations of factors and the confidence interval for each combination are presented in APPENDIX I. Negative predicted values are zero.

Figure 32 – DOE #02 Multi Vari Chart for Y₁.

Table 11 – DOE #02 Main effects and second order interactions contrasts for y_1 .

CONTRAST SUMMARY			
Term	Contrast	Term	Contrast
E	0,2679	AE	-0,0357
D	-0,2232	AC	0,0357
A	-0,1429	B	0,0268
BC	-0,1339	BE	0,0268
DE	-0,0804	BD	0,0179
C	-0,0714	CD	-0,0089
AB	-0,0625	CE	0,0000
AD	-0,0446		

4.2.2 Y_2 : Path distance

Y_2 is a numerical continuous variable that indicates the severity of the event by measuring the travel distance until the accident. The sooner the accident occurs, the more severe the event will be.

Figure 33 shows the worst and the better run for each treatment of lateral path deviation *versus* travelled distance for events with accidents. The dotted horizontal line indicates the maximum lateral displacement for a safe condition. After crossing that line, data were discarded. Solid lines are the shortest travelled distance of each treatment and dotted lines are the longest ones. The color scale indicates different treatments. All curves are shown in APPENDIX J.

The Multi Vari Chart (Figure 34) shows the travelled distance. For events where lateral displacement was equal to 1,0m, the travelled distance where considered as the point of the stretch where the accident occurred. Red line indicates where the curve begins and yellow line indicates its end. 86 of 96 events with accident occurred on the curve. In the others 10 events the accident occurred until 25 m after the end of the curve. Treatments T1, T6, T7 and T13 had the lowest travelled distance, and so, are the treatments with higher severity. The lower the distance, the sooner the accident and so, more dangerous is the combination of factors. All four treatments have *Speed (E)* at the highest level.

Figure 33 – Lateral path deviation *versus* travelled distance for runs with $Y_1 = 1$.

Table 12 shows calculated contrasts for Y₂. Contrasts equal or higher to 10 were considered as relevant. Factors E (speed), D (period), interaction ED, factor A (curve radius), interaction CB and factor C (driver skill), in this order, were the higher contrasts.

Table 12 – Calculated contrasts of Y₂.

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 prediction expression for Y_2 , calculated using linear regression method is:

$$y_2 = \begin{cases} 430,00 - 51,26\mathbf{E} + 42,29\mathbf{D} + 23,81\mathbf{A} - 13,88 - 9,67\mathbf{B} + 25,33\mathbf{ED} - 15,00\mathbf{BC}, & \text{for } \mathbf{C}=-1 \\ 430,00 - 51,26\mathbf{E} + 42,29\mathbf{D} + 23,81\mathbf{A} + 13,88 - 9,67\mathbf{B} + 25,33\mathbf{ED} + 15,00\mathbf{BC}, & \text{for } \mathbf{C}=+1 \end{cases}$$

where 430,00 is Y_2 overall mean, y_2 is the predicted value, and factors **A**, **B**, **D**, **E**, **BC** and **ED** shall be replaced by the code (+1) and (-1) as per treatment combination. Factor **C** is categorical and for each code of C a different expression must be used.

4.3 Vehicle state variables

Variables analyzed to monitor vehicle's control and handling stability in a steady-state cornering includes sideslip, roll and pitch angles, yaw rate, slip angle of the wheels and lateral velocity. Roll angle and yaw rate were chosen to be monitored because they are easy to understand and they are correlated to the others variables. This chapter will discuss roll angle and yaw rate. Graphics are fully presented in APPENDIX K and APPENDIX L, respectively. Usually these graphics are plotted in time. Since events were performed at different speeds, resulting in different times to travel the same route, they were plotted in vehicle travelled distance.

4.3.1 Roll angle

Figure 35 has a summary of roll angles behavior observed in DOE #2. Red line (1) is an event with no occurrence. The other curves represent different accident situations. In these cases there is a hop oscillation before the accident. The oscillation has different frequencies and amplitudes, pointing the severity of the event. Blue line (2) is an event where the driver was able to do the curve, but with wheels outside the curve (front and rear) off track, that is, with a greater radius. Green line (3) is a similar situation, but with wheels inside and outside the curve off track. Pink curve (4) is a rollover but keeping contact tire/ground with wheels outside the curve. Finally, the orange line (5) is a completely rollover.

Figure 35 – Summary of roll angle behavior observed in DOE #02

Roll angle can also point when wheels outside the curve are out off track. The highlighted area in Figure 36 is a disturbance in the system not observed in others curves. The disturbance is the lower friction coefficient of “shoulder”, resulting in different reaction forces in the tire and consequently a roll angle variation

Figure 36: Roll angle variation due to normal forces variation when off track.

4.4 Yaw rate

Yaw rate is how fast yaw angle is changing in time. The faster the changes more likely is to loose vehicle's control and handling. So, when analyzing yaw rate it is important to look for high frequencies and amplitudes. When yaw rate reaches value zero, it means that the vehicle is no longer spinning, and when in a cornering, it means that the vehicle lost the trajectory.

The analysis of the yaw rate curve is similar to the analysis done for the roll angle. An abrupt and sudden variation in the curve indicates that wheels outside the curve have lost contact with the track. In Figure 37 there are some examples of yaw rate behavior observed in DOE #02 that can summarize yaw rate. Orange line (1) is a typical event, without any occurrence, while red (2), blue (3), green (4) and pink (5) curves are events with accidents. Arrows indicate the moment when wheel from outside the curve went off track.

Figure 37 – Summary of yaw rate behavior observed in DOE #2.

5 DISCUSSION

This section includes discussion about the designed experiments and simulations, one subchapter for response variables, one subchapter for vehicle states variable and a last subchapter with discussion about the virtual driver.

5.1 Designed experiments and simulation

When working with two-level factorial designs, an important and not trivial work is to correctly choose levels. It is suggested that some experiments (most of the time, fractional factorial) be done in order to understand and evaluate levels before the study of factors itself. By performing a Screening Design with a large number of factors it is possible to get a good direction in factors and its levels' effect on response variable. In DOE #01 *Speed (E)* was responsible for almost half of y_1 value. When levels were adjusted in DOE #02 it was possible to better understand how factors and interactions influence the response variable.

Simulations are a complete controlled environment and there are no nuisances among events. To deal with that, *Speed (E)* and *friction coefficient (C)* were used as sources of variations. It made possible to replicate treatments instead of study only a single value of these two factors, increasing space of inference.

Replication and the completely randomized design also allowed study how capable of replicate is a virtual environment. Once *Speed (E)* was randomly defined within each level, there were events with same speed values within same treatment. Table 13 shows events grouped by speed at the same treatment. Each group corresponds to a speed value. Between treatments speeds may change. By analyzing results (roll angle, yaw rate or lateral path deviation), is possible to notice that those passages with same speed has exactly the same behavior. It occurs because simulations are deterministic events and every time a same configuration is performed, it will produce the same result.

