UNIVERSITY OF SÃO PAULO
SÃO CARLOS SCHOOL OF ENGINEERING

Fernando José Piva

Freeway level of service criteria based on travelers' perception

São Carlos

2022

Fernando José Piva

Freeway level of service criteria based on travelers' perception

Thesis presented to the São Carlos School of Engineering of the University of São Paulo in partial fulfillment of the requirements for the degree of Doctor of Science in Department of Transportation Engineering

Subject area: Transport Planning and Operations

Advisor: Prof. Dr. José Reynaldo Setti

REVISED VERSION

São Carlos

2022

Catalog card prepared by Patron Service at “Prof. Dr. Sergio Rodrigues Fontes Library” at EESC/USP

P693f

Piva, Fernando José
Freeway level of service criteria based on travelers’ perception. / Fernando José Piva; advisor José Reynaldo Setti. -- São Carlos, 2022.

Doctoral (Dissertation) – Graduate Program in Transportation Engineering and Concentration area in Planning and Operation of Transport Systems. -- São Carlos School of Engineering, at University of São Paulo, 2022.

1. Level of service. 2. Freeway operations. 3. User perception. 4. Online surveys. 5. Microsimulation. I. Título.

Prepared by Eduardo Graziosi Silva- CRB-8/8907

FOLHA DE JULGAMENTO

Candidato: Engenheiro **FERNANDO JOSÉ PIVA.**

Título da tese: "Critérios de nível de serviço em rodovias com base na percepção dos usuários".

Data da defesa: 26/10/2022.

Comissão Julgadora	**Resultado**
Prof. Titular **Jose Reynaldo Anselmo Setti** **(Orientador)** (Escola de Engenharia de São Carlos – EESC/USP)	APROVADO
Prof. Dr. **Scott Stuart Washburn** (University of Florida/UF)	APROVADO
Profa. Dra. **Helena Beatriz Bettella Cybis** (Universidade Federal do Rio Grande do Sul/UFRGS)	APROVADO
Prof. Dr. **André Luiz Barbosa Nunes da Cunha** (Escola de Engenharia de São Carlos – EESC/USP)	APROVADO
Prof. Dr. **Flávio José Craveiro Cunto** (Universidade Federal do Ceará/UFC)	APROVADO

Coordenadora do Programa de Pós-Graduação em Engenharia de Transportes:
Profa. Associada **Ana Paula Camargo Larocca**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

Dedico esta tese aos meus pais.
Sem o apoio deles, nada disso seria possível!

Acknowledgements

Em primeiro lugar, agradeço à minha família. Meus pais, Celso e Teresinha, meus irmãos Marcelo e Fábio, minhas cunhadas Regiane e Carla e meus sobrinhos Luís, Henrique e Amanda. Vocês foram fundamentais durante esse processo;

Aos tios Celso, Teresa e Cida e aos primos Rogério e Ricardo, agradeço pelas ajudas em todas as fases da minha vida;

Ao Prof. José Reynaldo Setti, muito obrigado por todas as incansáveis horas de correções, sugestões e questionamentos ao longo desta pesquisa. Certamente elas tornaram essa pesquisa muito mais robusta e bem escrita;

To Dr. Scott Washburn, thank you for accepting me at the University of Florida as a visiting scholar, to participate in the proposal committee, for all meetings and corrections after I returned to Brazil;

Meus agradecimentos à professora Cira Pitombo pelas contribuições no exame de qualificação, pelas diversas reuniões e pelas conversas ao longo dessa pesquisa. Aos professores André Cunha, Cassiano Isler e Linda Ho, pelas ajudas nas soluções de problemas ao longo desta pesquisa;

Aos professores Flávio Cunto, Helena Cybis, Scott Washburn e André Cunha por participarem da defesa do doutorado e colaborarem com importantes comentários para a melhora da versão corrigida deste documento;

Ao Gustavo, Karla, Fábio e Renata, que me receberam em Gainesville durante o período do doutorado sanduíche. Com certeza a vida nos EUA ficou muito mais fácil com as suas incontáveis ajudas e conversas;

Agradeço a todos os amigos que participaram dessa jornada... seja do STT, do grupo de pesquisa do professor Setti, da Flórida, de São Paulo, de Matão... aqui me reservo ao direito de não citar nomes para não ser injusto com ninguém;

Agradeço ao suporte financeiro dado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001 e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), através da bolsa de doutorado e pela bolsa do Programa de Doutorado Sanduíche no Exterior (PDSE) – Processo 88881.131593/16-01;

À USP e à UF, por proporcionar as condições necessárias para a produção desta pesquisa de doutorado. Extendo também o agradecimento a todos os funcionários das duas universidades. De alguma forma, todos vocês foram de grande ajuda para sanar todos os problemas.

Abstract

PIVA, F. J. **Freeway level of service criteria based on travelers' perception**. 2022. 91p. Thesis (Doctoral) - São Carlos School of Engineering, University of São Paulo, São Carlos, 2022.

Level of service (LOS) classifications of traffic operational conditions play a large role in roadway-improvement funding decisions. Consequently, it is essential that traveler perception of LOS is consistent with the values determined through a traffic analysis. Otherwise, the public confidence in appropriate transportation agency funding decisions may be undermined. Research to measure travelers' perceptions of traffic operational conditions is a relatively immature field and methods to obtain such data are still evolving. Methods used to collect data on the travelers' perception in previous studies include driver interviews at rest stops, focus groups, and rating video clips recorded from the driver's field-of-view. These videos are recorded on real roadways for greater realism but the capture of the full range of operational conditions in a timely manner is its greatest problem. Another significant challenge that all the methods in the literature face is collecting a sufficiently large sample of responses. This dissertation describes an approach to incorporate travelers' perception of trip quality that addresses the limitations of the previous studies by the use of a combination of realistic looking 3-dimensional traffic stream visualizations and an online survey to reach a large number of people. The use of traffic simulation software provides full control of the traffic and roadway characteristics, and allows for the creation of realistic computer generated animations over the full range of operating conditions in a fast, efficient and economical manner. Furthermore, modern microsimulation software is able to record video clips from the driver's viewpoint, which is likely to elicit more accurate rankings from the study participants than an overhead view. The creation of the roadway environment to produce a realistic view from the vehicle's cabin interior and an automated method for choosing a representative vehicle from the traffic stream are presented. Data on drivers' perception of the quality of the trips depicted in the video clips are obtained inviting participants to respond to a web-based survey. In the web survey, participants watched a set of freeway trips under the desired range of roadway and traffic conditions and rate each trip using a visual analog scale that varies from 0 to 100. The travelers' ratings are then discretized into the desired number of service levels using cluster analysis and the service level boundaries and the associate confidence intervals are estimated using logistic regression.

The feasibility of the proposed approach was demonstrated with a case study with 977 participants that rated 10,228 trips depicting 128 different combinations of traffic density, truck percentage, posted speed limit, grade steepness, and number of lanes, chosen according to a fractional factorial design. After a series of filters, the final sample consisted in 6231 ratings by 554 participants. A statistically significant correlation between traffic density and rating was found ($\rho = -0.532$, $N = 6231$). The case study results indicate that the HCM-7 freeway LOS boundaries are within the confidence intervals estimated using the proposed approach, in the exception of the threshold between LOS D and E. This suggests that, at least for the participants in the study, that drivers perceive LOS D and E as very similar. The effect of sociodemographic, traffic and roadway factors on the participants' perception of trip quality was investigated using a fractional factorial design. The sample consisted of 7,004 ratings by 625 participants. A total of 16 factors were analyzed using bivariate correlation and traffic density was the only factor that showed a significant, strong correlation with rating ($\rho = 0.520$, $p < 0.001$). Truck percent was also correlated with rankings, at a much lower level ($\rho = -0.116$, $p < 0.001$) and no significant influence from sociodemographic factors was found. These results were confirmed by a stepwise multiple linear regression model calibrated to the data. A CART decision tree model indicated that participants' perception of the trip quality tend to be affected by the presence of trucks in the stream when traffic density is greater than 7 veh/km/ln.

Keywords: Level of service. Freeway operations. User perception. Online surveys. Microsimulation.

Resumo

PIVA, F. J. **Critérios para estimar níveis de serviço em rodovias de pista dupla pela percepção dos usuários**. 2022. 91p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2022.

As classificações de nível de serviço (NS) das condições operacionais de tráfego realizam um papel importante nas decisões de financiamento da melhoria das estradas. Consequentemente, é essencial que a percepção do usuário no NS seja consistente com os valores determinados por meio de uma análise do tráfego. Caso contrário, a confiança nas decisões apropriadas de financiamento das agências públicas de transporte podem ser prejudicadas. Pesquisas que medem a percepção dos usuários sobre as condições operacionais do tráfego é um campo relativamente novo e os métodos para obter esses dados ainda estão sendo desenvolvidos. Os métodos usados para coletar a percepção dos usuários em estudos anteriores incluem entrevistas com motoristas em paradas de descanso, grupos focais e avaliação de vídeos gravados do ponto de vista do motorista. Esses vídeos são gravados em estradas reais para maior realismo, mas a captura de todas as condições de tráfego em tempo hábil é seu maior problema. Outro desafio que os métodos encontrados na literatura enfrentam é coletar uma amostra suficientemente grande de respostas. Esta dissertação descreve um método para coletar a percepção dos usuários na qualidade da viagem, suprindo as limitações dos estudos anteriores. Utilizou-se a combinação de vídeos realistas tridimensionais do fluxo de tráfego e uma pesquisa *online* para alcançar muitas pessoas. O uso de *software* de simulação de tráfego fornece controle total das características do tráfego e da rodovia, além de permitir a criação de animações realistas geradas por computador em todas as condições de operação de maneira rápida, eficiente e econômica. Além disso, os *softwares* de microssimulação modernos são capazes de gravar vídeos do ponto de vista do motorista, o que provavelmente obterá coleta da percepção dos participantes mais precisas do que uma visão aérea. Foram apresentados a criação do ambiente para produzir uma visão realista do interior do veículo e um método automatizado para escolher um veículo representativo do fluxo de tráfego. As percepções dos motoristas sobre a qualidade das viagens apresentados nos vídeos são obtidas convidando participantes a responder a uma pesquisa *online*. Nessa pesquisa, os participantes assistem a um conjunto de vídeos que representam viagens em rodovias sob desejadas condições de tráfego e avaliam cada viagem usando uma escala visual analógica que varia de 0 a 100. Essa avaliação dos usuários são então discretizadas no número desejado de níveis de serviços usando análise de cluster e os limites dos níveis de serviço e os intervalos de confiança associados são estimados usando regressão logística.

A viabilidade do método proposto foi demonstrada com um estudo de caso com 977 participantes que avaliaram 10.228 viagens representando 128 combinações diferentes de densidade de tráfego, porcentagem de caminhões, limite de velocidade da rodovia, inclinação da rodovia e número de faixas de tráfego, escolhidos de acordo com um desenho fatorial fracionário. Após uma série de filtros aplicados nos dados coletados, a amostra final consistiu em 6231 avaliações de 554 participantes. Foi encontrada uma correlação estatisticamente significativa entre densidade de tráfego e a avaliação dos vídeos ($\rho = -0,532$, $N = 6231$). Os resultados do estudo de caso indicam que os limites de NS apresentado pelo HCM-7 estão dentro dos intervalos de confiança estimados na presente pesquisa, com exceção do limite entre NS D e E. Isso sugere que, para os participantes do estudo, os motoristas percebem os limites de NS D e E como muito semelhantes. O efeito de fatores sociodemográficos e fatores de tráfego na percepção dos usuários sobre a qualidade da viagem foi investigado usando um desenho fatorial fracionário. A amostra consistiu em 7.004 avaliações de 625 participantes. Um total de 16 fatores foram analisados usando correlação bivariada. A densidade de tráfego foi o único fator que mostrou uma correlação significativa e forte com a percepção dos usuários ($\rho = 0,520$, $p < 0,001$). A porcentagem de caminhões também foi correlacionada com as notas, em um nível muito inferior ($\rho = -0,116$, $p < 0,001$) e não foi encontrada influência significativa de fatores sociodemográficos. Esses resultados foram confirmados por um modelo de regressão linear múltipla *stepwise*. Um modelo de árvore de decisão CART indicou que a percepção dos participantes sobre a qualidade da viagem tende a ser afetada pela presença de caminhões na corrente de tráfego quando a densidade de tráfego é maior que 7 veíc/km/fx.

Palavras-chave: Nível de serviço. Operações de rodovias de pista dupla. Percepção do usuário. Pesquisas *online*. Microssimulação.

Contents

Chapter 1

Introduction

Essential highways in Brazil are run by private sector companies under the supervision of regulatory agencies. Currently, 21 of these companies operate more than 10,000 km of the Brazilian federal highway system [1]. Specifically in the state of São Paulo (the most populous state with the highest gross domestic product in the country) [2], some of its main highways are managed by 20 private companies, covering 11,700 km [3], of which more than 5,000 km are freeways [4].

Regulatory agencies are responsible for assessing the quality of service of highway segments. According to the LOS (level of service) of these segments, additional lanes are built by private companies, following the instructions contained in contracts signed by companies and regulatory agencies (concession agreements). In Brazil, these regulatory agencies adopt different HCM (Highway Capacity Manual) versions to evaluate the LOS [5]. Likewise, the DNIT (Brazilian National Transport Infrastructure Department) also recommends the use of HCM methods in the absence of adaptations to local conditions [6]. However, since the HCM is based on default values and typical applications calculated for North America, particularly the United States, its use in other locations requires calibration of parameters, equations, and procedures [7].

Germany and the Netherlands already adapted the HCM to their local conditions by recalibrating parameters, curves, and equations using data collected in the respective countries [8, 9]. These "local" highway capacity manuals represent drivers, infrastructure, and vehicles more precisely. Previous research has shown the need to adapt the HCM procedures to Brazilian conditions due to the inaccuracy of parameters, vehicle conditions, and highway infrastructure [10–13].

First presented in the second edition of HCM 1965 [14], the LOS concept measures the quality of service in terms of operating speed and volume-to-capacity ratio (v/c). In its third edition, released in 1985 [15], basic segments of freeways were assessed for the first time using traffic density. In the 2010 edition [16], the LOS was based not only on the traditional measures but also on how users perceived the quality of service [17]. A study introduced the users' perception to estimate the LOS criteria for different modes of transportation, such as automobiles, buses, walking, and bicycles. The travelers watched videos from the users' point of view, with various travel conditions. After watching the videos, they were asked to rate the quality of service in each video [18].

The seventh HCM edition describes the quality of service as "how well a transportation facility or service operates from the travelers' perspective", while the LOS is defined as a "quantitative stratification of a performance measure or measures that represent the quality of service". Historically, the HCQS (Highway Capacity and Quality of Service) Committee have established service measures and threshold values to define the LOS [7]. However, it never considered the users' perception when setting the LOS thresholds for freeways. The differences in v/c for each LOS are almost equal, especially among levels B, C, D and E, in disagreement with research results on how drivers perceive the quality of service [19–21].

Some studies included users' perception of trip quality to evaluate different transportation system services, such as urban streets [22, 23], signalized intersections [24, 25], multimodal transportation [26–28], pavement serviceability of urban roads [29], and pedestrian quality of service [30, 31]. In these studies, the researchers combined data collected from the transportation system with the users' answers to a questionnaire about their opinion on the quality of service.

In other studies, in addition to travelers' perception of toll plaza services [32, 33], freeway merging conflicts [34], and freeway segments [21, 35–38], the authors also collected some sociodemographic information to evaluate freeway facilities. Unfortunately, most of these studies did not include in their results analyses of sociodemographic factors.

In order to create a framework to understand the travelers' perception of the quality of service, some researchers have conducted studies focusing on traffic facilities using different techniques, such as fitted utility functions [39, 40], participation in focus groups [41–44], and interviews with drivers after traveling on a highway [21, 35, 45] or watching videos recorded from inside a car [19, 20, 37, 46, 47]. Nonetheless, these studies share the same shortcomings: (a) the difficulty in obtaining a large sample of respondents that would be representative of the driving population; and (b) the difficult in presenting to the respondents a sufficiently large set of driving conditions. Ideally, a comprehensive range of operating conditions would be presented to a representative sample of drivers to obtain the necessary data to correlate a measure of service with the users' perception of the quality of service.

Videos have been created in traffic simulators using different operating conditions to assess users' perception of the quality of service. For instance, Obelheiro, Cybis and Ribeiro[32] used a website to collect data on the users' perception of the quality of Brazilian toll plazas services. In this study, a series of videos of a toll plaza operating at various congestion levels was produced using VISSIM. The videos were made from a bird's eye view to show the queue length at toll booths. The respondents were invited to watched them and rate the quality of service. Paiva and Setti[36] proposed the use of traffic simulation software to create videos from the perspective drivers. Such approach allows a complete control of traffic variables and the creation of different scenarios to describe the full range of operating conditions. The drivers were asked to rate the trip quality in the video clips shown. The feasibility of this method was demonstrated by a pilot study using a convenience sample of university students. The scores attributed by the respondents were associated with traffic density.

1.1 Research objectives and questions

The main objective of this research was to develop a method to incorporate an extensive quantity of data on the drivers' perception of trip quality into the estimation of LOS thresholds for freeway segments, besides investigating traffic and sociodemographic factors that might affect such perception. Thus, the following research questions were proposed:

1. How can the users' perception of the quality of service on freeways can be measured?
2. Is there a correlation between quality of service perceived by drivers and any service measure?
3. How can density-based LOS thresholds be estimated from the drivers' perception of the quality of service?
4. Do freeway and traffic stream characteristics (density, grade magnitude, number of lanes, truck percentage, and speed limit) affect how users perceive the trip quality?
5. Do sociodemographic characteristics (e.g., age, sex and travel frequency) affect the drivers' perception of the trip quality?

The method proposed to achieve the main objective and answer the research questions is based on the approach previously used by Paiva and Setti[36].

1.2 Method

The instrument proposed for data collection consisted of a socioeconomic questionnaire to characterize the respondents as well as the presentation of a sequence of video clips depicting different traffic operating conditions. The questionnaire was an online survey in order to reach a large sample of respondents from different locations at a low cost and in a relatively short time. The respondents were asked to rate the trip quality according to the video clips shown using a continuous visual analog scale.

A traffic simulator generated videos showing different freeway and traffic characteristics and various operating conditions. Parameters for local conditions were used to create accurate video clips, that is, as close as possible to the reality faced by travelers. Thus, this works is structured as follows: section 2.4 describes the elements and network details to create the freeway environment, section 2.5 addresses the vehicle's internal parts shown in the videos; and section 2.6 presents the method used to find a vehicle in a traffic stream that better represents the traffic conditions recorded in the video clips.

The scenarios used to create the video clips included five traffic factors, which would generate 1024 scenarios if created with the full factorial design. Therefore, to reduce the number of scenarios to 128, a 1/8 fractional factorial design was used. A set of short video clips showing the trip from the drivers' perspective was created to represent the range of operating conditions, as shown in Chapter 2.

The online questionnaire included 11 questions about sociodemographic data and the ratings of 12 video clips on a continuous scale. Section 3.3.2 discusses the website framework that allowed users to access it from different devices, whilst sections 3.4.3 and 4.4 describe the characteristics of the freeway users who responded to the survey. Sections 3.4.4 and 4.4 present the data checks used to discard outlier data.

The LOS boundaries were estimated using cluster analyses and logistic regression. Section 3.3.3 shows the method proposed to estimate the LOS thresholds, whereas section 3.4.5 details how they were calculated for a case study in Brazil. An exploratory analysis to evaluate the impact of sociodemographic and traffic factors on the travelers' perception of the quality of service using the respondents' ratings as a dependent variable for bivariate correlation, linear regression, and decision tree analyses is included in section 4.5.

1.3 Thesis structure

This thesis contains five chapters. This first chapter presents the introduction, the main research objective, the research questions, and the proposed method. Chapter 2, which is a paper published in *Promet – Traffic&Transportation* [48], explains the method used to create video clips with the desired traffic and highway characteristics and partially answers the first research question. Chapter 3 details the proposed method for estimating LOS thresholds based on the travelers' perception of the quality of service and addresses research questions 1, 2 and 3. Chapter 4 shows an exploratory analysis of the factors that affect travelers' perception of the quality of service, covering the two final research questions. Lastly, the last chapter brings the conclusions and final remarks. Chapters 2, 3, and 4 contain the literature review, conclusions, recommendations, and suggestions for future research pertaining to the topic of each chapter.

Chapter

The creation of video clips for data collection

2

This chapter describes a novel method to create the video clips used for data collection with greater time and cost efficiency. This new method combines traffic microsimulation and 3-D visualization capabilities. The focus of this chapter is to provide guidance on how to apply traffic microsimulation and computer 3-D visualization to evaluate highway trip quality from a traveler's perspective. It discusses the creation of the simulation environment to produce a realistic view from the vehicle's cabin interior, including the network creation, landscaped area, dashboard speedometer and rear-view mirror. The chapter also presented an automated method for choosing an appropriate vehicle within the simulated traffic stream, such that the desired overall traffic stream conditions are conveyed to the traveler vehicle within the field of view.

The contents of the chapter were presented as a poster in the 99th Annual Meeting of the Transportation Research Board (TRB), in 2020. Subsequently, a revised and expanded version of that paper was published by *Promet – Traffic & Transportation* [48].

2.1 Introduction

The concept of level of service (LOS) was introduced in the second edition of the Highway Capacity Manual (HCM), in 1965 [14]. LOS refers to assigning a letter grade (from A, the best, to F, the worst) that corresponds to the general operational conditions, as measured by one or more performance measures, referred to as service measures. In the 1965 HCM, the selected service measures for freeways were operating speed and volume-to-capacity ratio [17]; in the third edition of the HCM, released in 1985 [15], basic freeway segments were assessed for the first time, with density chosen as the service measure [17], which was kept in the subsequent editions.

While the service measures and thresholds for LOS rankings have historically been chosen by transportation engineers involved in the development of the HCM, the intent is that these ranking thresholds would also be reasonably consistent with how the traveling public would perceive the quality of the operational conditions. Given the very large amount of funding typically involved in transportation infrastructure decisions, it is vitally important to transportation agencies that there be consistency between the quantitative changes to the operational conditions as a result of the investment and the perception of these changes by the users of the transportation system. However, until relatively recently, research on user perceptions of traffic operational conditions was lacking.

Some initial progress was made in this area in the HCM 2010 edition [16]. In that edition, LOS assessment based on user perception research was introduced for the bicycle, pedestrian, and transit modes on urban streets. Research on user perceptions of traffic operational conditions has gained momentum over the last 15 years, but is still a relatively immature area of research. Particularly challenging is developing a research approach to accurately measure user perceptions of traffic operational conditions that is also time and cost efficient.

This chapter describes such an approach, which uses a combination of traffic microsimulation and three-dimensional (3-D) visualization. Until relatively recently, computer visualizations within traffic microsimulation programs have been in only two dimensions or of low visual quality if in three dimensions. On the other hand, dedicated computer visualization programs usually provided realistic representations of the roadway environment but the included traffic conditions were not grounded in traffic flow theory. Desktop computing performance has now reached a level where realistic-looking 3-D visualizations can be created directly within the traffic simulation environment. Given the novelty and nontrivial nature of this research approach, it is important to share the lessons learned in developing this approach so as to assist the research community in applying this promising approach to future efforts. Specifically, we addressed the topics of creating the simulation network, creating the roadway view from the vehicle's cabin interior, and developing an automated method for choosing a representative vehicle from the overall simulated traffic stream.

2.2 Literature review

The literature shows two different approaches to understand how travelers perceive and are affected by the quality of trips. The first one is focused on the individual's subjective well-being, looking for elements that explain travel satisfaction [49] and tries to assess the individual benefits from travel improvements [50] through a psychological scale such as the satisfaction with travel scale [51].

The second approach tries to correlate engineering measures of effectiveness to users' perception of the quality of the trip aiming to somehow include the traveling public perception in the performance evaluation. Since the 2000s, several studies have been conducted to create a framework to understand the traveler's perception of the quality of service, across a variety of traffic facilities – e.g., freeways [19, 35, 52], two-lane highways, intersections [53, 54], roundabouts, sidewalks [55] and bicycle paths [56]. The study [18] that culminated with the introduction of user perception for multi-modal LOS in the HCM 2010 utilized a procedure for assessing LOS that consisted of having travelers view and rate video recordings of a variety of travel conditions. More specifically, a broad range of users were shown video clips recorded on typical urban street segments from the point-of-view of automobile drivers, pedestrians, and bicycle riders and were asked to rate the quality of service in each video clip. For this study, 26 to 35 videos were recorded for each travel mode. Multiple linear regression and cumulative logistic regression methods were used to fit models that could estimate the mean rating obtained for each video. The best model used three independent variables (stops per mile, presence of trees, and presence of an exclusive left-turn lane) and explained approximately 75% of the variation in mean observed LOS ratings. This model predictive performance was noticeably better than that of HCM LOS method (which explained only 46% of the variation in mean observed LOS ratings), clearly indicating the importance of including the users' perception in LOS evaluation.

For freeways and highways, other researchers have also used video clips to evaluate how drivers perceive the quality of service [19, 20, 37, 46, 47, 52]. In all these cases, the video clips were recorded on real roadways, under the prevailing operational conditions at the time; consequently, the range of conditions were limited. Other techniques used include focus groups [41–44] and interviews with drivers after traveling on a highway [21, 35, 45]. These approaches generally share the same challenge: presenting to participants a sufficiently large set of driving conditions that would include the full range of operating conditions. A study using a convenience sample of 20 university students [38] investigated drivers' behavior and perceived quality of travel using a driving simulator in Brazil. Each participant drove through three different scenarios consisting of a 6-lane divided highway under traffic volumes and densities corresponding to the HCM LOS A, C, and E levels. The simulator software recorded speed, acceleration, and braking during each scenario. Participants answered a set of questions after driving in each scenario. Two of the questions asked drivers to rate the easiness of driving and of maintaining the desired speed, using a scale from 1 (hardest) to 4 (easiest). In the other two questions, drivers rated the trip quality with respect to traffic conditions (density and speed) and psychological comfort (stress level, etc.), by means of a 10-point scale, where 1 was the worst and 10 the best. The results showed a significant correlation between traffic density and the drivers' perception of trip quality. This study illustrates the major difficulties in trying to incorporate drivers' perception in LOS evaluation, namely obtaining a sample sufficiently large to be statistically significant and the need to use a limited number of scenarios, due to the complexity in generating a set of scenarios covering the full range of densities from LOS A to E.

These challenges can be overcome using computer generated video clips. Previous research has shown the strong potential of using 3-D animated clips generated from traffic simulation [57] and virtual reality [58] as tools in empirical research. Specifically for freeway facilities, several studies were based on animated video clips created using traffic simulators. Obelheiro, Cybis and Ribeiro[32] proposed a method to develop criteria for LOS at toll plazas. Video clips representing a variety of conditions (number of open booths, truck percent, flow rates, etc.) were created using microsimulation (Vissim), from a bird's eye viewpoint, and a website was used to reach a large survey sample with low costs. Thus, many users from different locations and with varying demographic characteristics were able to participate in the survey. Another study [36], also using animated videos generated by microsimulation software (Aimsun), proposed a method to estimate LOS thresholds based on the users' perception of the quality of the trip. The authors decided to use video clips recorded from the driver's viewpoint rather than an aerial viewpoint because the latter perspective does not represent the view automobile occupants experience when traveling on a highway. Study participants watched a set of animated videos and rated the quality of the trip positioning a cursor along a line between "very poor" to "excellent" – a visual analog scale [59]. A convenience sample of university students was used to demonstrate the feasibility of the method.

The literature shows that, while there is a need to include users' perceptions in LOS evaluation, it would be extremely challenging, expensive, and time-consuming to record video clips covering the full range of operational conditions. The recruitment of a sufficiently large number of participants to watch video clips or to participate in a driving simulator-based experiment is in every way as challenging. The use of realistic 3-D animated video clips, generated by microsimulation software, is an appealing way to avoid these difficulties. Using microsimulation, one can easily generate a set of video clips covering the full range of operating conditions for any scenario. A website-based survey can be used to collect data on user ratings and to reach a large number of participants, avoiding the problems associated with small samples. In this chapter, it was discussed how to create realistic video clips using simulation and how to select a vehicle whose driver's view reflects the operational conditions associated with a certain traffic density.

2.3 Microsimulation approach

Traffic microsimulation has the potential to overcome one of the main challenges of field-based in-vehicle video recordings – obtaining a wide range of traffic conditions. Using traffic simulation software allows for the complete control of all factors that affect the traffic stream, making it possible to generate the full range of operating conditions, and requires only modest resources [57]. Furthermore, some modern microsimulation programs are capable of generating realistic animated videos (i.e., computer-generated animations) from the driver's viewpoint, which is likely to elicit more accurate rankings from the study participants than an overhead view [36]. In this study, PTV Vissim Version 11 [60] was selected for the simulation platform: 1) because of its ability to render very realistic traffic animations from driver's field-of-view perspective, both forward and rearward looking; and 2) because of the student's familiarity with the tool. Any other traffic microsimulation software with similar capabilities can also be used, provided that it has been previously calibrated to represent driving behavior and heavy vehicle performance in a sufficiently realistic way [13, 61].

The following elements must be considered carefully for the development of realistic video clips: 1) creation of the simulated road environment, 2) the roadway view from the vehicle's cabin interior, and 3) the choice of the vehicle within the traffic stream for creating the driver's view. These three aspects are discussed in the following sections. For brevity's sake, "videos" will be used in place of "animated videos" henceforth.

2.4 Simulated road environment

To reproduce the driving experience as realistically as possible, the simulation environment must contain elements that help the participant to view the roadway and traffic environment just as drivers would in a real car [57, 58]. Trees and plants placed beyond the shoulder and in the median help to convey the sense of travel speed. Horizontal and vertical curves must be represented by sufficiently small segments to ensure the feeling of a smooth ride. The perception of traveling on an upgrade requires the inclusion of suitable topographic scenery elements, such as hillsides, trees, plants, and grass.

The student's experience suggests that including reverse horizontal curves connected by a short straight segment (as in the example shown in Figure 1) in the network works best to generate a realistic view of the landscape all the way to the horizon, while blocking the view of the parts of the network that are void of scenery elements. Other aspects of the network creation include a median and lanes in the opposite direction. To increase the realism of the video, a new graphic representation of the traffic lanes (with the proper lane width, realistic longitudinal pavement markings, left and right shoulders) replaced the default traffic lanes used by the simulator. Because Vissim creates the simulated roadway connecting (x, y, z) points by straight lines, horizontal and vertical curves along the road alignment were defined in the simulation using closely spaced points, corresponding to a 1-m increment on the x-axis. These short segments ensure that the vehicle movement along the road in the video clip is smooth and thus more realistic.

Figure 1 illustrates a representative image for the network created for this study, showing the segment used for the video clip creation (between points A and B), surrounded by the landscaped area, and the section used for feeding vehicles into the network. Points A and B must be carefully chosen to ensure that the views from the driver's field of vision include only the landscaped area. The section A–B used for the creation of the videos must be long enough for a 1-minute clip. If the study includes the study of the effect of steep inclines on the users' perception of the quality of service, the network should include an upgrade. The starting point for this upgrade should be such that it appears ahead of the car in the videos, so the respondent is able to notice that the car will climb a grade.