Table 13 – Passages from DOE #02 grouped by speed.

T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16
P2	P15	P30		P64	P78	P91	P99	P116	P128	P141	P155	P174	P188	P199	P214
P12	P23	P32	-	P66	P79	P93	P103	P123	P133	P154	P164	P182	P190	P204	P216
P1	P16	P33		P57	P72	P96	P105	P114	P135	P147	P163	P173	P189	P207	P211
P9	P22	P35	-	P63	P77	P98	P106	P121	P139	P153	P165	P181	P194	P210	P218
					P83		P112		P140						
-	P17	P39		P65	P73	P85		P118	P131		P157	P169	P183	P203	P219
	P24	P42	-	P70	P74	P94	-	P119	P132	-	P162	P180	P191	P206	P220
	P28							P122					P196		
-	-	-	-	P59											P217
				P62	-	-	-	-	-	-	-	-	-	-	P222

Figure 38 is an example of two events with the same configuration but ran in separately. Event P 105 was the 10th event performed, while P 112 was the 222nd. Despite of randomization, yaw rate curves of both events were overlapping. It means that when using simulation, the run order is not important. This statement does not apply to experiments with human drivers, as in the case of driving simulators.

Figure 38 – Example of overlapping curves of DOE #02.

DOE #01 was a screening design and results must be carefully used. It can only be used to give a direction for further experiments. By increasing data available for analysis (including more replications) it would also increase the chance of occurrence of accident and possibly we would have a less screening design. To illustrate that, Figure 39 makes a comparison between the fitted models for both

5.2 Response variable

The transformation of Y_1 into a probability function made possible the estimation of a model to predict the probability of occurrence of accident within the factors considered. It was used linear probability model. Another method that could be used is the logistic regression, where parameters of the response functions are estimated by using the maximum likelihood method. But this method will not be discussed at this study.

The result was not possible with Y_2 , despite of being a continuous variable. Residual and Predicted plots (Figure 41) shows that the model is not well adjusted and results should be used carefully. On residual plots it is expected that points falls randomly on positive and negative side of the blue line and that no pattern can be recognizable in this distribution. The pattern presented indicates a nonconstant variance. It can be a due to the categorical variable, factor C. Predicted plot have a low number of observations within the range of predictor values. More statistical analysis needs to be done before any conclusion about Y_2 . This variable was expected to indicate the severity of each event, what can also be done by monitoring vehicle states variables roll angle and yaw rate.

Figure 41 – DOE #02 - Residual and predicted plot for Y_2 .

5.3 Vehicle state variables

In a cornering, there is lateral load transfer from the inside wheels of the vehicle to the outside wheels (Gillespie, 1992). It means that tires normal forces monitoring can improve the capacity of identify imminence of accident. Figure 42 shows tire normal force plot in different situations.

By monitoring only roll angle is not possible to know where wheel of inside the curve loss the contact with ground. Figure 43 illustrate the difference in using roll angle and normal forces. The red circle indicates the same instant when the vehicle starts to rollover. By analyzing roll angle it not possible to know that the car is already on two wheels. When tire normal force is plotted the point where normal reaches zero is clearly defined. In all analysis normal forces were plotted, although it was not necessary at this case.

A similar situation is observed for yaw rate. Usually yaw rate and sideslip angle are analyzes together when there are lateral forces acting affecting vehicle's motion. But will not be discussed in this research. Yaw rate, unlike roll angle, provide sufficient information to monitor lateral instability.

The influence of speed in the roll angle and yaw rate can be clearly seen in Figure 44. They are all results from treatment 12. Speed can amplify roll angle and

yaw rate behavior. It might be an explanation why speed is considered one of the most significant factors for traffic accident.

Figure 43 – Tire normal force: identification of the point where tires loose contact with ground.

Figure 44 – Influence of speed in roll angle and yaw rate.

Yaw rate and roll angle are measured at local coordinate system. When there is a rollover it is expected that the vehicle also spins at considered rate around its vertical axis. Figure 45 illustrated two events from the same treatment that differs only by speed value. In P 143 there is a rollover while in P 149 there is a slight off track. When the vehicle starts to hop there is a sudden change in yaw rate and as the driver keep driving, there is more load transfer (that is why roll angle continue to increase) and there is a discrete oscillation of yaw rate. These oscillations need to be avoided so driver can have the control of the vehicle.

Figure 45: Roll angle and Yaw Rate relation.

5.4 Virtual driver

A short discussion on the virtual driver will be presented, although this analysis is not part of the scope of the study. Figure 46 has information about steering angle and throttle signals for a same maneuver changing only the driver skill. Notice the different behavior in responsiveness and in terms of oscillation in both signals.

The different behavior resulted in 8 meters difference in path lateral deviation between drivers (Figure 47). At the end, standard driver were capable of controlling the car again while novice driver rollover.

Database used as input doesn't have driver's information, but it is of general knowledge that novice drivers are more likely to have traffic accidents due to their unfamiliarity with driving dynamics.

As speed increases, frequency and amplitude of steering inputs, which came from the driver, have major impact in lateral instability and that is why the driver skill might be an important factor to be taken into account. Unfortunately, when using human driver, it is almost impossible to have a sample with the same behavior (or skill) among all volunteers. Methods used for sample filtration are kilometers driven per week, driver's license time, habitual usual driving place (road, highway, urban area). But none of them is capable to standardize or measure the driver skill.

Figure 46: Driver demands.

Figure 47: Path lateral deviation for different drivers skill.

6 CONCLUSION

This study was aimed on making an assessment of risk factors, responsible for the occurrence of traffic accidents. Based on the traffic accident database from a Brazilian Federal Highway, the study was performed in a virtual environment, using validated vehicle models to reproduce the vehicle dynamic behavior. The study is a part of a big research group focused on the application of simulations for road safety studies.

One of the pronounced contributions of the current study was to be able to identify active those factors that affect road traffic accidents. In the two experiments, speed was the factor with major influence on occurrence of accidents, like pointed WHO in her 2015 Global Report for Road Safety. The first experiment had screening purposes and pointed the direction for the next experiment. Three main effects (speed, curve radius and period of the day) and four second order interactions had influence on the occurrence of accidents. These main effects were considered in the second experiment. Two factors were also part of the second experiment despite of their low significance: driver skill and car load, because of their interaction. For the second experiment speed levels were adjusted for lower ranges. Speed, period of the day, curve radius and driver skill were the factors, in this order, with relevant influence on the occurrence of accidents. For the linear probability model, a fifth main effect (car load) was considered because of its interaction with curve radius and driver skill.