Figure 1 – Horizontal and vertical alignments of the simulation network, containing the links for traffic input and the landscaped area where the videos are created

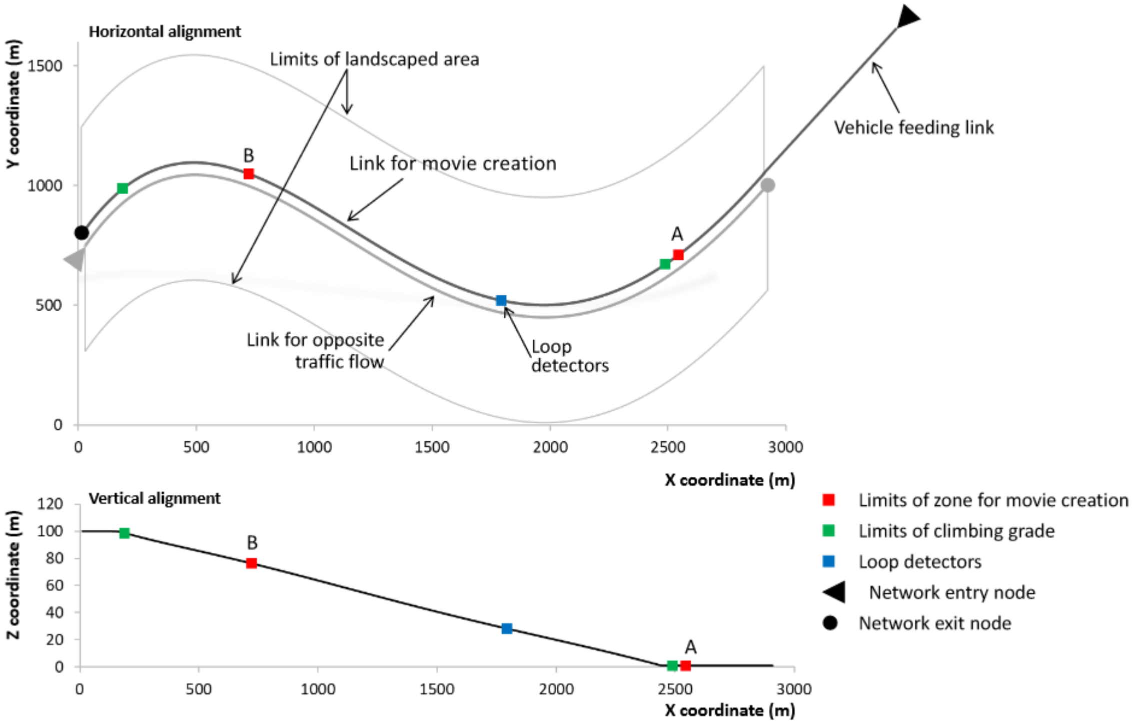
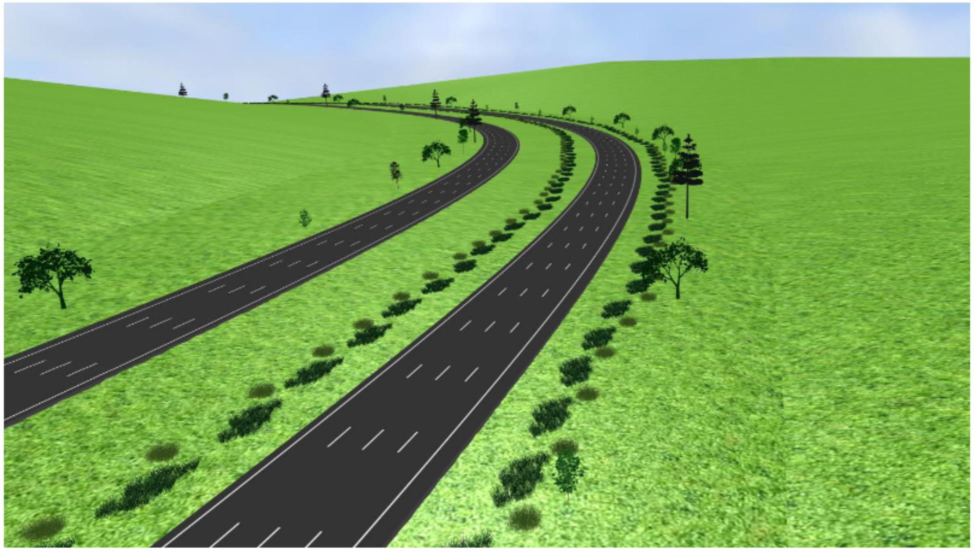


Figure 2 shows a general view of the landscape scenery elements along the roadway segment, such as grass, plants, and trees. In the figure, it is possible to observe both freeway directions, with pavement markings, shoulders, and the median.

Figure 2 – Overview of the landscaped area

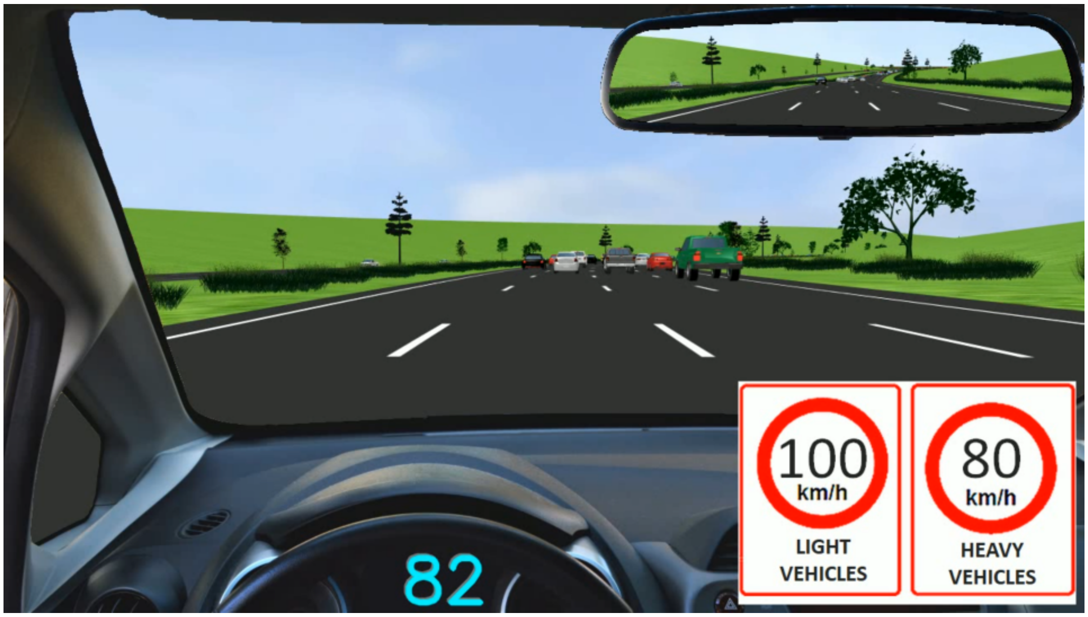


2.5 Roadway view from vehicle interior

It is also important that the generated driver's field-of-view from within the vehicle cabin be as realistic as possible [52]. The dashboard speedometer must be displayed because vehicle speed is frequently checked by a driver [38]. Drivers typically also want to know the posted speed limit of the roadway – this information is displayed in the recorded video (lower right corner) for 10 seconds at the beginning and 10 seconds at the end of the clip, rather than through occasional roadside speed limit signs that could be obscured by other vehicles. The windshield rear-view mirror should also be included since many drivers frequently check traffic approaching from behind. A separate video has to be created for the rear-view mirror and then overlaid with the front video, with time synchronization. The combination of the two videos provides a realistic representation of overtaken and overtaking vehicles around the subject vehicle, as well as the intensity of the traffic flow.

Figure 3 shows the road from the driver's viewpoint, after the front-view and rear-view movies are combined and the instantaneous speed is placed over the dashboard. Both the forward- and the rearward-looking videos should be generated from a point-of-view corresponding to the typical eye height of a passenger car driver, 1.08 m above the pavement [62, p. 3-15].

Figure 3 – Driver's view of the road and surrounding landscaped area through the windshield, with speed indication (in km/h) on the dashboard and a traffic stream presenting a density of 16.8 veh/km/lane



Once the simulation network is properly set up and the desired level of realism for the animated videos is achieved, the next step is the selection of a representative vehicle within the traffic stream for the production of the video clips, as discussed next.

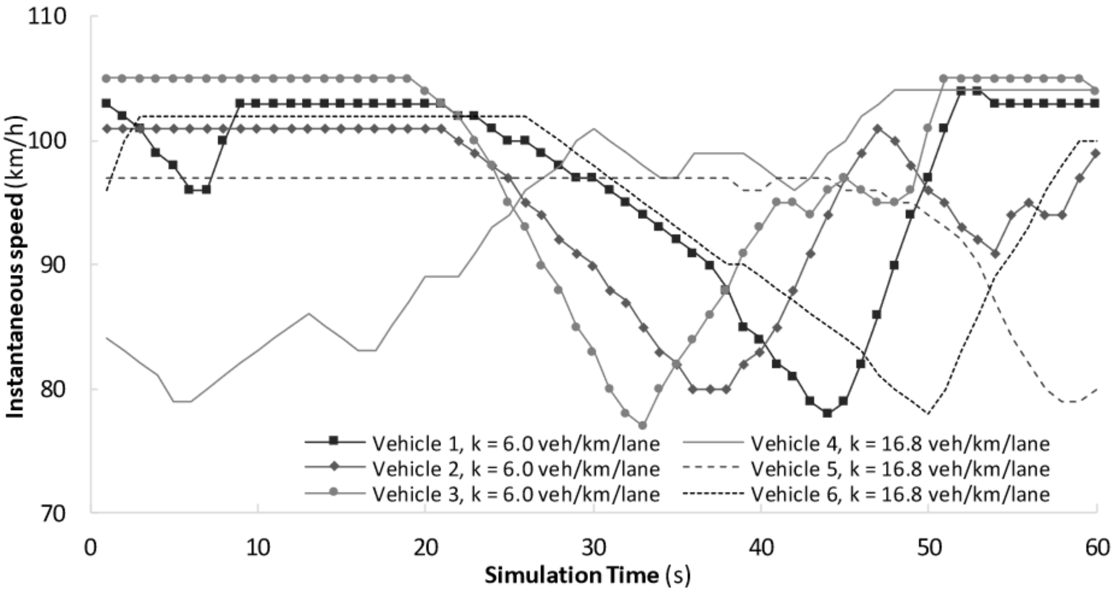
2.6 Vehicle selection in traffic stream

The other significant issue is related to the variance of driver behavior within the simulated traffic stream: which, among all the vehicles in the simulation, should be selected for the creation of the video clip? If the traffic stream is to be observed from overhead, this would not be a problem because the entire stream itself would be the object of interest. A video clip made from a driver's viewpoint, however, requires a careful choice of a vehicle that is really representative of the operational conditions that one wants to represent or, in other words, the density within the driver's field of view.

Previous studies have shown that density, speed variance and percentage of free-flow speed are strongly correlated to the perceived quality of service [20, 36, 52]. Therefore, it is necessary to find a car whose speed and distance to the vehicle ahead vary as little as possible around the expected values for the operational conditions portrayed in the video clip.

Figure 4 illustrates the problem of randomly choosing any vehicle to create the video clip. It shows the variation of second-by-second speed of six vehicles over 60 seconds, for three simulation scenarios representing two different densities (6.0 and 16.8 veh/km/lane), with 0% trucks and 100 km/h speed limit. At a density of 6.0 veh/km/lane (near the threshold between LOS A and B), one would expect very little variation in speed, which should be close to the speed limit (100 km/h), as observed for vehicle 2, whose speed drops to 80 km/h and then goes back to its initial value. Vehicle 1 would not be a good choice: its speed varies considerably during this simulation interval and might not represent, for a person watching that clip, the traffic conditions expected for high levels of service.

Figure 4 – Variation of instantaneous speeds of randomly selected vehicles during 60 s of simulation, for traffic streams with densities of 6 and 16.8 veh/km/lane.



At a density of 16.8 veh/km/lane (close to the boundary between LOS D and E), the average speed should be lower than for the 6.0 veh/km/lane scenario; however, this may not be the case for any individual vehicle, as is observed for vehicle 5, which travels very close to the speed limit for a large percentage of the time interval. Vehicle 6, also representing a heavy traffic scenario, started the simulation with instantaneous speed higher than the speed limit. A participant, watching these three 60-second video clips representing the heavy traffic, might not think that they represent trips within LOS D operational conditions.

To solve the problem of selecting a representative vehicle among all the vehicles in the simulation, we developed an approach based on the relationship between speed (u) and density (k). Initially, simulation output is used to fit a speed-density function, which is then used to choose a vehicle in the simulation that is traveling under the desired conditions.

2.6.1 Using the speed-flow relationship to select a representative vehicle

If the objective is to determine how drivers perceive the quality of service under a range of conditions, then each simulation scenario is a combination of factors (e.g., speed limit, truck percent, grade magnitude, number of lanes etc.) for a range of densities representing the LOS spectrum.

The developed approach uses a speed-density relationship fitted for each scenario to determine the flow rate and average speed that correspond to the desired density level. Since speed-density functions generated from simulation are very dependent on the employed car-following and lane-changing models and often do not match very closely empirical relationships, they have to be fitted to the output of simulation runs for each scenario, otherwise the selected vehicle might not be representative of the desired density conditions.

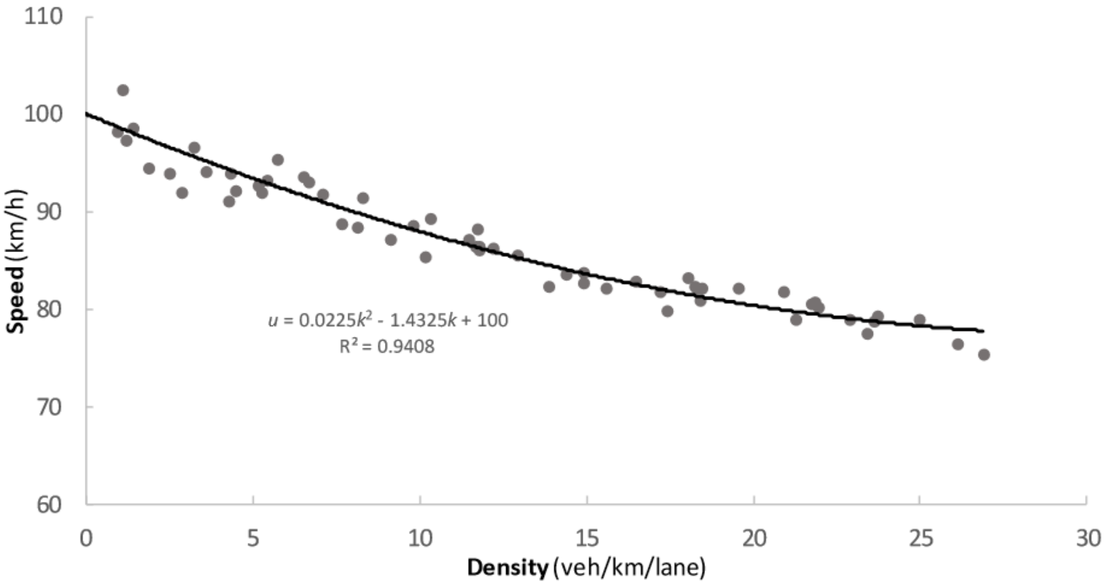
The data for fitting the speed-density function is obtained from loop detectors placed at a representative location in the network (e.g., at the location indicated as “Loop detectors” in Figure 1), where heavy vehicles would be traveling at their crawl speed on the upgrade, should a steep incline be part of the network. A series of simulations, starting with an input flow rate of 100 veh/h/ln, increased by 100 veh/h/ln every 15 minutes until reaching capacity, provides the data to fit the speed-density model. From the 5-min simulation reports, density is calculated using the relationship $k = q/u$ and the model is fitted using regression analysis. Any monotonically decreasing function that has a good fit to the simulation outputs can be used.

Figure 5 illustrates the proposed approach. In this example, a parabolic function was chosen to represent the speed-density function for a simulation scenario with 4% grade, 4 lanes, 100 km/h speed limit and only passenger-cars.

The input flow rate for the simulation is calculated using the relationship $q = ku$, for the desired density level. For instance, considering the scenario presented in Figure 5, to record video clips that represent a density of 7.2 veh/km/ln, the required flow rate can be obtained using the fitted speed-density function:

$$u = 0.0255k^2 - 1.4325k + 100. \tag{2.1}$$

Figure 5 – Example of a speed-density function fitted using simulation results for a segment with 4 lanes, 4% grade, 100 km/h speed limit and 0% trucks.



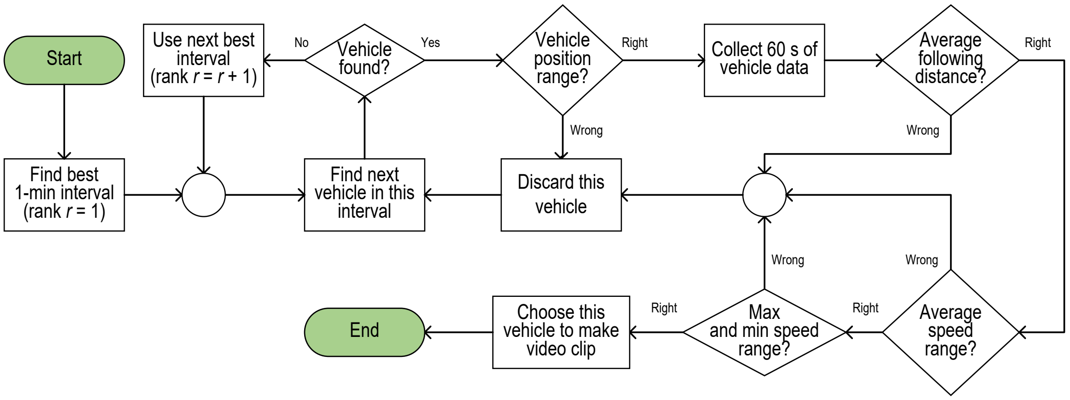
Setting $k = 7.2$ veh/km/ln in Equation 2.1 results in $u = 91.0$ km/h; thus, the required traffic flow rate is $q = 7.2 \times 91.0 = 655$ veh/h/ln. At this flow rate, a representative vehicle would be traveling at an average of 91 km/h and the average spacing to a vehicle in front of it would be the inverse of the density, or approximately 140 m.