Another important contribution was to design an experiment that allowed evaluating not only the influence of one factor at a time, but also to evaluate the influence of their interactions on the response variable. The use of fractional factorial experiments allowed an increase in the number of factors considered and the number of replicates performed, resulting in a more knowledgeable scientific research.

Although it was not the focus of this research, a third contribution was to introduce the usage of virtual drivers for road safety studies. There is still no mathematical formulation capable of reproducing human behavior with accuracy, and so naturalistic researches are needed. Herein, the virtual driver used was able to

differentiate driver skill. For screening design proposes this tool might be helpful by reducing time and human resources in early stages of research. It can also be used to test experiments before their startup, as a kick-off tool. It would allow researchers to test response variables, levels of factors, geometric design variations, among others possibilities.

Last but not the least contribution, the study proposed the use of two possible vehicle state variables which can help to identify imminence of accidents in a virtual experiment: roll angle and yaw rate. The state variables can also be used in driving simulators. However, it is necessary to have a representative vehicle model, preferably one that contemplates tires and suspension formulation. Otherwise, state variables can be overestimated or even worst, underestimated. Tire normal forces can be used with roll angle when it is important to define road critical points.

In this study two hypotheses were tested to study risk factor for the occurrence of traffic accidents, which are as follows:

- a.) Virtual environment can reproduce same results as can be observed in the field, regarding the risk factor for occurrence of traffic accidents;
- b.) Virtual driver can replace volunteers in experiments where driver behavior is not the focus of the study, but driver' skill needs to be accounted.

The first hypothesis was concluded as it is possible to have similar results in the field and virtual environment. The quality of the results depends on the quality of input data used in the modeling process and the assumptions adopted for modeling.

For the second hypothesis, virtual driver cannot completely replace volunteers in experiments. The study emphasized that virtual driver is a helpful tool for screening experiments when only driver's skill is needed.

It is important to notice that all presented results are valid only for the vehicle model being used. It is expected that similar vehicles, from the same category however, from others companies, have similar performance. Still it cannot be extrapolated without an initial analysis. Some road boundary conditions which may alter results are: slope, bank consideration, curves with radius below 130 m, lane width, shoulders (and its friction coefficient), among others. The preview time is also a factor that deserves to be used very carefully. Values used don't have any validation with real values.

6.1 Recommendation for further studies

In any research, during the study there were uncertainties that could not be solved and hypotheses were taken to simplify situations and/or configurations. Such uncertainties and simplifications can be seen as suggestions for future studies, which are as following for this research:

- Validate results using a driving simulator, with human driver. It would also allow study the difference between human and virtual driver and their skills. The study may also provide information on how each driver interacts with vehicles commands by analyzing driver demands such as steering angle, throttle and brake. An additional contribution would be the beginning of the construction of a database containing information on how Brazilians drive.

- Application of suggested vehicle states variables in a study using driving simulator or simulation to monitor accident. It is also possible to differentiate the non-occurrences by severity scale based on state variables and perform a statistical analysis with data. It would be the beginning of a of the definition of a metric for road safety;

- Replicate the study adding road geometry elements such like: superelevation, grade, horizontal and vertical curves, among others. Vehicle state variables and driver demands can be used to differentiate the way of each human driver behavior and how they interact to each element in road geometry.

- Any study focused in studying only geometric design and highway surrounding in a comparable way (such like previous and after any change) and that does not need the cognitive ability of the driver, can use virtual driver instead of human driver.

REFERENCES

- AASHTO. **Highway Capacity Manual**. Washington, D.C.: TRB, National Research Council 2000.
- ALLEN, R. W.; ROSENTHAL, T. J.; COOK, M. L. **A short history of driving simulation**. Boca Raton, USA: CRC Press, 2011.
- BALCI, O. Validation, verification, and testing techniques throughout the life cycle of a simulation study. **Annals of Operations Research**, v. 53, n. 1-4, p. 121-173, 1994. ISSN 0254-5330.
- BAYLE, P. et al. ILLUSTRATION OF SIX SIGMA ASSISTANCE ON A DESIGN PROJECT. **Quality Engineering**, v. 13, n. 3, p. 341-348, 2001. ISSN 0898-2112.
- BLANA, E. Driving Simulator Validation Studies: A Literature Review., 1996-12 1996. Available at: < <http://eprints.whiterose.ac.uk/2111/> >.
- BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for Experimenters**. 2. ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005. ISBN 0-471-71813-0.
- BOYLE, L. N.; LEE, J. D. Using driving simulators to assess driving safety. **Accident Analysis and Prevention**, v. 42, n. 3, p. 785-787, 2010.
- CANALE, A. C. **Automobilística: dinâmica, desempenho**. 10. ed. São Paulo: Érica, 1993.
- CHAN, E. et al. Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills? , v. 13, n. 5, p. 343–353, September 2010 2010.
- CHENG, C.-S. Projection Properties of Factorial Designs for Factor Screening. In: DEAN, A. e LEWIS, S. (Ed.). **Screening: Methods for Experimentation in Industry, Drug Discovery, and Genetics**. New York, NY: Springer New York, 2006. p.156-168. ISBN 978-0-387-28014-1
- COSTA NETO, A. **Application of multibody system (MBS) techniques to automotive vehicle chassis simulation for motion control studies**. 1992. 213 (Doctor of Philosophy). Engineering Department, University of Warwick, Coventry, UK.
- _____. **Sistemas Multicorpos: curso Dinâmica e controle de sistemas robóticos** 2005.
- _____. **Dinâmica Veicular**. São Carlos, SP: 2015.
- DEMING, W. E. **Out of the Crisis**. Massachusetts Institute of Technology, Center for Advanced Engineering Study, 2000. ISBN 9780262541152.

DNER. **Manual de projeto geométrico de rodovias rurais**. Rio de Janeiro: IPR: 195 p. 1999.

EDARA, R.; SHIH, S. Effective use of multibody dynamics simulation in vehicle suspension system development. **SAE Technical Paper**, 2004.

EVGENIKOS, P. et al. WHO | Data systems: a road safety manual for decision-makers and practitioners. **WHO**, 2010. ISSN 978 92 4 159896 5.

FELÍCIO, L. C. **Modelagem da dinâmica de sistemas e estudo da resposta**. RiMa, 2007. ISBN 8576561182.

FHWA. How Do Weather Events Impact Roads? , 2016. Available at: < http://ops.fhwa.dot.gov/weather/q1_roadimpact.htm >. Accessed on: 27/jan/2016.

FISHER, D. L. et al. Use of a Fixed-Base Driving Simulator to Evaluate the Effects of Experience and PC-Based Risk Awareness Training on Drivers' Decisions. **Human Factors: The Journal of the Human Factors and Ergonomics Society**, v. 44, n. 2, p. 287-302, June 1, 2002 2002.