The search for a vehicle satisfying such conditions, among all vehicles in the simulation, would be very tedious and time-consuming to perform manually; thus, an automated search is the only reasonable way to facilitate this task.

2.6.2 Automated search for a representative vehicle for video clip creation

To ensure that the selected vehicle represents the target traffic conditions, the search for such vehicle requires checking a series of conditions, as illustrated in the flowchart in Figure 6.

Figure 6 – Flowchart of method to search for a representative vehicle



The search for the suitable one-minute vehicle trip for the video clip creation starts by ranking every one-minute interval in the simulation according to traffic density and truck percentage. Using a sufficiently long simulation of the desired density scenario (i.e., using the input flow rate calculated as shown in the previous section), the loop detectors are used to collect 1-min interval data on the percentage of trucks, flow rate, average speed, and density (computed as $k = q/u$). These 1-min intervals are then ranked using the function:

$$F_{\text{rank}}(i) = |k(i-1) - k_{\text{des}}| + 3 \times |k(i) - k_{\text{des}}| + |k(i+1) - k_{\text{des}}| + 10 \times |p(i) - p_{\text{des}}|, \tag{2.2}$$

where $F_{\text{rank}}(i)$ is the ranking function value for the i-th 1-min simulation interval; $k(i-1)$ is the density in the previous simulation minute; k_{des} is the desired traffic stream density; $k(i)$ is the density (veh/km/ln) in the i-th 1-min simulation period; $k(i+1)$ is the density in the next simulation minute; $p(i)$ is the proportion of trucks, in decimal, in the i-th 1-min simulation period; and p_{des} is the desired proportion of trucks. The coefficients in the ranking function reflect the relative importance of each parameter: the density in the previous or in the next simulation minute is not as important as during the current (i-th) minute and the truck proportion is less important than the density.

To avoid selecting periods when density or percentage of trucks is too different from the desired levels, an arbitrarily large penalty (e.g., 1000) is added to the ranking function value obtained from EquationEq. 2.2 if $|k(i) - k_{\text{des}}| > 0.5\, k_{\text{step}}$ or $|p(i) - p_{\text{des}}| > 0.5\, p_{\text{step}}$. These two checks ensure that traffic density and proportion of trucks during that minute is sufficiently close to the desired levels, assuming that k_{step} is the density step and p_{step} is the truck percentage used to produce the videos.

After this, the 1-minute interval $F_{\text{rank}}(i)$ values for the simulation are ranked, from the smallest (best) to the largest (worst). The notation used to represent the sorted order is r, where $r = 1$ is the rank for the best ranked 1-minute interval, which is the one with the lowest $F_{\text{rank}}(i)$ value.

The next step in the procedure consists of finding a vehicle in the best-ranked minute that satisfies all criteria in the flowchart shown in Figure 6, with regard to its initial position in the network, its speed and average distance to the vehicle traveling ahead during the 60-s simulation interval. The initial position is relevant because, depending on the initial or final position of the vehicle on the segment, the video might include undesirable views of empty space beyond the landscaped region.

The procedure uses second-by-second data for each simulated vehicle, as shown in Table 1, with outputs of simulation second, vehicle number, position from the beginning of the link used for the video creation, speed at the end of the time step, and distance behind its leader vehicle (following distance).

The procedure starts by finding the best ranked 1-min interval. The next step consists of searching the data shown in Table 1 to find the first vehicle whose simulation time (first column

Table 1 – Example of the vehicle tracking data exported by simulation tool

Simulation time (s)	Vehicle number	Position (m)	Speed (km/h)	Following distance (m)
300.05	77	1083.17	80.17	250.00
300.05	79	1436.12	87.80	250.00
300.05	90	562.56	77.72	250.00
300.05	97	736.23	88.87	250.00
301.05	77	1105.44	80.17	250.00
301.05	79	1460.51	87.80	250.00
301.05	90	584.15	77.72	250.00
301.05	97	760.92	88.87	250.00
302.05	77	1127.71	80.17	250.00
...	...	...	...	...

in Table 1) is within that 1-min interval. Upon finding a vehicle that satisfies this condition, the program checks the position of that car in the link. In the network used by the student, only cars whose position is between 500 m and 530 m from the start of the link are desired, because the rear view mirror image must only show the landscaped area and the video must show the beginning of the slope. If the vehicle's position is within this range, then the average speed, maximum speed, minimum speed, and the average following distance are collected for the vehicle for the next 60 simulation seconds; otherwise, this vehicle is discarded and the program looks for the next vehicle in this 1-min interval.

The next restriction to be checked is the average following distance, s, which must be within the interval:

$$\frac{1000}{(k_{\text{des}} + k_{\text{step}})} \leq s \leq \frac{1000}{(k_{\text{des}} - k_{\text{step}})}, \tag{2.3}$$

where s is the average following distance (m) for this vehicle during the 60 s of collected data. This check guarantees that the average spacing will be within the predefined density level during the video clip; if not, this vehicle is discarded and another one is tested.

The last two checks verify that the vehicle's speed is compatible with the desired density range and does not vary too much during the 60-s period of interest. The first condition checks if the average speed is within the speed calculated per the developed speed-density function (as illustrated in Figure 5 and Equation 2.1), using the desired density, plus or minus a small tolerance (e.g., 1 km/h). Thus, it guarantees that the average speed of the vehicle is always lower than the average speed of a vehicle chosen for lower densities. Also, the maximum and minimum instantaneous speeds must not differ by more than a reasonably small range (e.g., 10 km/h) from the average speed during the 60 seconds that the video clip lasts. Failing either of these two checks, the vehicle is discarded and the program searches for the next one within that 1-min interval.

This procedure can be automated using, for instance, Visual Basic for Applications (VBA) in an Excel spreadsheet or Python. Once the search procedure finds a vehicle that passes the four checks, a video clip of these 60 seconds is recorded from the driver's point of view, facing forward (to show the windshield view) and facing backwards to create the image in the rear-view mirror.

We used a Python program to place the speedometer, a speed limit sign at the beginning and end of the video, the forward-looking video and the rearward-looking video in the proper places on an image of a car interior. The final result is shown in Figure 3. The speedometer shows the speed second-by-second, using data from the simulation output (Table 1). The procedure is then repeated until video clips for all scenarios are recorded.

2.7 Results

The best way of demonstrating the effectiveness of the proposed approach is to show two vehicle trajectories, one created from a vehicle selected using this procedure and another one created from a randomly chosen vehicle. Figure 7 shows the instantaneous speed of six vehicles selected using the proposed approach (three in each density level) during a 60-s video clip. For low density (6 veh/km/ln, speeds were constant for almost the whole duration of the video and near the speed limit (100 km/h), as one would expect. Vehicle 1 experiences some interference from other vehicles but the speed drop is within what has been defined as acceptable. Higher density traffic conditions (k = 16.8 veh/km/lane) produce lower vehicle speeds. There is also interference from slower vehicles causing a reduction in speed, as in the first 17 seconds of the video clip of vehicle 4. All vehicles' maximum instantaneous speeds are lower than the speed limit and consistent with the traffic stream density: vehicle speeds are different for different density levels and similar for vehicles traveling within the same density level. Comparing the vehicle trajectories in Figure 7 to those previously shown in Figure 4 makes evident the improvement brought by the proposed approach.

Table 2 compares information gathered from individual vehicle tracking data (shown in Table 1) and traffic stream data (from loop detectors) to show that the operational conditions during the 60-second trips are similar to the point measures. Note that the average density experienced by the selected vehicles in the scenario with density of 16.8 veh/km/ln are higher than those for the 6 veh/km/ln scenario, as expected. The average speeds of vehicles in the scenario with density 6 veh/km/ln are higher than those for the selected vehicles in the 16.8 veh/km/ln density scenario. The density derived from detector data is very close to the desired density, for both scenarios.

A sample video [63], available for download, demonstrates the comparison of trips by randomly selected vehicles and vehicles chosen using the developed procedure. The data shown in Table 2 and Figure 7 were obtained from the simulations used to create the sample video.

Figure 7 – Instantaneous speeds for six vehicles selected using the proposed approach. Simulation conditions for density of 6 veh/km/ln are +1% grade, 4 lanes, 100 km/h and 0% trucks and the same for 16.8 veh/km/ln, except for grade (+4%).

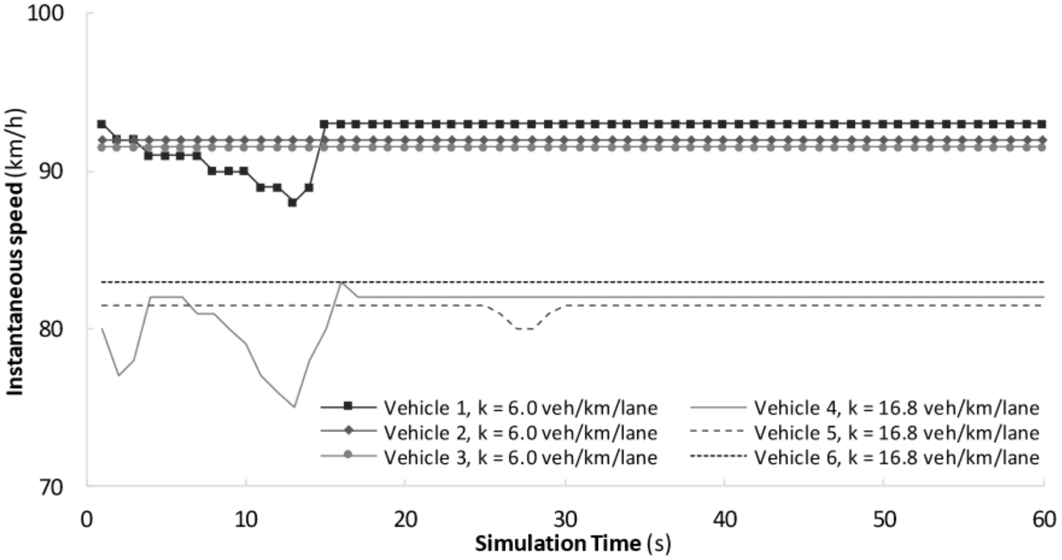


Table 2 – Data for six vehicles selected by the developed procedure: vehicle data are obtained from tracking individual vehicles over the 60-s trip; traffic stream data are obtained from the loop sensors over the same 60 s time interval.

	Vehicle data		Traffic stream data		
Id	Speed (km/h)	Density$^{(1)}$ (veh/km/ln)	Speed (km/h)	Density$^{(2)}$ (veh/km/ln)	Target density (veh/km/ln)
1	92	5.1	85	6.0	6.0
2	92	6.2	94	5.9	6.0
3	92	6.2	95	6.0	6.0
4	82	17.2	81	16.9	16.8
5	82	17.5	84	16.9	16.8
6	82	16.4	83	16.9	16.8

(1) Density = (Average vehicle spacing)$^{-1}$
(2) Density = Flow rate /Average speed during the 60-s interval

We used the proposed approach to create 417 one-minute video clips covering 128 traffic stream scenarios for a follow-on study to collect LOS ratings from a large audience. Study participants visited a website where they watched and rated a subset of these video clips (total of 12) depicting a wide range of traffic densities, from very light traffic flow to capacity flow. The website-video clip strategy allowed for reaching a large number of participants (977 persons) at a very low cost over a short time. The results of the analysis of the obtained 10,228 ratings will be reported in a subsequent chapter.

2.8 Concluding remarks

This chapter describes guidelines and processes that can be used to facilitate the application of traffic microsimulation software to the study of user perception of traffic operational conditions. The topics covered included network creation, field-of-view setup from the vehicle cabin interior, and development of an automated procedure to select representative vehicles in the traffic stream. The end result is animated videos that realistically portray a 60-s vehicle trip under the desired traffic characteristics. The proposed vehicle-selection procedure was developed and tested using a program coded in VBA, but any programming language can be used. The program was used to select vehicles to create 60-s video clips for different combinations of traffic stream density, truck percentage, grade magnitude, number of lanes and speed limit. The proposed approach was found to generate video clips that represent the operating conditions of the traffic stream with greater fidelity than those created using randomly chosen vehicles. Researchers interested in obtaining the VBA and Python source code are encouraged to contact the student.

The flow rate and speed for any given scenario are set to yield a specific density for the overall simulated traffic stream. However, one limitation of the proposed method for selecting the vehicle in the simulated stream is that it only uses information about the car ahead of the candidate vehicle and does not include information about vehicles travelling behind the candidate vehicle or on adjacent lanes. This limitation could be overcome by means of second-by-second evaluation of a “local” density, based on the number of vehicles within the driver’s field of view, to the front, side, and behind the candidate vehicle (through front windshield and rear- and side-view mirrors). Furthermore, additional checks on the selection of the representative vehicle could, for instance, include measures of how many seconds the vehicle is traveling outside the upper and lower limits for instantaneous speed and spacing.

Chapter 3

Estimation of service level thresholds

Research to measure travelers' perceptions of operational traffic conditions is relatively not mature, and methods to obtain such data are still developing. Methods used to collect data on the users' perceptions in previous studies include focus groups, rating recorded in-field video clips, and driver interviews at a rest stop. However, these methods' challenges are: (1) supplying a large set of operational traffic combinations in prerecorded video clips and (2) collecting a sufficiently large sample of responses.

This chapter describes an approach to the problem that addresses the limitations of the previous studies. This approach uses a combination of realistic-looking 3-dimensional traffic stream visualizations and an online survey to reach many people. In the proposed method, participants watch video clips depicting trips along freeways under controlled traffic conditions and rate trips using a visual analog scale ranging between 0 (worst possible trip) and 100 (best possible trip). Using a continuous scale for rating trips avoids any *a priori* assumption about the number of service levels. A k-means clustering and logit analysis determined the number of LOS classifications and corresponding density threshold values[1].

3.1 Introduction

Consistency between travelers' perception of trip quality and the level of service (LOS) classification of traffic operational conditions is important both to the management of the highway system and roadway improvement funding decisions. The seventh edition of Highway Capacity Manual (HCM-7) defines level of service (LOS) as the "quantitative stratification of a perfor-

---

[1] This chapter has been written in an article format and contains its own literature review and conclusions.

mance measure or measures representing quality of service" [7]. Historically, the members of the Transportation Research Board's (TRB) Committee on Highway Capacity and Quality of Service have chosen service measures and LOS thresholds, expecting that these LOS strata would be consistent with the travelers' perception of the quality of the operational conditions. The first time that this approach changed was in the definition of the 2010 HCM [16] LOS criteria for bicycle, pedestrian, and transit modes on urban streets, which was based on user perception research [18].

While research on the correlation between traffic operational conditions and users' perception of trip quality has been increasing, it is still a relatively immature field. One of the greatest challenges in this area is to find a way to collect reliable data in a manner that is not only time and cost efficient, but would also provide a sample sufficiently large to provide statistically significant estimates for LOS limits.

This chapter describes an approach to incorporate travelers' perception into the definition of LOS threshold values for freeways and divided multilane highways. The travelers' perception of the quality of service are obtained using an online survey to reach a large number of people. Participants are invited to watch and rate video clips depicting trips under controlled traffic conditions. The video clips are realistic looking 3D visualizations of traffic streams created using microsimulation software grounded in traffic flow theory. The ratings are then used to estimate the service level boundaries and associated confidence intervals. The text is organized as follows: this introduction is followed by a literature review; next, the proposed approach is described and a case study demonstrates its feasibility; the chapter concludes with a discussion of the results and suggestions for further research.

3.2 Literature review

User's perception of the quality of a trip can be analyzed from the individual's perspective, using a psychological scale [51] to find aspects that correlate to travel satisfaction and subjective well-being [49] to assess individual benefits connected to changes or improvements in travel [50]. A different approach consists of looking for correlations between traffic operational conditions (e.g., delay, travel time) and traveler's ratings to incorporate the users' perception of the quality of the trip in the evaluation of system performance [18, 19, 53, 55, 56, 64]. The second approach was adopted in NCHRP 3-70 to create the criteria for multimodal LOS assessment in the HCM 2010: a wide range of participants were invited to watch video clips recorded from the viewpoint of urban street users (pedestrians, bicycle riders, and car drivers) and to rate the LOS in each video [18]. The set of ratings thus obtained were the base from which LOS criteria were developed. Similar studies have focused on how freeway users subjectively rate the quality of service [19, 35, 52].