FRITTELLI, C. et al. Effects of Alzheimer's disease and mild cognitive impairment on driving ability: a controlled clinical study by simulated driving test. **International Journal of Geriatric Psychiatry**, v. 24, n. 3, p. 232-238, 2016. ISSN 1099-1166.

GEORGE, M. L. et al. **The lean six sigma pocket toolbox**. United States of America: McGraw-Hill, 2005. 282.

GILLESPIE, T. D. **Fundamentals of vehicle dynamics**. Warrendale, USA: Society of Automotive Engineers Inc., 1992. 470.

HALL, J. et al. Guide for pavement friction. **Transportation Research Board of the National Academies, Washington DC, USA**, 2009.

HARRY, M.; SCHROEDER, R. **Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations**. New York: Doubleday, 2000. ISBN 0-385-49437-8.

HARVEY, A. et al. **Sistema de dados: um manual de segurança viária para gestores e profissionais da área**. Geneva: World Health Organization 2010.

HORST, R. V. D.; RIDDER, S. Influence of Roadside Infrastructure on Driving Behavior: Driving Simulator Study. <http://dx.doi.org/10.3141/2018-06>, 2008-01-25 2008.

KEMENY, A.; PANERAI, F. Evaluating perception in driving simulation experiments. **Trends in Cognitive Sciences**, v. 7, n. 1, p. 31-37, 2003

KOCH, P. N.; YANG, R.-J.; GU, L. Design for six sigma through robust optimization. **Structural and Multidisciplinary Optimization**, v. 26, n. 3-4, p. 235-248, 2004. Available at: < <http://link.springer.com/article/10.1007/s00158-003-0337-0> >.

KWAK, Y. H.; ANBARI, F. T. Benefits, obstacles, and future of six sigma approach. **Technovation**, v. 26, n. 5, p. 708-715, 2006. ISSN 0166-4972.

LEE, H. C.; CAMERON, D.; LEE, A. H. Assessing the driving performance of older adult drivers: on-road versus simulated driving. **Accident Analysis & Prevention**, v. 35, n. 5, p. 797–803, September 2003 2003.

MAZE, T.; AGARWAI, M.; BURCHETT, G. Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. **Transportation research record: Journal of the transportation research board**, n. 1948, p. 170-176, 2006. ISSN 0361-1981.

MOEN, R. D.; NOLAN, T. W.; PROVOST, L. P. **Quality improvement through planned experimentation** 3. ed. Canadá: MCGRAW-HILL, 2012. ISBN 9780071759670.

MONTGOMERY, D. C. **Design and analysis of experiments**. 8ª. Arizona: Wiley, 2012. 730 ISBN 978-1-118-14692-7.

PACEJKA, H. **Tire and vehicle dynamics**. Elsevier, 2005. ISBN 0080543332.

RAISINGHANI, M. S. et al. Six Sigma: concepts, tools, and applications. **Industrial Management & Data Systems**, v. 105, n. 4, p. 491-505, 2013-04-12 2013.

RANGEL, M. A. C. **Análise da percepção da sinalização vertical rodoviária em ambientes simulados de direção. Um estudo de caso na rodovia BR-116**. São Carlos: Departamento de Engenharia de Transportes, Escola de Engenharia de São Carlos, Universidade de São Paulo: 167 p. 2015.

RONEN, A.; YAIR, N. The adaptation period to a driving simulator. **Transportation Research Part F: Traffic Psychology and Behaviour**, v. 18, p. 94-106, 2013. ISSN 13698478.

SAS Institute Inc. **Discovering JMP 13®**. Cary, NC: SAS Institute Inc 2017.

SHECHTMAN, O. Validation of driving simulators. **Advances in Transportation Studies**, n. SPEC. ISSUE, p. 53-62, 2010.

STINE, J. S. et al. Analyzing the influence of median cross-section design on highway safety using vehicle dynamics simulations. v. 42, n. 6, p. 1769–1777, November 2010 2010. Available at: < <http://dx.doi.org/10.1016/j.aap.2010.04.018> >.

TORRES, A. L. M. **Análise de consistência de traçado de uma rodovia de múltiplas faixas**. 2015. (Thesis (Master degree in Transportation Engineering)). Engenharia de Transportes, Escola Politécnica, Universidade de São Paulo, São Paulo.

UNDERWOOD, G.; CRUNDALL, D.; CHAPMAN, P. Driving simulator validation with hazard perception. **Transportation Research Part F: Traffic Psychology and Behaviour**, v. 14, n. 6, p. 435-446, 2011.

VI-GRADE. **VI -CarRealTime 17.0: Documentation - VI -CarRealTime Framework - Build mode**. Germany: VI-grade engineering software & services 2015a.

_____. **VI-CarRealTime 17.0: Documentation - VI-Driver Theory**. Germany: VI-grade engineering software & services 2015b.

WHO. Global status report on road safety 2013. 2015-10-19 2013. Available at: <
http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/>.
 Accessed on: 19/oct/2015.

_____. Global status report on road safety 2015. **World Health Organization**, 2015-12-01 2015. Available at: <
http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/>.
 Accessed on: 01/dec/2015

APPENDIX A – Best practices on road safety legislation according to WHO

The best practices and its application can be completely found in the Global Status Report on Road Safety 2015 (Who, 2015). The following is only brief information.

- Speed in urban areas.

Reducing speed to 50 km/h in urban areas is based on pedestrian, cyclists and motorcyclists deaths and injuries reduction. According to WHO “an adult pedestrian has less than a 20% chance of dying if struck by a car at less than 50 km/h but almost a 60% risk of dying if hit at 80 km/h”. Local authorities also may have permission to change speed limit as needed.

- Motorcycle helmet use

A good quality helmet can reduce in 40% the death and in 70% the injury of any motorcycle passenger when used. The helmet quality standard must be specified by helmet Law. The Law should also cover all riders and be clear about the obligation of the use, despite the age of the rider.

- Driving - drinking

Drinking before drive increases the chance of an accident. Limit blood alcohol concentration to 0,05 g/dl may reduce road traffic crashes. The major challenge is not only Law creation, but also insures its application.

- Seat-belt use

Wearing a seat-belt reduces the risk of a fatality among drivers and front-seat occupants by 50% and by 25% for back-seat occupants. It also decreases the chance of being thrown from the vehicle in a crash event. The enactment of seat-belt law for all passengers and its enforcement are essentials in this case.

- Child restrain use

The usage of child restrain equipment reduces the risk of road traffic death in 90% for babies and infants (under 1 year) and 54% - 80% among children. The compliance to law is low. One of the reasons is the cost of child restrains what can be prohibitive to many families. In this case, government should work not only in enforcement but also to increase the access to the equipment.