The techniques used to collect data on the users' perception include interviews with drivers after traveling along a road segment [21, 35, 45], focus groups [41–44] and video clips recorded in-field [20, 24, 37, 46, 47, 64]. All of these methods share the same limitation, namely the difficulty in recording video clips covering the full spectrum of operational conditions comprised by the service levels and/or obtaining a large sample size.

Since the early 2010s, traffic simulation software has been capable of producing a 3-D visualization of simulated traffic conditions. These visualizations/video clips have been used to collect data on users' perception of the quality of service. Obelheiro, Cybis and Ribeiro[32] used Vissim to record video clips from a bird's eye viewpoint to analyze toll plaza service level from the users' and operators' perspectives. Paiva and Setti[36] used Aimsun to generate video clips from the viewpoint of the driver of a car traveling along a freeway segment that were comparable to those recorded in-field by Washburn and Kirschner[20] and Jensen[37], but with full control of the depicted operational conditions. A convenience sample of university students demonstrated the feasibility of using such video clips for quality of service evaluation and the existence of a strong correlation between users' ratings and the selected performance measure (density).

An important aspect of incorporating users' perception into LOS evaluation is obtaining a large sample of users. Previous studies [18, 20, 37, 41, 42, 44] used relatively small sample sizes, given the inconveniences and costs associated with obtaining larger samples. More recently, online-based surveys have become a useful tool to reach a large number of participants from different locations and diverse sociodemographic characteristics [65–67]. In transportation engineering, video clip based online surveys have already been used [32, 68–70]. Thus, an online survey using realistic computer generated video clips could be a powerful tool to present a wide range of traffic operational conditions to a large number of participants.

The literature presents several different methods to estimate LOS thresholds from user ratings, including piece-wise linear regression [21], ordered probit model [19, 20, 35], fuzzy clustering [24], fuzzy *c*-means data clustering [47] and logistic regression [71]. Logistic regression estimates the probability of a certain event as a function of the relationship between a dependent variable and one or more independent variables. It has been widely used in different areas of transportation engineering to model user behavior: LOS at pedestrian crosswalks [71], pedestrian red-light running [72], discretionary lane changes [73, 74] and mode choice [75, 76]. Given the natural variability among users' perception of the quality of a trip, a previous study [36] has shown that logistic regression is capable to model LOS thresholds based on users' ratings. In the next sections, the student propose a method to estimate freeway LOS thresholds based on user ratings of realistic computer generated video clips produced by traffic microsimulation software.

3.3 The proposed approach

The major steps in the proposed method are: (1) creation of a set of video clips depicting the full range of operating conditions and scenarios; (2) data collection, including the creation of the web-based survey form to collect the data and the recruitment of participants; and (3) level-of-service thresholds calculation.

The next sections present each step in the proposed method. A case study illustrates in greater detail how this approach was used to calculate LOS limits for freeways in Brazil.

3.3.1 Video clip creation

Each video clip (i.e., traffic simulation animation) shows a trip made under predetermined conditions. The video is recorded from the driver's point of view. Figure 8 illustrates what a participant sees in a typical video clip: the view through the windshield, the rear-view mirror, and the current speed (in km/h). A speed limit sign also appears for 10 seconds at the beginning and end of the video clip.

Figure 8 – Typical video clip frame showing a view of road traffic ahead and behind; speedometer is updated every second and speed limit is superimposed on the lower right side of frame.

Because the quality of the results hinges on the realism and accuracy of the video clips, special care must be taken for their creation. The process to create the video clips requires three preliminary steps: the calibration of the traffic simulator to represent vehicles and drivers characteristics for the desired local conditions; the creation of the roadway environment in a traffic simulator; and the definition of the scenarios to represent the desired range of operating conditions. Once all scenarios are simulated, the next step is the creation of the video clips that will be used in the data collection. More detail about this process can be found in Piva, Setti and Washburn[48].

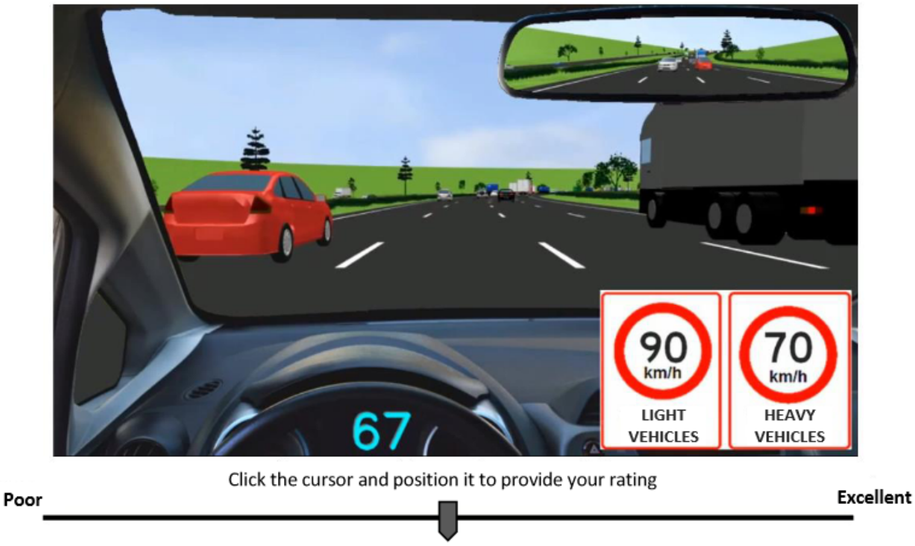
3.3.2 Data collection

The process selected for data collection is a web-based survey form, to enable the greatest number of people to participate with the lowest cost. This brings a series of advantages, besides the ease of access and cost: (1) participants' responses are automatically stored in a database, facilitating data processing and analysis; (2) a device-responsive website enables participation using computers, tablets, or smartphones, which increases the number of potential respondents; (3) the website programming can include the collection of visitor behavior data that assists in identifying spurious responses; (4) as the survey is a self-administered questionnaire, there is no need for supervision during the data collection, other than the regular administration of the web server hosting the website; and (5) the simplicity of the whole process is such that anyone familiarized with web polls would be able to respond to the survey.

Participants can be recruited by sending short invitations through email, messaging apps, newsletters, and social and professional networks, as in any web-based survey.

The website contains, along with a sociodemographic questionnaire to characterize participants, a short training video explaining how to rate the trip quality and what to watch in the video clips. Participants then watch a sequence of video clips and rate the quality of the trips using a visual analog scale (VAS) [59], moving a slider over a scale that varies from "poor" to "excellent". Figure 9 illustrates a page on the website with the video clip and the VAS, with the cursor used to rate the trip. Responses obtained using a VAS are not restricted to a predetermined number of choices [36, 59]. The position of the cursor on the rating scale translates into a score between 0 (poor) and 100 (excellent).

Figure 9 – The driver's view of the road and the visual analog scale used to rate the trip. The participants move the cursor to the point between "poor" and "excellent" that represents their perception of the trip quality.


Click the cursor and position it to provide your rating

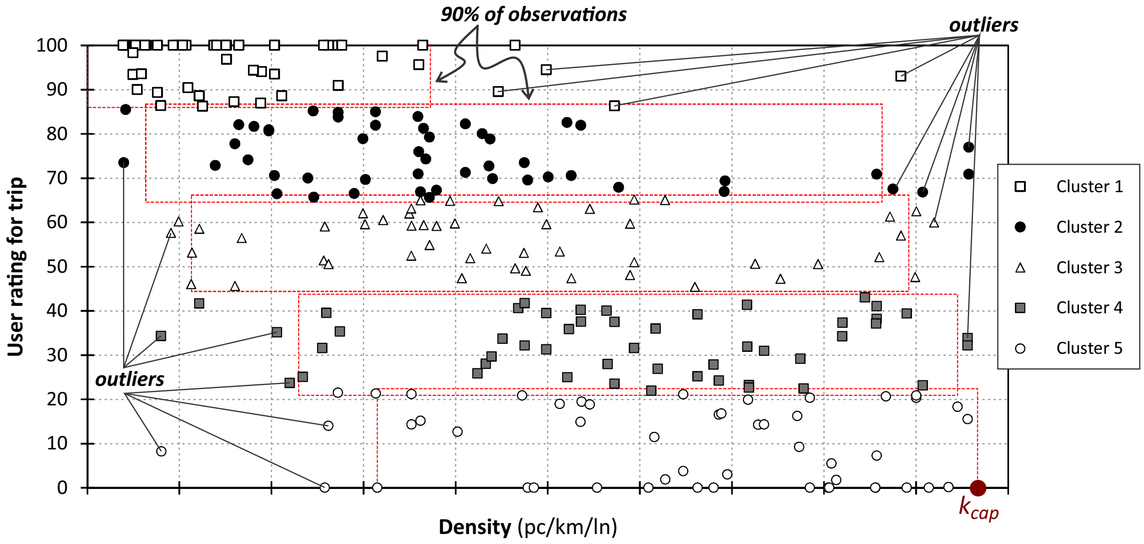
Besides the participants' anonymized identifiers and their sociodemographic characterization, the following information is recorded for each video clip watched by a participant: the date and time the participant accessed the video clip page; the video clip watched; and the rating (a value between 0 and 100).

3.3.3 Level-of-service threshold calculation

The objective of this step is to find the values of the service measure that represent the limit between consecutive LOS categories using the ratings given by the participants.

The data collection will supply a set of scores with their associated value of density. As demonstrated in a previous study [36], it is possible to consider each rating an independent observation. Initially, these observations are discretized into clusters corresponding to the desired number of service levels for uncongested flow (usually five, from A to E) using, for instance, the k-means method, as illustrated in Figure 10.

Figure 10 – Example of discretized users' ratings assuming five levels of service. The ratings outside the boxes are the outliers for each cluster and not included in the set used for establishing the LOS boundaries.



Because users do not always rate trip conditions in a rational way, the sample will include a few good ratings for high density conditions and, likewise, some poor ratings for light traffic. Outliers, such as these, can introduce errors in the calculation of the LOS boundaries and should be excluded from the sample. Figure 10 illustrates the proposed approach to exclude outliers, which consists of deleting a predefined percentage of outliers from the sample. Only the remaining ratings will be used for the LOS boundary definition. Assuming that a fraction of the ratings are outliers (e.g., $\delta_o = 0.1$), the ratings outside the boxes in Figure 10 are deleted from the sample. Assuming $n = 5$ service levels, all ratings for densities above the 90^{th} percentile of the density values distribution obtained for cluster 1 (LOS A) are eliminated from the sample. The procedure is repeated for the next clusters. For each cluster, the percentile boundaries ρ_ℓ (lower boundary) and ρ_u (upper boundary) are calculated by:

$$\rho_\ell = 100\frac{(i-1)\times\delta_o}{(n-1)} \quad \text{and} \tag{3.1}$$

$$\rho_u = 100\left[(1-\delta_o) + \frac{(i-1)\times\delta_o}{(n-1)}\right], \tag{3.2}$$

where i is the cluster number ($i = 1$ for LOS A, ... , $i = 5$ for LOS E). All ratings for densities lower or greater than the densities corresponding to the lower and upper percentile boundaries are discarded, as exemplified in Figure 10.

The probability of users rating a trip made under density k as LOS $\ell = \mathrm{A}$, $P[\ell = \mathrm{A}](k)$, is estimated by a binary logit function and the probability of users rating that same trip as a LOS worse than A is the complementary probability:

$$P[\ell = \mathrm{A}](k) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot k)}} \qquad \text{and} \qquad (3.3)$$

$$P[\ell \neq \mathrm{A}](k) = 1 - P[\ell = \mathrm{A}](k) = 1 - \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot k)}}, \qquad (3.4)$$

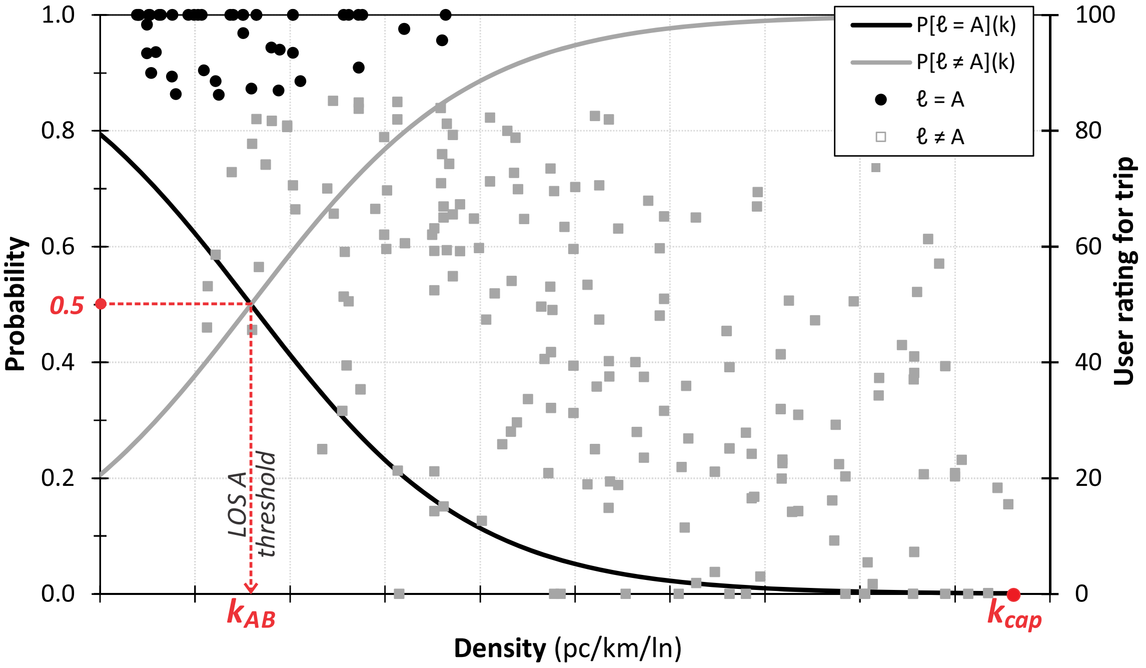
where β_0 and β_1 are calibration constants estimated by logistic regression from the users' ratings, assuming that cluster 1 data in Figure 10 correspond to users rating the trip as level of service $\ell = \mathrm{A}$, and data in all other clusters correspond to $\ell \neq \mathrm{A}$.

The threshold between LOS A and LOS B would be the density $k = k_{AB}$ where

$$P[\ell = \mathrm{A}](k_{AB}) = P[\ell \neq \mathrm{A}](k_{AB}), \qquad (3.5)$$

i.e., the density level at which the probability of users rating the trip as LOS A is equal to the probability of users rating the trip not being in LOS A. Figure 11 illustrates the procedure, showing that the threshold between LOS A and LOS B corresponds to $P[\ell = \mathrm{A}](k_{AB}) = 0.5$.

Figure 11 – The level of service threshold is the density value k_{AB} that makes the probability of users rating the trip in the desired level of service or better, $P[\ell = \mathrm{A}](k_{AB})$, to be equal to its complementary probability, $P[\ell \neq \mathrm{A}](k_{AB})$. The black dots represent trips rated as LOS A (cluster 1) and the grey squares represent trips rated as LOS lower than A (clusters 2 to 5).



For the calculation of the next threshold, between LOS B and C, the logit function represents the probability of users rating the trip as LOS A or B and the data used for its calibration corresponds to observations in clusters 1 and 2. Clusters 3, 4 and 5 are used to estimate the probability of users rating the trip as a LOS worse than B.

$$P[\ell = \text{A or B}](k) = \frac{1}{1 + e^{-(\beta_2 + \beta_3 \cdot k)}} \qquad \text{and} \tag{3.6}$$

$$P[\ell \neq \text{A or B}](k) = 1 - P[\ell = \text{A or B}](k) = 1 - \frac{1}{1 + e^{-(\beta_2 + \beta_3 \cdot k)}}, \tag{3.7}$$

The threshold between LOSs B and C is the density k_{BC}, for which:

$$P[\ell = \text{A or B}](k_{BC}) = P[\ell \neq \text{A or B}](k_{BC}). \tag{3.8}$$

The subsequent thresholds k_{CD} and k_{DE} are calculated in a similar manner.

The proposed procedure can be used for any number of levels of service. The desired number of levels of service defines the number of clusters by which the user ratings will be divided into and the calibrated logit functions reflect this corresponding number of LOSs.