APPENDIX B – Parameters of the analyzed highway curves

Curve Id	Begining (kilometer + meter)	End (kilometer + meter)	Length [m]	Radius [m]
C1	508 + 740	509+110	347	130
C2	509 + 175	509+510	338	130
C3	509 + 550	509+725	174	130
C4	509 + 960	510 + 125	142	615
C5	510 + 435	510 + 600	166	190
C6	510 + 875	511 + 120	261	130
C7	511 + 170	511 + 510	341	180
C8	511 + 575	511 + 835	263	230
C9	512 + 080	512 + 425	345	615
C10	512 + 540	512 + 910	369	190
C11	512 + 960	513 + 345	352	190
C12	513 + 645	513 + 810	162	230
C13	513 + 860	514 + 070	170	605
C14	514 + 640	514 + 845	205	130
C15	515 + 030	515 + 300	270	230
C16	515 + 515	515 + 805	288	190
C17	516 + 135	516 + 350	217	155
C18	516 + 395	516 + 735	342	215
C19	517 + 130	517 + 657	564	285
C20	517 + 825	518 + 209	424	185

APPENDIX C – Available information in Database

Field	Description	
Date	Date of the accident	day / month / year
Description	Type of damage caused by the accident	Property damage Victims Fatal victims Not defined
Type	Consequences of the accident	Collision Roll over Run off Overturning Not defined
Probable cause	Probable cause of the accident	Cutted Mechanical / Electrical Defect Distracted driving Aquaplaning Sideslip Oil on the track Drowsy driving Speeding Assault Others Parked vehicle on the road Not Defined Rain Object on the track

		Tire blowout
		Performance error
		Previous accident
		Cargo shifted
		Applied the brake suddenly
		Disregarding traffic signs
		Pedestrian on the road
		Flat tire
		Lane change
		Fog
		Driving under influence of liquor
		Bicyclist
		Jam
		Irregular turn of the road
Time	Time of occurrence of the accident	Hour:minutes
km	Kilometer of the highway where the accident occurred	Number between 509 and 518
mt	Meter of the highway where the accident occurred. Markings are 50 to 50 meters	Number between 0 and 950
Direction	Orientation of the highway. Data are only for southerly direction.	South
Highway	Highway identification. Data are only for a specific Federal Highway	--
Number of vehicles	Sum of all vehicles involved in the event	Cardinal number
Passenger Car	Number of passenger cars involved in the event	Cardinal number
Heavy vehicle	Number of heavy vehicles (trucks) involved in the	Cardinal number

	event	
Van	Number of vans involved in the event	Cardinal number
Motorcycle	Number of motorcycle involved in the event	Cardinal number
Bus	Number of bus involved in the event	Cardinal number
Pick up	Number of pick-ups involved in the event	Cardinal number
Others	Number of others types vehicles involved in the event	Cardinal number
Visibility condition		Good
		Partial
		Bad
		Not defined
Special condition	Description of abnormal conditions at the time of the accident	--
Weather condition	Description of climate conditions at the time of the accident	Normal condition
		Rain
		Drizzle
		Fog
		Cloudy
		Not identified
		Not defined
Track condition	Description of track conditions at the time of the accident	Wet
		Dry
		Not defined
Plan	Description of the plan type at the point of the accident	Straight
		Sharp curve
		Gentle curve
		Not defined
Profile	Description of track profile at the point of the	Uphill

accident

Downhill

In level

Not defined

APPENDIX D – DOE#01 Planning Form

DOE Planning Form

Title DOE #1 - Screening design to isolate relevant factors and levels understanding

Objective Understand how factors influence response variable, adjust levels, discard factors with low influence and so discovery a desired direction for further experiments.

Information acquired The analysis of accident data indicates factors not listed in the experiment as possible causes, for example, previous accident, mechanical failure, drowsiness at the wheel, overtaking in a prohibited place, etc. The maximum speed allowed in the stretch of interest of the highway is 80 km/h for light vehicles.

Experimental strategy and variables

Response variables	Measuring technique
Y ₁ . Number of occurrence of accident	Count (binary: 0 = no occurrence; 1 = accident)

Factors	Levels		Theory
	(-)	(+)	
A. Curve radius	130 m	230 m	Smoother curves (larger radii) have less risk of accidents.
B. Path profile	Downhill	Uphill	Slope contributes to vehicle acceleration, causing higher speeds to be achieved.
C. Path conditions	Wet (0,3 - 0,5)	Dry (0,7 - 0,9)	Wet track has lower coefficient of friction between tire and pavement.
D. Driver skill	Novice	Standard	Novices have greater reaction times and make more mistakes.
E. Speed	50 - 70 km/h	110 - 130 km/h	Higher speeds require greater control / dominance over the vehicle.
F. Period	Night	Day	The night time has reduced visibility (lack of light).
G. Load	1 person	4 persons	Increased number of occupants results in slower vehicle responses.

Nuisance variables	Control method
None	N/A

Forecasts

Main effects

(-) A (+)	(-) B (+)	(-) C (+)	(-) D (+)
(-) E (+)	(-) F (+)	(-) G (+)	