3.4 Case study: LOS thresholds for freeways in Brazil

A case study to identify the LOS thresholds for freeways in Brazil was undertaken to demonstrate the application of the proposed method.

3.4.1 Video clips creation

The creation of video clips from the driver's viewpoint followed the guidelines proposed by Piva, Setti and Washburn[48], using a version of VISSIM that had been recalibrated to better represent the behavior of local drivers [61] and the performance of Brazilian trucks [13]. A typical rural freeway segment, comprising a straight section bounded upstream and downstream by two curves was used to create the video clips. Its horizontal alignment is shown in Figure 12(a) and Figure 12(b) is a bird's-eye view of the network, showing the limits of the landscaped area.

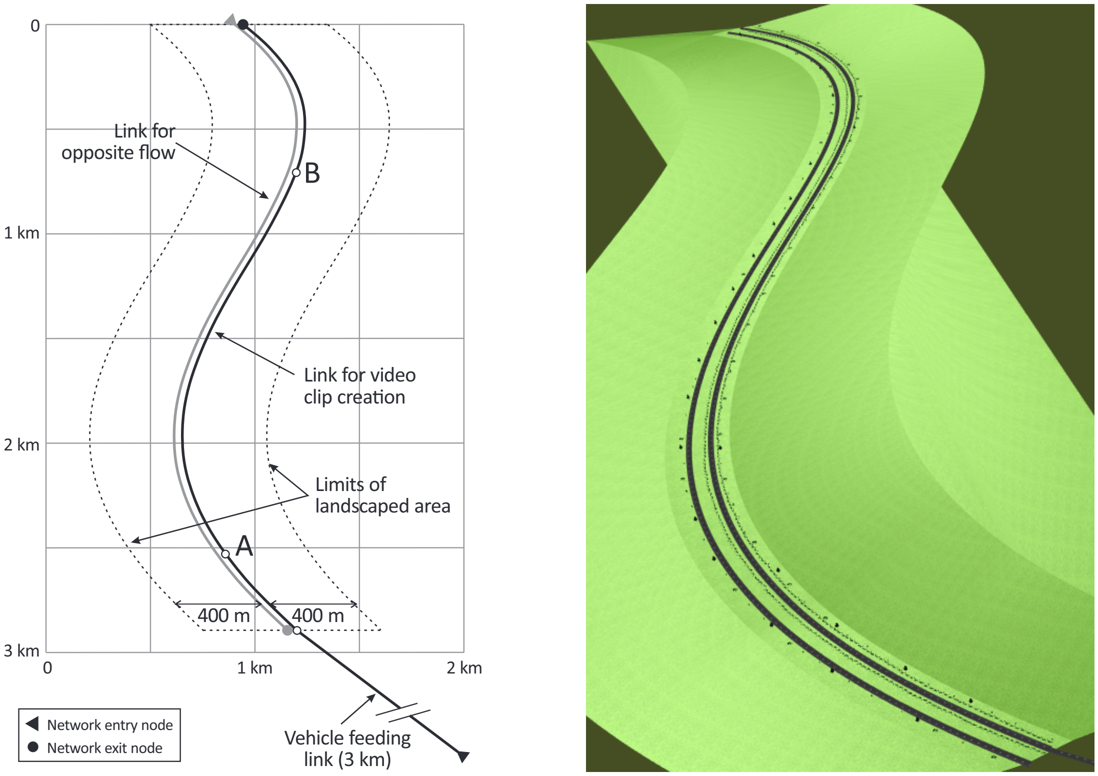
The video clips show a view through the windshield, the rear-view mirror and the dashboard with a speedometer displaying the current speed, as shown in Figure 9. The front view and the rear view reflect the degree of congestion around the car and the speedometer provides information about the travel speed.

The traffic stream density used in the video clip creation process was obtained from loop detectors placed 1.3 km downstream from point A. Average speed and flow rate data were collected at 1-minute intervals and density was estimated using the traffic flow fundamental relationship, $k = q/u$.

3.4.2 Definition of simulation scenarios: design of the experiment

Five factors were selected to characterize the traffic stream and roadway conditions in the video clips: traffic stream density, truck percentage, posted speed limit, grade magnitude and

Figure 12 – The simulation network used to generate the animated video clips: (a) horizontal alignment; and (b) bird's-eye view showing the limits of the surrounding landscaped area.



(a) (b)

number of lanes. Several studies have shown that density is the most important traffic stream characteristic affecting users' perception of trip quality in freeways [7, 35, 43]; therefore, the first factor (*A*) was density, with 16 levels. The next two factors were truck percentage (*B*) and speed limit (*C*), with four levels each; the other two factors were grade magnitude (*D*) and number of lanes (*E*), with 2 levels each. Table 3 shows the factors, number of levels, increments, and values defined to create the video clips.

Table 3 – Traffic stream and roadway conditions selected to create the video clips

Factor		Levels	Values
A	Traffic density	16	1.2, 2.4, 3.6, ..., 19.2 veh/km/lane
B	Truck percent	4	0%, 10%, 20% and 30%
C	Speed limit	4	90, 100, 110 and 120 km/h
D	Grade magnitude	2	1% and 4%
E	Number of traffic lanes	2	3 and 4 lanes

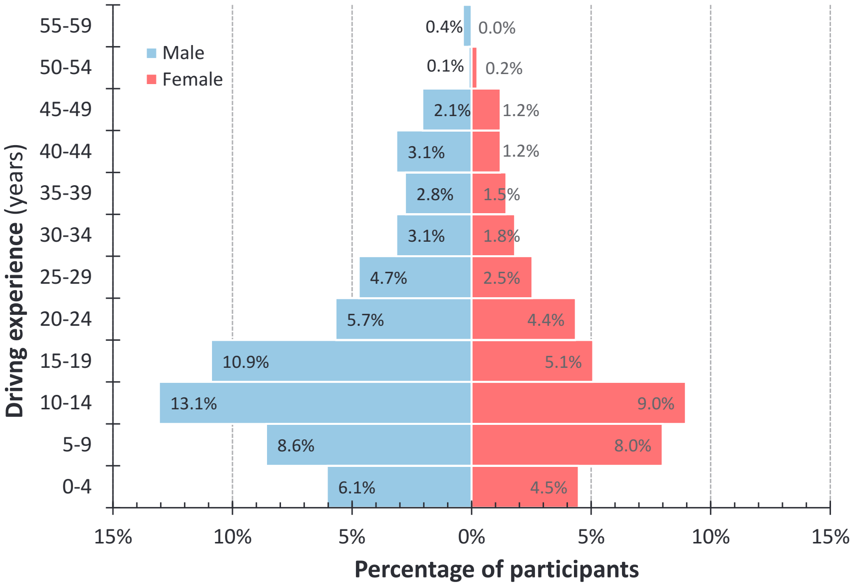
A full factorial design would result in $16 \times 4 \times 4 \times 2 \times 2 = 1024$ video clips, which would require not only a large amount of effort, time, and cost, but also a very large number of participants. Thus, a fractional factorial design (FFD) was used to simplify the experimental design [77]. Such a design allows the main effects and second-order interactions of factors to be evaluated. Higher-order interactions are unlikely to be statistically significant. The total number of main effects is 23 $[15\text{ (from factor }A)+3\ (B)+3\ (C)+1\ (D)+1\ (E)]$. For second-order interactions, the effects of interaction between factors A and B were considered the most relevant ($15 \times 3 = 45$ effects), resulting in 68 effects of interest. The smallest fractional factorial design that would satisfy this is $1/8 \times 1024 = 128$ scenarios. A detailed discussion of the FFD is presented in section 4.3.1.

After the 128 scenarios were defined, at least three repetitions of each one were recorded, resulting in a total of 417 one-minute video clips. Some densities had more than three repetitions to ensure that the number of respondent ratings were similar over the full range of LOS.

3.4.3 Data collection

All ethical guidelines for research with human subjects were followed, and the research proposal was approved by the Brazilian Committee for Ethics in Research on Human Beings (CEP-EACH-USP protocol number: 2034830). The data collection was based on a website hosted in one of the University of São Paulo servers. Messages sent through social networks (Facebook, Instagram, Whatsapp) and mailing lists invited people to visit the website and participate in the survey. This approach proved to be highly effective and provided a convenience sample of 977 respondents who watched and rated 10,228 video clips. Figure 13 shows the respondent sample composition in terms of years of driving experience and sex.

Figure 13 – Sex and years of driving experience for survey respondents



The website had four main sections. The first one is the informed consent form to which respondents must agree to participate in the data collection. The second section consists of a sociodemographic questionnaire. In the third section, the participants watch a short training video that explains how to use the VAS cursor to rate the trip. The fourth section contained a sequence of 12 video clips to be rated by the participant. YouTube was used to host and play the video clips in the questionnaire.

The full range of traffic density scenarios was divided into four sub-ranges: 1.2–5.4, 5.4–10.2, 10.2–15.0 and 15.0–19.2 veh/km/ln. Each participant watched at least one video clip from each sub-range, to ensure that all participants rated the full spectrum of operating conditions.

After the respondent moved the VAS cursor to rate the trip, the cursor position was converted to a scale from 0 to 100 and stored in the web server database. The respondents could not proceed to the next page until they moved the VAS cursor from its initial position, at the center of the scale.

3.4.4 Data checks

To ensure that the ratings were representative of the users' perception of the trip quality, a series of checks were applied to the participants' responses. The first filter was the time spent watching the video clip: a minimum viewing time of 12 seconds was required to accept the rating, otherwise the rating was discarded. The minimum "duration" was selected based on subjective tests employing a convenience sample of university students, and is referenced to the time between the page loading finished and the participant clicked on the "next page" button. The application of this filter reduced the sample from 10228 ratings by 977 respondents to 9736 ratings by 921 participants.

The second filter eliminated respondents who are not drivers or do not drive on highways or never travel on freeways. This filter removed another 2545 ratings by 245 participants from the sample. The last filter checked the difference of ratings given to the trips made under the lowest and highest densities by the participant: if this difference was less than 50, all ratings by this participant were discarded because of their inability to reasonably distinguish between these two conditions. A total of 960 ratings by 122 participants was removed by this filter; at the end of this step, the sample consisted of 6231 ratings by 554 drivers. Based on a previous study, each rating was considered an independent observation [36].

Once density values were converted from veh/km/ln to pc/km/ln using the HCM model and parameters [7], the 16 density levels resulted into 24 bins of 1.2 pc/km/ln width, from 0.6 to 29.2 pc/km/ln. The number of ratings for these bins ranged from 182 to 376, averaging 259.6 ratings/bin. The correlation between density (in pc/km/ln) and rating was found to be significant (Spearman's $\rho = -0.532$, $N = 6231$); therefore, density can be used as a performance measure.

3.4.5 Level-of-service boundary estimation

Any number of level-of-service categories can be used because the data consist of ratings between 0 and 100. The next step in the method is to discretize the respondents' ratings into levels of service. This can be achieved using a variety of procedures; in this study, k-means clustering was applied. Considering that the HCM uses five LOS categories for uncongested traffic, the 6231 ratings were initially clustered into five groups. Outliers (traveler ratings that are highly incompatible with the operating conditions) were eliminated using the strategy explained in section 3.3.3. Figure 14 displays the ratings in each cluster and Table 4 shows the limits for valid ratings and the corresponding density ranges for each of the five clusters in Figure 14: 90% of the ratings for each cluster were within the lower and upper density values. Sample size refers to the total number of ratings and the number of ratings within the valid density range for each cluster.

Figure 14 – Respondents' ratings discretized into five clusters, with outliers removed.

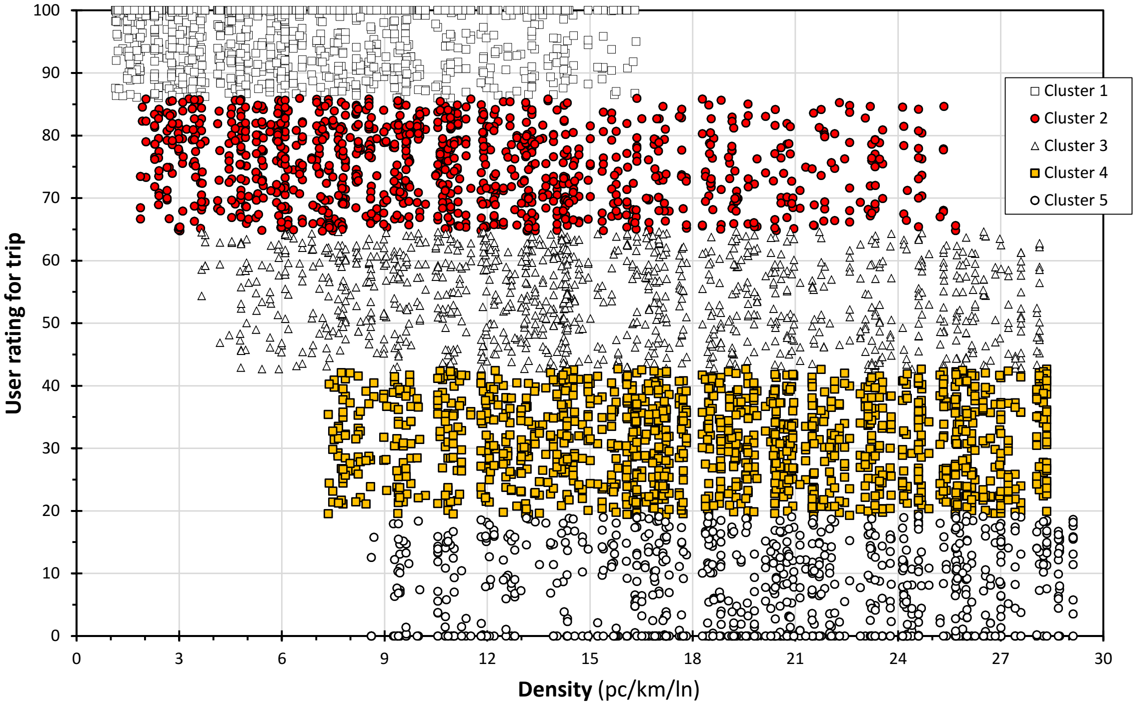


Table 4 – Sample size and limits for valid ratings and their corresponding density ranges for each of the five clusters in Figure 14

	Limits for valid data					
	Ratings		Density (pc/km/ln)		Sample size	
Cluster	Lower	Upper	Lower	Upper	Total	Valid
1	85.99	100.00	1.14*	16.36	1524	1371
2	64.72	85.98	1.87	25.68	1142	1027
3	42.69	64.71	3.65	28.13	1158	1048
4	19.27	42.68	7.33	28.35	1421	1256
5	0.00	19.26	8.62	29.12*	986	886

* These are the lowest and highest density values in the video clip set

Four logit functions were fitted to these data using the procedure detailed in section 3.3.3 (Eqs. 3.3 and 3.6), using log-likelihood. The LOS boundaries k_{AB}, k_{BC}, k_{CD} and k_{DE} of the five LOS are defined by the density k that results in a probability equal to 0.5. Table 16, Appendix (A) presents the coefficients found by SPSS for the fitted logit functions, including Nagelkerke's ρ^2. All coefficients are significant at $\alpha = 5\%$. The results are summarized in Figure 15. The dotted lines in Figure 15 are the lower and upper limits for the 95% confidence interval for the LOS threshold estimates.

The boundary between levels of service D and E, k_{DE}, lies very close to the density at capacity (in this case study, around 29 pc/km/ln), suggesting that, at least for this particular sample, users have difficulty distinguishing between LOS E and F (congestion).

3.4.6 Discussion of results

Table 5 compares the LOS thresholds found in the literature for five service levels (A to E). The HCM thresholds were arbitrarily defined by the members of TRB's Highway Capacity and Quality of Service Committee, based on engineering judgment, and are the "standard" values for LOS analysis, while the other lines in Table 5 refer to values based on the users' perception of the quality of the trip: [19] used 24 video clips taped from overpasses and 193 participants; Washburn and Kirschner[20] used 13 video clips recorded from the driver's viewpoint and 126 participants; and this study was based on 417 computer-animated video clips from the driver's perspective and 554 respondents.

Figure 15 – Logistic functions used to estimate level of service boundaries considering five levels of service (A to E) and their 95% confidence interval.

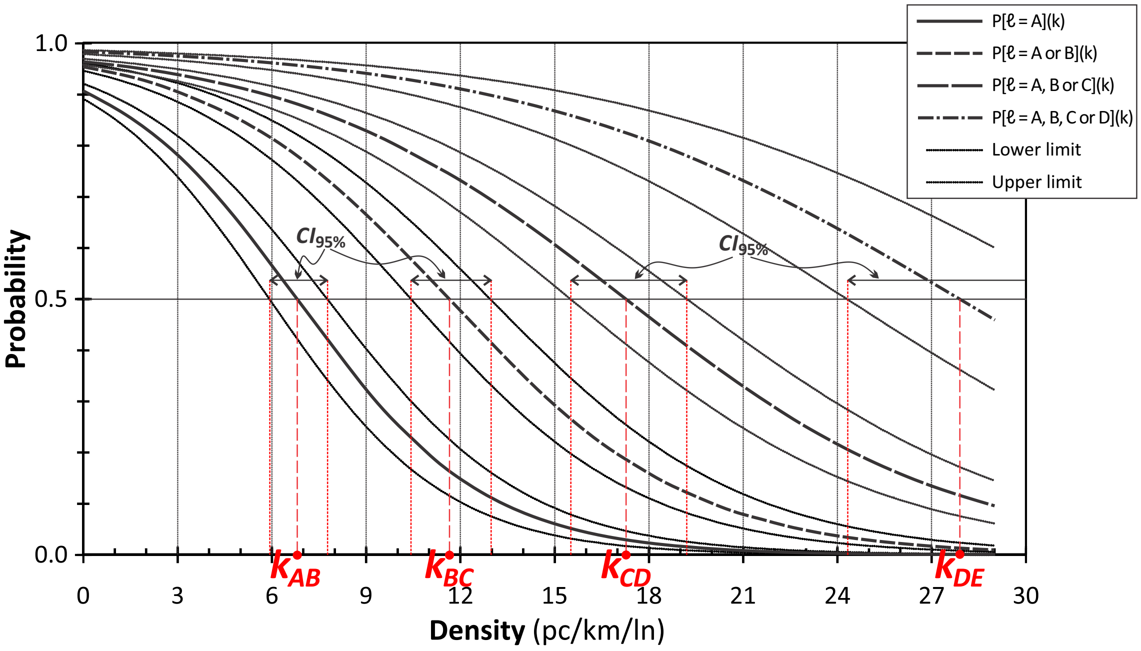


Table 5 – Comparison of LOS thresholds, considering five service levels

	LOS boundaries (pc/km/ln)			
Source	k_{AB}	k_{BC}	k_{CD}	k_{DE}
HCM	6.8	11.2	16.2	21.7
Choocharukul, Sinha and Mannering[19]	4.3	13.0	21.1	30.4*
Washburn and Kirschner[20]	3.7	8.7	13.7	18.0
This study	6.8	11.6	17.3	27.9
This study $CI_{95\%}$	5.9–7.8	10.4–12.9	15.5–19.2	24.3–32.0*

*Greater than $k_{\text{cap}} = 29$ pc/km/ln, the upper limit of LOS E

As expected, the LOS boundaries vary among studies – after all, participants have their own set of personal characteristics that affect how they perceive the quality of a trip. Washburn and Kirschner[20] reported the lowest values for all thresholds, significantly below the HCM values, and attributed this to drivers having different expectations when driving on rural freeways, compared to driving on urban freeways. Choocharukul, Sinha and Mannering[19] found values for k_{BC}, k_{CD} and k_{DE} that are higher than the ones in the HCM and that k_{DE} lies beyond capacity. The LOS boundaries found in this study are virtually the same as the HCM's, with the exception of k_{DE}, which is much higher for the participants in this study sitting very close to capacity. The higher value found for k_{DE} suggests that, at least for this study's participants, HCM's LOS D and E are very similar and that LOS E comprises a narrow range of densities very close to capacity.

3.4.7 How many service levels can users differentiate?

Previous studies have suggested that drivers cannot perceive five levels of service. In a footnote, Choocharukul, Sinha and Mannering[19] report that, when using cluster analysis (as in this study), "participants seemed to differentiate only three levels of freeway traffic conditions". Papadimitriou, Mylona and Golias[21], using piecewise linear regression, found that "drivers perceive no more than two or three categories of traffic conditions".

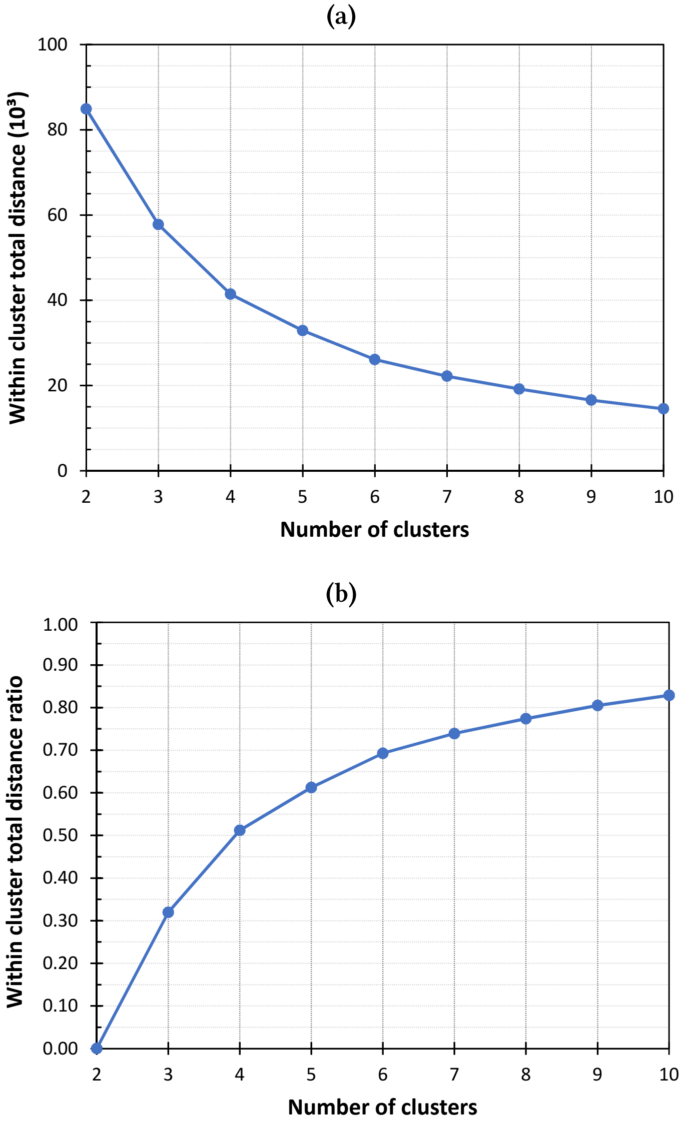
One of the ways to estimate how many LOS levels are perceived by drivers is by determining the "ideal" number of clusters. The elbow method is a heuristic to find the best number of clusters to classify data, based on a scree plot that shows the value of a distortion score as a function of the number of clusters: the optimum number of clusters is the one after which an additional cluster adds little reduction to the distortion score [78, 79].

The full sample ($N = 6231$) was used in the analysis. The number of clusters varied between 2 and 10 and the distortion score used was the within cluster total distance (WCTD), defined as

$$\text{WCTD} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} ||x_{ij} - \bar{x}_i||, \tag{3.9}$$

where x_{ij} is the j-th rating in the i-th cluster; $\bar{x}_i$ is the i-th cluster's center; k is the number of clusters; and n_i is the number of ratings in the i-th cluster. Figure 16(a) shows that while there is not a clear "elbow" in the scree plot, the reduction in WCTD markedly decreases after 4 clusters. This is confirmed by Figure 16(b), which shows the variation of the WCTD ratio, defined as $(\text{WCTD}_2 - \text{WCTD}_i)/\text{WCTD}_2$, as a function of the number of clusters: after 4 clusters, the reduction for increasing the number of clusters from 4 to 5 is significantly less than the reduction observed for the change from 3 to 4 clusters.

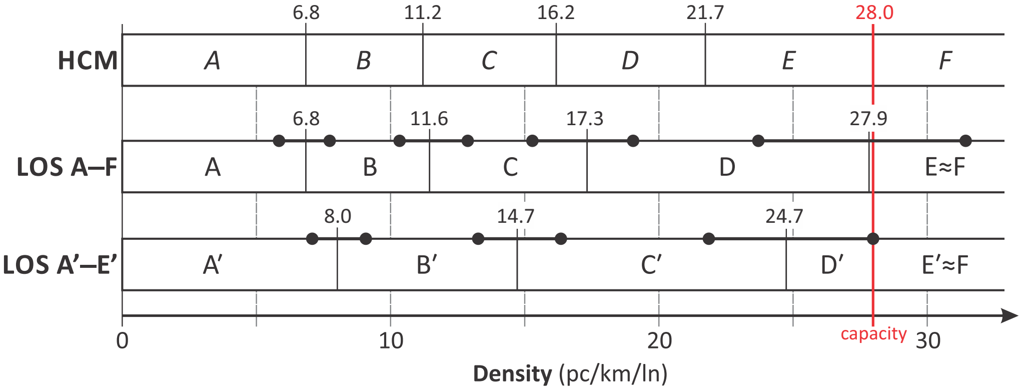
Figure 16 – Number of LOS perceived by drivers in the sample: curves tend to flatten out for more than four clusters.



Assuming that drivers would perceive only four levels of service (A', B', C' and D'), three logit functions were fitted to the data using the proposed approach. All coefficients are significant at $\alpha = 5\%$ (Table 17, Appendix A) and the resulting LOS boundaries are $k_{A'B'} = 8.0$ pc/km/ln ($CI_{95\%} = [7.1, 9.1]$), $k_{B'C'} = 14.7$ pc/km/ln ($CI_{95\%} = [13.3, 16.3]$) and $k_{C'D'} = 24.7$ pc/km/ln ($CI_{95\%} = [21.8, 28.0]$).

Figure 17 compares the current HCM LOS boundaries to the boundaries estimated using the proposed approach, for both five and four LOS levels. The results suggest that users in our sample cannot very precisely perceive five service levels (A to E), as the threshold between LOS D and E (27.9 pc/ln/km) corresponds the HCM capacity (28 pc/ln/km). Assuming five LOS levels, there is a fairly good correspondence between the HCM boundaries and the ones found with our approach for LOS A, B and C. The HCM's LOS D and E, however, were perceived as the same by our sample of users, unless the lower end of the 95% confidence interval for k_{DE} is used. Even in this case, the LOS E range of density is narrower than the one used in the HCM.

Figure 17 – Comparison of estimated LOS thresholds and their 95% confidence interval using five or four LOS.



Comparing the HCM's LOS boundaries to those obtained by the proposed approach assuming four LOS levels (A' to D'), one can see that LOS A' comprises HCM's LOS A and LOS B's lower end, while LOS B' includes the upper range of HCM's LOS B and about half of LOS C. The other half of HCM's LOS C, all of LOS D and part of LOS E belong to LOS C'. LOS D' corresponds to a narrow range of densities close to capacity. These limits suggest that, when using four instead of five LOS levels, there are changes in all LOS thresholds, rather than a subdivision of one LOS into two service levels. The results summarized in Figure 17 indicate that the participants in this study are able to reasonably perceive only four service levels for under-saturated traffic flow and that a fifth service level would lie beyond capacity, within congested flow.

While it would be reasonable to make a case for using only four service levels to report traffic conditions to travelers, Figure 14 shows that traveler ratings seem to be uniformly distributed within each cluster rather than follow a normal distribution with a reasonably small standard deviation. The dispersion in traveler ratings could be the reason for the differences in LOS boundaries shown in Figure 17 and the wide range of densities associated with LOSs D and C'. From the point of view of a transportation agency, a continuous scale of values for the service measures, perhaps combined with other performance measures, would facilitate monitoring and the prioritization of improvement projects.

3.5 Summary and conclusions

This chapter presents a method to estimate LOS boundaries based on drivers' perception of the quality of a trip, using a combination of video clips generated using traffic simulation software and a web survey. This approach proved to be an effective and cost-efficient way to obtain the required data. The 6231 ratings have shown that, even if other aspects might also influence the drivers' perception of LOS, there is a strong correlation between density (volume/average speed) and the perceived quality of service. The results demonstrated that logistic regression is a valid and efficient way to estimate LOS boundaries and their associated confidence intervals.

The case study results indicate that the HCM freeway LOS boundaries are within the confidence intervals estimated using the proposed approach, with the exception of the threshold between LOS D and E that is higher than the one used in the HCM. This suggests that, at least for the participants in this study, HCM's LOS D and E are perceived to be very similar, including densities that are closer to capacity. Based on the results of the elbow method, this study has found that participants were clearly able to distinguish four service levels.

From a practical point-of-view, the proposed approach can be easily adapted to collect data and estimate LOS thresholds based on the users' perception of the quality of service, in an inexpensive and effective fashion.

Chapter

4

Factors affecting trip quality perception

4.1 Introduction

This chapter aims to assess the importance of independent variables on participants' perception of trip quality on freeways. Concerning data collection, participants answered a sociodemographic questionnaire including 11 questions and subsequently evaluated a sequence of video clips depicting a range of traffic conditions, as explained in Chapter 3. These video clips were randomly chosen from a total of 384 produced in a traffic simulator. In total, 830 respondents watched 8,639 video clips. To assess the trip quality presented in each video clip, participants indicated their opinion on a continuous analog scale, ranging from "very poor" to "excellent". The ratings were used as a dependent variable to analyze the effect of 16 factors. The importance of each factor was analyzed by bivariate correlation, multiple linear regression and decision trees. This chapter shows the results of these analyses.

4.2 Literature review

User perception of service quality is an aspect that ideally should be considered in the evaluation of transport systems [7, cap. 5]. Research related to inclusion of users' perceptions on mode choice behavior [26, 27, 80], pedestrian level of service [30, 31], serviceability of road pavements [29] quality of service at signalized intersections [24, 25] and urban street segments [22, 23] can be found in the literature. In these studies, a questionnaire was used to find out the users' opinion on the quality of service at certain times, and the data were used to find the correlation between a performance measure and the quality of service perceived by the system users (travelers).

Some studies, in addition to the perceived quality of service, also collected participants' sociodemographic data. Such approach was used in surveys focusing pedestrians [81], public transport riders [82,83], drivers traveling through toll plazas [32,33], freeway merging conflicts [34] and rural freeway segments [21,35,37,38]. In these studies, the main objective was to assess the importance of a set of performance measures and sociodemographic factors in the user's perception of service quality.

The literature suggests that, in basic freeway segments, the main factor influencing user's perception of service quality is traffic density, which interferes with freedom to maneuver within the traffic stream [7,35,43]. However, few of these studies have evaluated how or whether sociodemographic factors affect the users' perception of trip quality [21,38].

Among the statistical tools to assess the effect of independent variables on a dependent variable, three methods were considered to assess the effects of sociodemographic and traffic factors on the users' perception of trip quality on freeways: bivariate correlation, linear regression and decision trees. Bivariate correlation is a technique that was developed in the late 19th century and is still used today in several areas of science. It can be used to statistically express the relationship between two variables [84]. The most used correlation coefficients are the Pearson correlation coefficient r (when the data follow a normal distribution) and Spearman's ρ for cases where the data do not follow a normal distribution. In traffic engineering, correlations were found in road safety to evaluate the transferred Safety Performance Function (SPF) and the local SPF segment ranking [85], public opinion on automated driving comparing the descriptive statistics of responses between age, gender, mileage and country level [86], comparing perceived congestion level and performances of velocity and local densities in pedestrian comfortability perception studies [81] and comparing driving perception of the trip quality with traffic conditions in a freeway using a driving simulator [38].

Multivariate linear regression is a tool used to obtain the importance that each independent variable imposes to obtain a predictive model of the dependent variable [87]. Researchers have used regression to evaluate the relationship between drivers' physiological responses, situational factors and takeover request lead time in a simulated driving environment [88], to estimate pavement condition indices as a function of the International Roughness Index (IRI) and the Pavement Condition Index (PCI) [89], to predict annual average daily traffic function of road, land use and demographic and socioeconomic characteristics [90,91] and to explore the relationship between on-ramp flow and capacity [92]. In terms of regression models, the most used is linear regression and one of the methods that can be adopted is stepwise. This method starts with a model defined only by a constant. Afterward, the factor that best predicts the model output is identified, selecting the factor with the highest simple correlation coefficient with the dependent variable. If this first predictor increases the model's ability to predict the output, it is maintained and a second predictor is analyzed. This routine continues until no independent variable meets the criteria to be added to the model [93].