Generation and resolution

ABCD = F ABCE = G	2^{7-2}_{IV}

Method of nuisance treatment

Completed Random Design 3 replicates

APPENDIX E – FRD DOE #01

APPENDIX F – DOE #01 Prediction values based on linear regression model

Event	A	B	C	D	E	F	G	Prediction	LL	UL
1	-1	-1	-1	-1	-1	-1	-1	-0,1250	-0,3631	0,113095
2	-1	-1	-1	-1	-1	-1	1	0,0000	-0,2381	0,238095
3	-1	-1	-1	-1	-1	1	-1	0,0000	-0,2381	0,238095
4	-1	-1	-1	-1	-1	1	1	0,1250	-0,1131	0,363095
5	-1	-1	-1	-1	1	-1	-1	1,0000	0,761905	1,238095
6	-1	-1	-1	-1	1	-1	1	1,1250	0,886905	1,363095
7	-1	-1	-1	-1	1	1	-1	0,8750	0,636905	1,113095
8	-1	-1	-1	-1	1	1	1	1,0000	0,761905	1,238095
9	-1	-1	-1	1	-1	-1	-1	0,0000	-0,2381	0,238095
10	-1	-1	-1	1	-1	-1	1	-0,1250	-0,3631	0,113095
11	-1	-1	-1	1	-1	1	-1	0,1250	-0,1131	0,363095
12	-1	-1	-1	1	-1	1	1	0,0000	-0,2381	0,238095
13	-1	-1	-1	1	1	-1	-1	1,1250	0,886905	1,363095
14	-1	-1	-1	1	1	-1	1	1,0000	0,761905	1,238095
15	-1	-1	-1	1	1	1	-1	1,0000	0,761905	1,238095
16	-1	-1	-1	1	1	1	1	0,8750	0,636905	1,113095
17	-1	-1	1	-1	-1	-1	-1	-0,1250	-0,3631	0,113095
18	-1	-1	1	-1	-1	-1	1	0,0000	-0,2381	0,238095
19	-1	-1	1	-1	-1	1	-1	0,0000	-0,2381	0,238095
20	-1	-1	1	-1	-1	1	1	0,1250	-0,1131	0,363095
21	-1	-1	1	-1	1	-1	-1	1,0000	0,761905	1,238095
22	-1	-1	1	-1	1	-1	1	1,1250	0,886905	1,363095
23	-1	-1	1	-1	1	1	-1	0,8750	0,636905	1,113095
24	-1	-1	1	-1	1	1	1	1,0000	0,761905	1,238095
25	-1	-1	1	1	-1	-1	-1	0,0000	-0,2381	0,238095
26	-1	-1	1	1	-1	-1	1	-0,1250	-0,3631	0,113095
27	-1	-1	1	1	-1	1	-1	0,1250	-0,1131	0,363095
28	-1	-1	1	1	-1	1	1	0,0000	-0,2381	0,238095
29	-1	-1	1	1	1	-1	-1	1,1250	0,886905	1,363095
30	-1	-1	1	1	1	-1	1	1,0000	0,761905	1,238095
31	-1	-1	1	1	1	1	-1	1,0000	0,761905	1,238095
32	-1	-1	1	1	1	1	1	0,8750	0,636905	1,113095
33	-1	1	-1	-1	-1	-1	-1	-0,1250	-0,3631	0,113095
34	-1	1	-1	-1	-1	-1	1	0,0000	-0,2381	0,238095
35	-1	1	-1	-1	-1	1	-1	0,0000	-0,2381	0,238095
36	-1	1	-1	-1	-1	1	1	0,1250	-0,1131	0,363095
37	-1	1	-1	-1	1	-1	-1	1,0000	0,761905	1,238095
38	-1	1	-1	-1	1	-1	1	1,1250	0,886905	1,363095
39	-1	1	-1	-1	1	1	-1	0,8750	0,636905	1,113095
40	-1	1	-1	-1	1	1	1	1,0000	0,761905	1,238095
41	-1	1	-1	1	-1	-1	-1	0,0000	-0,2381	0,238095
42	-1	1	-1	1	-1	-1	1	-0,1250	-0,3631	0,113095
43	-1	1	-1	1	-1	1	-1	0,1250	-0,1131	0,363095
44	-1	1	-1	1	-1	1	1	0,0000	-0,2381	0,238095
45	-1	1	-1	1	1	-1	-1	1,1250	0,886905	1,363095
46	-1	1	-1	1	1	-1	1	1,0000	0,761905	1,238095
47	-1	1	-1	1	1	1	-1	1,0000	0,761905	1,238095
48	-1	1	-1	1	1	1	1	0,8750	0,636905	1,113095
49	-1	1	1	-1	-1	-1	-1	-0,1250	-0,3631	0,113095
50	-1	1	1	-1	-1	-1	1	0,0000	-0,2381	0,238095
51	-1	1	1	-1	-1	1	-1	0,0000	-0,2381	0,238095
52	-1	1	1	-1	-1	1	1	0,1250	-0,1131	0,363095
53	-1	1	1	-1	1	-1	-1	1,0000	0,761905	1,238095
54	-1	1	1	-1	1	-1	1	1,1250	0,886905	1,363095
55	-1	1	1	-1	1	1	-1	0,8750	0,636905	1,113095
56	-1	1	1	-1	1	1	1	1,0000	0,761905	1,238095
57	-1	1	1	1	-1	-1	-1	0,0000	-0,2381	0,238095
58	-1	1	1	1	-1	-1	1	-0,1250	-0,3631	0,113095
59	-1	1	1	1	-1	1	-1	0,1250	-0,1131	0,363095
60	-1	1	1	1	-1	1	1	0,0000	-0,2381	0,238095
61	-1	1	1	1	1	-1	-1	1,1250	0,886905	1,363095
62	-1	1	1	1	1	-1	1	1,0000	0,761905	1,238095
63	-1	1	1	1	1	1	-1	1,0000	0,761905	1,238095
64	-1	1	1	1	1	1	1	0,8750	0,636905	1,113095

Event	A	B	C	D	E	F	G	Prediction	LL	UL
65	1	-1	-1	-1	-1	-1	-1	0,0000	-0,2381	0,238095
66	1	-1	-1	-1	-1	-1	1	0,1250	-0,1131	0,363095
67	1	-1	-1	-1	-1	1	-1	-0,1250	-0,3631	0,113095
68	1	-1	-1	-1	-1	1	1	0,0000	-0,2381	0,238095
69	1	-1	-1	-1	1	-1	-1	0,8750	0,636905	1,113095
70	1	-1	-1	-1	1	-1	1	1,0000	0,761905	1,238095
71	1	-1	-1	-1	1	1	-1	0,5000	0,261905	0,738095
72	1	-1	-1	-1	1	1	1	0,6250	0,386905	0,863095
73	1	-1	-1	1	-1	-1	-1	0,1250	-0,1131	0,363095
74	1	-1	-1	1	-1	-1	1	0,0000	-0,2381	0,238095
75	1	-1	-1	1	-1	1	-1	0,0000	-0,2381	0,238095
76	1	-1	-1	1	-1	1	1	-0,1250	-0,3631	0,113095
77	1	-1	-1	1	1	-1	-1	1,0000	0,761905	1,238095
78	1	-1	-1	1	1	-1	1	0,8750	0,636905	1,113095
79	1	-1	-1	1	1	1	-1	0,6250	0,386905	0,863095
80	1	-1	-1	1	1	1	1	0,5000	0,261905	0,738095
81	1	-1	1	-1	-1	-1	-1	0,0000	-0,2381	0,238095
82	1	-1	1	-1	-1	-1	1	0,1250	-0,1131	0,363095
83	1	-1	1	-1	-1	1	-1	-0,1250	-0,3631	0,113095
84	1	-1	1	-1	-1	1	1	0,0000	-0,2381	0,238095
85	1	-1	1	-1	1	-1	-1	0,8750	0,636905	1,113095
86	1	-1	1	-1	1	-1	1	1,0000	0,761905	1,238095
87	1	-1	1	-1	1	1	-1	0,5000	0,261905	0,738095
88	1	-1	1	-1	1	1	1	0,6250	0,386905	0,863095
89	1	-1	1	1	-1	-1	-1	0,1250	-0,1131	0,363095
90	1	-1	1	1	-1	-1	1	0,0000	-0,2381	0,238095
91	1	-1	1	1	-1	1	-1	0,0000	-0,2381	0,238095
92	1	-1	1	1	-1	1	1	-0,1250	-0,3631	0,113095
93	1	-1	1	1	1	-1	-1	1,0000	0,761905	1,238095
94	1	-1	1	1	1	-1	1	0,8750	0,636905	1,113095
95	1	-1	1	1	1	1	-1	0,6250	0,386905	0,863095
96	1	-1	1	1	1	1	1	0,5000	0,261905	0,738095
97	1	1	-1	-1	-1	-1	-1	0,0000	-0,2381	0,238095
98	1	1	-1	-1	-1	-1	1	0,1250	-0,1131	0,363095
99	1	1	-1	-1	-1	1	-1	-0,1250	-0,3631	0,113095
100	1	1	-1	-1	-1	1	1	0,0000	-0,2381	0,238095
101	1	1	-1	-1	1	-1	-1	0,8750	0,636905	1,113095
102	1	1	-1	-1	1	-1	1	1,0000	0,761905	1,238095
103	1	1	-1	-1	1	1	-1	0,5000	0,261905	0,738095
104	1	1	-1	-1	1	1	1	0,6250	0,386905	0,863095
105	1	1	-1	1	-1	-1	-1	0,1250	-0,1131	0,363095
106	1	1	-1	1	-1	-1	1	0,0000	-0,2381	0,238095
107	1	1	-1	1	-1	1	-1	0,0000	-0,2381	0,238095
108	1	1	-1	1	-1	1	1	-0,1250	-0,3631	0,113095
109	1	1	-1	1	1	-1	-1	1,0000	0,761905	1,238095
110	1	1	-1	1	1	-1	1	0,8750	0,636905	1,113095
111	1	1	-1	1	1	1	-1	0,6250	0,386905	0,863095
112	1	1	-1	1	1	1	1	0,5000	0,261905	0,738095
113	1	1	1	-1	-1	-1	-1	0,0000	-0,2381	0,238095
114	1	1	1	-1	-1	-1	1	0,1250	-0,1131	0,363095
115	1	1	1	-1	-1	1	-1	-0,1250	-0,3631	0,113095
116	1	1	1	-1	-1	1	1	0,0000	-0,2381	0,238095
117	1	1	1	-1	1	-1	-1	0,8750	0,636905	1,113095
118	1	1	1	-1	1	-1	1	1,0000	0,761905	1,238095
119	1	1	1	-1	1	1	-1	0,5000	0,261905	0,738095
120	1	1	1	-1	1	1	1	0,6250	0,386905	0,863095
121	1	1	1	1	-1	-1	-1	0,1250	-0,1131	0,363095
122	1	1	1	1	-1	-1	1	0,0000	-0,2381	0,238095
123	1	1	1	1	-1	1	-1	0,0000	-0,2381	0,238095
124	1	1	1	1	-1	1	1	-0,1250	-0,3631	0,113095
125	1	1	1	1	1	-1	-1	1,0000	0,761905	1,238095
126	1	1	1	1	1	-1	1	0,8750	0,636905	1,113095
127	1	1	1	1	1	1	-1	0,6250	0,386905	0,863095
128	1	1	1	1	1	1	1	0,5000	0,261905	0,738095

APPENDIX G – DOE#02 Planning Form

DOE Planning Form

Title DOE #02 - Analysis of factor that affect traffic accidents

Objective Evaluate the probability of occurrence of an accident for different combinations of factors. More refine experiment. Identify possible variables to identify the imminence of accident (from vehicle dynamics point of view).

Information acquired Speed is has strong interaction with regards to occurrence of accidents. Second order interaction might be relevant when dealing with driver skill and load.

Experimental strategy and variables

Response variables	Measuring technique
Y ₁ . Number of occurrence of accident	Count (binary: 0 = no occurrence; 1 = accident)
Y ₂ . Path distance	Linear measurement of distance traveled until accident (continuous)

Factors	Levels		Theory
	(-)	(+)	
A. Curve radius	130 m	230 m	Smoother curves (larger radii) have less risk of accidents
B. Load	1 person	4 persons	Increased number of occupants results in slower vehicle responses
C. Driver skill	Novice	Standard	Novices have greater reaction times and make more mistakes.
D. Period	Night	Day	The night time has reduced visibility (lack of light)
E. Speed	60 - 80 km/h	90 - 120 km/h	Vehicles at high speed are more unstable and may increase the probability of accident

Nuisance variables	Control method
None	N/A

Forecasts

Main effects

(-) A (+)	(-) B (+)	(-) C (+)	
(-) D (+)	(-) E (+)		

Generation and resolution

$ABCD = E \quad 2_{IV}^{5-1}$

Method of nuisance treatment

Completed Random Design Number of replicates: 14 per treatment

APPENDIX H – FRD DOE #02

APPENDIX I – DOE #02 Prediction values based on linear regression model

Event	A	B	C	D	E	Prediction	LL	UL
1	-1	-1	-1	-1	-1	0,2946	0,0072	0,5821
2	-1	-1	-1	-1	1	0,9911	0,7036	1,2785
3	-1	-1	-1	1	-1	0,0089	-0,2785	0,2964
4	-1	-1	-1	1	1	0,3839	0,0965	0,6714
5	-1	-1	1	-1	-1	0,4196	0,1322	0,7071
6	-1	-1	1	-1	1	1,1161	0,8286	1,4035
7	-1	-1	1	1	-1	0,1339	-0,1535	0,4214
8	-1	-1	1	1	1	0,5089	0,2215	0,7964
9	-1	1	-1	-1	-1	0,7411	0,4536	1,0285
10	-1	1	-1	-1	1	1,4375	1,1501	1,7249
11	-1	1	-1	1	-1	0,4554	0,1679	0,7428
12	-1	1	-1	1	1	0,8304	0,5429	1,1178
13	-1	1	1	-1	-1	0,3304	0,0429	0,6178
14	-1	1	1	-1	1	1,0268	0,7393	1,3142
15	-1	1	1	1	-1	0,0446	-0,2428	0,3321
16	-1	1	1	1	1	0,4196	0,1322	0,7071
17	1	-1	-1	-1	-1	0,1339	-0,1535	0,4214
18	1	-1	-1	-1	1	0,8304	0,5429	1,1178
19	1	-1	-1	1	-1	-0,1518	-0,4392	0,1357
20	1	-1	-1	1	1	0,2232	-0,0642	0,5107
21	1	-1	1	-1	-1	0,2589	-0,0285	0,5464
22	1	-1	1	-1	1	0,9554	0,6679	1,2428
23	1	-1	1	1	-1	-0,0268	-0,3142	0,2607
24	1	-1	1	1	1	0,3482	0,0608	0,6357
25	1	1	-1	-1	-1	0,3304	0,0429	0,6178
26	1	1	-1	-1	1	1,0268	0,7393	1,3142
27	1	1	-1	1	-1	0,0446	-0,2428	0,3321
28	1	1	-1	1	1	0,4196	0,1322	0,7071
29	1	1	1	-1	-1	-0,0804	-0,3678	0,2071
30	1	1	1	-1	1	0,6161	0,3286	0,9035
31	1	1	1	1	-1	-0,3661	-0,6535	-0,0786
32	1	1	1	1	1	0,0089	-0,2785	0,2964

APPENDIX J – DOE #02 Y2: Path distance per treatment

APPENDIX K – DOE #02: Roll Angle by Treatment

The following results are grouped by treatment.

TREATMENT 2

TREATMENT 3

TREATMENT 4

TREATMENT 5

TREATMENT 6

TREATMENT 7

TREATMENT 8

TREATMENT 9

TREATMENT 10

TREATMENT 11

TREATMENT 12

TREATMENT 13

TREATMENT 14

TREATMENT 15

APPENDIX L – DOE #02: Yaw Rate by Treatment

The following results are by treatment.

TREATMENT 2

TREATMENT 3

TREATMENT 4

TREATMENT 5

TREATMENT 6

TREATMENT 7

TREATMENT 8

TREATMENT 9

TREATMENT 10

TREATMENT 11

TREATMENT 12

TREATMENT 13

TREATMENT 14

TREATMENT 15

TREATMENT 16

ADDENDUM A – List of two-level fractional factorial designs
(Box et al., 2005, p.272)

Table 6.22. Minimal Aberration Two-Level Fractional Factorial Designs for k variables and N Runs
(Number in Parentheses Represent Replication)

		number of variables k									
		3	4	5	6	7	8	9	10	11	
4	2^{3-1}_{III}										
8	2^3	2^{3-1}_{III}	2^{4-1}_{IV}	2^{5-1}_{V}	2^{6-1}_{VI}	2^{7-1}_{VII}	2^{8-1}_{VIII}	2^{9-1}_{IX}	2^{10-1}_{X}	2^{11-1}_{XI}	
		$\pm 3 = 12$		$\pm 4 = 12$ $\pm 5 = 13$	$\pm 4 = 12$ $\pm 5 = 13$ $\pm 6 = 23$	$\pm 4 = 12$ $\pm 5 = 13$ $\pm 6 = 23$ $\pm 7 = 123$	$\pm 5 = 123$ $\pm 6 = 234$ $\pm 7 = 134$ $\pm 8 = 1234$	$\pm 5 = 123$ $\pm 6 = 234$ $\pm 7 = 134$ $\pm 8 = 1234$ $\pm 9 = 12345$	$\pm 5 = 123$ $\pm 6 = 234$ $\pm 7 = 134$ $\pm 8 = 1234$ $\pm 9 = 12345$ $\pm 10 = 123456$	$\pm 5 = 123$ $\pm 6 = 234$ $\pm 7 = 134$ $\pm 8 = 1234$ $\pm 9 = 12345$ $\pm 10 = 123456$ $\pm 11 = 1234567$	
16	2^4	2^4	2^{5-1}_{V}	2^{6-1}_{VI}	2^{7-1}_{VII}	2^{8-1}_{VIII}	2^{9-1}_{IX}	2^{10-1}_{X}	2^{11-1}_{XI}	2^{12-1}_{XII}	
	2 times		$\pm 4 = 123$	$\pm 5 = 1234$	$\pm 5 = 123$ $\pm 6 = 234$	$\pm 5 = 123$ $\pm 6 = 234$ $\pm 7 = 134$	$\pm 6 = 123$ $\pm 7 = 124$ $\pm 8 = 2345$	$\pm 6 = 2345$ $\pm 7 = 1345$ $\pm 8 = 1245$ $\pm 9 = 1235$	$\pm 6 = 1234$ $\pm 7 = 1235$ $\pm 8 = 1345$ $\pm 9 = 1245$ $\pm 10 = 2345$	$\pm 6 = 123$ $\pm 7 = 234$ $\pm 8 = 345$ $\pm 9 = 134$ $\pm 10 = 145$ $\pm 11 = 245$	$(\frac{1}{128})$
32	2^5	2^5	2^5	2^6	2^{7-1}_{VII}	2^{8-1}_{VIII}	2^{9-1}_{IX}	2^{10-1}_{X}	2^{11-1}_{XI}	2^{12-1}_{XII}	
	4 times		2 times		$\pm 6 = 1234$ $\pm 7 = 1245$	$\pm 6 = 1234$ $\pm 7 = 1245$	$\pm 6 = 123$ $\pm 7 = 124$ $\pm 8 = 2345$	$\pm 6 = 2345$ $\pm 7 = 1345$ $\pm 8 = 1245$ $\pm 9 = 1235$	$\pm 6 = 1234$ $\pm 7 = 1235$ $\pm 8 = 1345$ $\pm 9 = 1245$ $\pm 10 = 2345$	$\pm 6 = 123$ $\pm 7 = 234$ $\pm 8 = 345$ $\pm 9 = 134$ $\pm 10 = 145$ $\pm 11 = 245$	$(\frac{1}{64})$
64	2^6	2^6	2^6	2^6	2^6	2^{7-1}_{VII}	2^{8-1}_{VIII}	2^{9-1}_{IX}	2^{10-1}_{X}	2^{11-1}_{XI}	
	8 times		4 times	2 times		$\pm 7 = 123456$	$\pm 7 = 1234$ $\pm 8 = 1256$	$\pm 7 = 1234$ $\pm 8 = 1356$ $\pm 9 = 3456$	$\pm 7 = 2346$ $\pm 8 = 1346$ $\pm 9 = 1245$ $\pm 10 = 1235$	$\pm 7 = 345$ $\pm 8 = 1234$ $\pm 9 = 126$ $\pm 10 = 2456$ $\pm 11 = 1456$	$(\frac{1}{32})$
128	2^7	2^7	2^7	2^7	2^7	2^7	2^{8-1}_{VIII}	2^{9-1}_{IX}	2^{10-1}_{X}	2^{11-1}_{XI}	
	16 times		8 times	4 times	2 times	$\pm 8 = 1234567$	$\pm 8 = 1234567$	$\pm 8 = 13467$ $\pm 9 = 23567$	$\pm 8 = 1237$ $\pm 9 = 2345$ $\pm 10 = 2345$ $\pm 11 = 1234567$	$\pm 8 = 1237$ $\pm 9 = 2345$ $\pm 10 = 2345$ $\pm 11 = 1234567$	$(\frac{1}{16})$