A decision tree is a non-parametric data mining tool that presents the underlying relationships in a dataset in a simple way. In addition to the relationship between dependent and independent variables, the division algorithm also identifies the independent variables that maximize homogeneity in the groups that form each child node, according to the dependent variable [94]. One of the models used is the classification and regression tree (CART) method [95]. This model was used to identify the key factors affecting bus transit service quality [96], when analyzing accident risks with frontal collisions using a driving simulator [97] and to study the main factors that influence accident risks on freeways [98], identifying the factors with the greatest impact on the severity of accidents [99, 100].

4.3 Method

As discussed in previous chapters, data were collected through an online survey. Surveys answered on the Internet, using devices such as cell phones or notebooks, manage to obtain more responses from users with different sociodemographic characteristics, from different locations and at a low cost [66, 67, 101].

The sociodemographic questionnaire contained 11 questions, including age, sex, education, and driving habits and experience. Afterward, a sequence of 12 video clips was presented, depicting different traffic and road conditions. To measure user perceptions of trip quality, respondents rated each video using a continuous visual analog scale [59], ranging from "very poor" to "excellent". Arbitrarily, this variable was defined as 0 corresponding to the worst rating and 100, which was the best rating. To define the scenarios for the video clips, a fractional factorial design was made to define the characteristics of the video clips.

4.3.1 Experimental design: traffic and road factors

To create the experiments that were presented to the respondents, five traffic and road factors were considered: traffic density, truck percent, posted speed limit, grade magnitude and number of traffic lanes. All had different amounts of levels and multiples of 2. Table 6 presents the factors, their levels and the number of pseudo-factors assigned to them.

Table 6 – Factors used in the fractional factorial design, along with their description and levels

Factor	Description		Levels	Number of pseudo-factors	
A	k	Traffic density	16	4	(A1, A2, A3 and A4)
B	P_T	Truck percentage	4	2	(B1 and B2)
C	SL	Speed limit	4	2	(C1 and C2)
D	g	Grade magnitude	2	0	
E	N	Number of traffic lanes	2	0	

As the use of the 1024 ($16 \times 4 \times 4 \times 2 \times 2$) possible combinations would require a great deal of effort, cost and time to collect the data for the experiment, a fractional factorial design was chosen to reduce the number of combinations used. The criteria for choosing the fraction were the possibility of testing the main effects of the five factors and the interaction effects of the most important factors, in this case, traffic density and truck percent [77, 102].

For this experiment, the total number of main effects was 23 (15 from factor A, 3 from factors B and C, and 1 from factors D and E). Regarding first-order interactions, the interaction effects between factors A and B were considered the most relevant, totaling 45 effects (15×3). Adding the main effects and the interaction effects, a total of 68 effects of interest were obtained ($23 + 45$). The closest value greater than 68 that was also a fraction of 1024 was 128 experiments, corresponding to a fraction of $1/8$ ($1024/8 = 128$).

The factor levels of the experiment were all chosen as multiples of 2. Thus, a 2^{k-p} fractional factorial design was used to select the $1/8$ fraction of the combination of factors, in which some adaptations were necessary. First, it was identified what the number of factors should be if they all had only 2 levels. Therefore, since the total number of possible combinations was $2^{10} = 1024$, we should have $k = 10$ factors. To obtain 128 combinations, we would have $p = 3$ and $2^{10-3} = 2^7 = 128$. Thus, pseudo-factors for factors A, B and C had to be created. Table 6 presents these pseudo-factors created for this experiment. It can be observed that factor A, which has 16 levels, was equivalent to having 4 pseudo-factors: $A1$, $A2$, $A3$ and $A4$. Factors B and C, with 4 levels each, required 2 pseudo-factors for each factor: $B1$, $B2$, $C1$ and $C2$.

In experiments with 2 levels, the most commonly used notation to designate levels is -1 at the lowest level and $+1$ at the highest level. There is a standard order to create the design matrix for the experiment [77], as can be seen in Table 7 (the complete design of the 128 rows can be found in Table 18 (Appendix B). As there were $2^7 = 128$ combinations, this table has 7 columns and 128 rows, and each row is a configuration of the experiment. The first column comprised $2^1 = 2$ blocks, the first block was the first $2^6 = 64$ rows filled with -1 and the second block also had 64 rows filled with $+1$. The second column consisted of $2^2 = 4$ blocks, and the first block was the first $2^5 = 32$ rows filled with -1 and the second block comprised the following 32 rows filled with $+1$, the third block was equal to the first and the fourth was equal to the second. The next columns were created in a similar way to the previous ones.

Considering Table 7 containing 128 rows and 7 columns, 3 columns had to be added, resulting in a table containing the $k = 10$ factors of the experiment plan (Table 8). In this 10-column table, as factors A and B were considered the most relevant, columns 1 to 4 were used for pseudo-factors $A1$, $A2$, $A3$, and $A4$ and columns 5 and 6 for pseudo-factors $B1$ and $B2$. Thus, it was ensured that the effects of these factors were not confounded with other factors, called free main effects. Columns 7 and 8 were assigned to pseudo-factors $C1$ and $C2$. To complete the table, the last two columns were used, with factors D and E. These main effects are not free effects but confounded with some interaction assumed to be non-active. For this plan, the following confounders were used: $C2 = A1 \cdot A4 \cdot B1 \cdot C1$; $D = A2 \cdot A4 \cdot B2 \cdot C1$; $E = A1 \cdot A3 \cdot A4 \cdot B2$. Table 8 presents the 10 columns with each pseudo-factor, which form

Table 7 – Design matrix for 128 runs, created with $2^{(column)}$ blocks of $2^{(column-1)}$ of -1 and $2^{(column-1)}$ of +1 in each column

	Columns						
Runs	1	2	3	4	5	6	7
1	−1	−1	−1	−1	−1	−1	−1
2	−1	−1	−1	−1	−1	−1	+1
3	−1	−1	−1	−1	−1	+1	−1
4	−1	−1	−1	−1	−1	+1	+1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
8	−1	−1	−1	−1	+1	+1	+1
9	−1	−1	−1	+1	−1	−1	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
16	−1	−1	−1	+1	+1	+1	+1
17	−1	−1	+1	−1	−1	−1	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
32	−1	−1	+1	+1	+1	+1	+1
33	−1	+1	−1	−1	−1	−1	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
64	−1	+1	+1	+1	+1	+1	+1
65	+1	−1	−1	−1	−1	−1	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
126	+1	+1	+1	+1	+1	−1	+1
127	+1	+1	+1	+1	+1	+1	−1
128	+1	+1	+1	+1	+1	+1	+1

the experimental plan of the 128 experiments. Table 19 (Appendix B) presents the complete design.

Table 9 (A) shows the association of the pseudo-factors $A1$, $A2$, $A3$ and $A4$, used to create the 16 original factor A levels. Similarly, Table 9 (B) shows the associations used to create the original 4 factors B and C levels. As factors D and E only have 2 levels, they were identified as -1 for level 1 and +1 for level 2.

Finally, Table 10 summarizes the final plan executed for 128 combinations of factors (the complete plan can be found in Table 20 (Appendix B)). Factor A, which is the most important factor of the experiment, represents the traffic density and had 16 levels. Afterward, factor B corresponds to the truck percent in the traffic stream. Factor C was chosen to represent the posted speed limit on the freeway. Finally, factors D and E were chosen to represent the grade magnitude and the number of traffic lanes, respectively.

Having defined the experiment plan, three replicates were made for each scenario, totaling 384 video clips, which depict, with a high level of fidelity, the corresponding combination of traffic and road conditions in the freeway segment, as detailed in Piva, Setti and Washburn[48].

4.4 Data collection

To reach as many people as possible, the survey was disseminated on social media such as WhatsApp, Instagram, Facebook and LinkedIn. In addition, mailing lists were used to publicize the site to specific groups of professionals, such as professors and transportation engineers.

Table 8 – Design matrix combining the 7 columns for 128 runs and pseudo-factors $C2$, D, and E, creating a 10 columns designs matrix

	Pseudo-factors									
Runs	$A1$	$A2$	$A3$	$A4$	$B1$	$B2$	$C1$	$C2$	D	E
1	−1	−1	−1	−1	−1	−1	−1	+1	+1	+1
2	−1	−1	−1	−1	−1	−1	+1	−1	−1	+1
3	−1	−1	−1	−1	−1	+1	−1	+1	−1	−1
4	−1	−1	−1	−1	−1	+1	+1	−1	+1	−1
5	−1	−1	−1	−1	+1	−1	−1	−1	+1	+1
6	−1	−1	−1	−1	+1	−1	+1	+1	−1	+1
7	−1	−1	−1	−1	+1	+1	−1	−1	−1	−1
8	−1	−1	−1	−1	+1	+1	+1	+1	+1	−1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
31	−1	−1	+1	+1	+1	+1	−1	+1	+1	−1
32	−1	−1	+1	+1	+1	+1	+1	−1	−1	−1
33	−1	+1	−1	−1	−1	−1	−1	+1	−1	+1
34	−1	+1	−1	−1	−1	−1	+1	−1	+1	+1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
63	−1	+1	+1	+1	+1	+1	−1	+1	−1	−1
64	−1	+1	+1	+1	+1	+1	+1	−1	+1	−1
65	+1	−1	−1	−1	−1	−1	−1	−1	+1	−1
66	+1	−1	−1	−1	−1	−1	+1	+1	−1	−1
67	+1	−1	−1	−1	−1	+1	−1	−1	−1	+1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
125	+1	+1	+1	+1	+1	−1	−1	−1	+1	−1
126	+1	+1	+1	+1	+1	−1	+1	+1	−1	−1
127	+1	+1	+1	+1	+1	+1	−1	−1	−1	+1
128	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1

Table 9 – Combinations of pseudo-factors to create levels for each factor: (a) Pseudo-factors $A1$, $A2$, $A3$ and $A4$ used to create 16 A levels and (b) pseudo-factors $B1$ or $C1$ and $B2$ or $C2$ combined to create 4 levels for factors B and C

(A)

$A1$	$A2$	$A3$	$A4$	Factor A levels
−1	−1	−1	−1	1
−1	−1	−1	+1	2
−1	−1	+1	−1	3
−1	−1	+1	+1	4
−1	+1	−1	−1	5
−1	+1	−1	+1	6
−1	+1	+1	−1	7
−1	+1	+1	+1	8
+1	−1	−1	−1	9
+1	−1	−1	+1	10
+1	−1	+1	−1	11
+1	−1	+1	+1	12
+1	+1	−1	−1	13
+1	+1	−1	+1	14
+1	+1	+1	−1	15
+1	+1	+1	+1	16

(B)

$B1$ or $C1$	$B2$ or $C2$	Factors B or C levels
−1	−1	1
−1	+1	2
+1	−1	3
+1	+1	4

Table 10 – Condensed planning matrix for the 128 scenarios used for the experiment

Scenario	Traffic density	Truck percent	Speed limit	Grade magnitude	Number of traffic lanes
1	1	1	2	2	2
2	1	1	3	1	2
3	1	2	2	1	1
4	1	2	3	2	1
5	1	3	1	2	2
6	1	3	4	1	2
7	1	4	1	1	1
8	1	4	4	2	1
⋮	⋮	⋮	⋮	⋮	⋮
31	4	4	2	2	1
32	4	4	3	1	1
33	5	1	2	1	2
34	5	1	3	2	2
⋮	⋮	⋮	⋮	⋮	⋮
63	8	4	2	1	1
64	8	4	3	2	1
65	9	1	1	2	1
66	9	1	4	1	1
67	9	2	1	1	2
⋮	⋮	⋮	⋮	⋮	⋮
125	16	3	1	2	1
126	16	3	4	1	1
127	16	4	1	1	2
128	16	4	4	2	2

Data were collected from June to October 2019 and 830 people responded to the survey, totaling 8,639 video clips watched and rated.

To eliminate possible spurious responses from the database, two filters were applied. The first filter referred to the time the respondent spent on each page to assign a rating to the video clip. If the user remained on the page for less than 12 seconds, this rating was discarded. This filter's objective was to discard ratings for which the participant did not watch the video clip long enough to be aware of the traffic and road conditions to provide a reliable rating. Using this filter, 396 evaluations were removed, corresponding to 4.6% of the sample. The second filter was related to the consistency of the participant's responses. If the participant watched the 12 video clips in the survey, this set depicted conditions ranging from light traffic to traffic close to capacity. Thus, it was expected that the ratings given to the worst and best scenarios would be sufficiently different from each other. In this case, if the difference between the lowest and highest rating attributed by the participant was less than 50 (half of the total range, from 0 to 100), that participant was excluded from the sample. After implementing the two filters, 625 participants remained who watched 7,004 video clips.

Table 11 summarizes the participants' sociodemographic characteristics, as well as the number of respondents and video clips watched, and their percentage in the sample. From the factors investigated, two variables were defined as continuous: age and years of driver's license, which were correlated ($r = 0.931$). Since these two factors have a strong correlation, only the respondents' age was used in the analyses. Five variables were identified as categorical: sex, have a driver's license, type of vehicle used, drive on freeways and state of residence. The other factors were classified as ordinal. From the people who answered the survey, 60% declared themselves as male, 56% were below 35 years old and 62% had a university undergraduate degree. Most respondents stated that they are the driver on their trips, travel up to 2 times a month, use a car, drive in the state of São Paulo and do not face congestion on freeways. As a result, the sample obtained should be considered a convenience sample.

As the sample is a convenience sample, some factors had few respondents at certain levels (e.g., respondents with low education level). As a result, it was decided to aggregate respondents into just two levels for some factors, as shown in Table 12.

From the five traffic factors in each video, three were defined as ordinal: posted speed limit at 90, 100, 110 and 120 km/h; grade magnitude, with 1% and 4%; and number of lanes, with 3 or 4 traffic lanes. The truck percent and the traffic density were considered continuous variables. The truck percent varied between 0% and 39.7% while the traffic density varied between 0.95 and 19.8 veh/km/ln.

4.5 Results

First, an initial analysis of the data collected from box plots was carried out, to investigate the spread and skewness of the ratings as a function of the traffic stream and road factors considered in the scenarios. Figure 18 summarizes the analysis results. The five box plots in Figure 18 show that:

(1) There seem to be little difference in the distribution of ratings due to the number of lanes or grade magnitude, as the box plots in Figures 18 (A) and (B) are quite similar for the different values of these two factors, suggesting that these two factors (grade magnitude and number of lanes) are not very important to the respondents' perception of the trip quality.
(2) The box plot in Figure 18 (C) suggests that ratings were higher as the posted speed limit increases.
(3) Figure 18 (D) indicates that ratings generally decrease as the fraction of trucks in traffic increases.
(4) The box plot in Figure 18 (E) shows an evident correlation between ratings and traffic density.

In short, the box plots indicate that the relationships between rating and the traffic factors are as one would expect.

A heat map was made with the frequency ratings as a function of traffic density and the ranges of the respondents' ratings, shown in Figure 19. The graph indicates that rating 0 and 100 are more frequent than the other ratings. It can also be observed that rating 100 is more frequent for low densities than for high densities, as expected, and that the opposite occurs for rating 0. In addition, it was observed that the heat map presents higher frequencies of high

Table 11 – Summarized sociodemographic data

Variable	Category	Respondents Total	%	Watched clips Total	%
Age	18 – 25	86	13.8	965	13.8
	26 – 30	141	22.6	1553	22.2
	31 – 35	122	19.5	1361	19.4
	36 – 40	89	14.2	1001	14.3
	41 – 45	43	6.9	474	6.8
	46 – 50	44	7.0	507	7.2
	51 – 55	26	4.2	296	4.2
	56 – 60	30	4.8	338	4.8
	> 60	44	7.0	509	7.3
Sex	Male	378	60.5	4215	60.2
	Female	245	39.2	2775	39.6
	Others/Do not declare	2	0.3	14	0.2
Education level	Incomplete primary education	1	0.2	3	0.0
	Complete primary education	0	0.0	0	0.0
	Incomplete high school	1	0.2	12	0.2
	Complete high school	20	3.2	199	2.8
	Incomplete undergraduate	83	13.3	944	13.5
	Complete undergraduate	281	45.0	3143	44.9
	Complete master degree	167	26.7	1891	27.0
	Complete doctoral degree	72	11.5	812	11.6
Driver’s license	Yes	608	97.3	6825	97.4
	No	17	2.7	179	2.6
Years of driver’s license	0 – 5	77	12.3	829	11.8
	6 – 10	119	19.0	1350	19.3
	11 – 15	135	21.6	1503	21.5
	16 – 20	91	14.6	1015	14.5
	21 – 25	55	8.8	632	9.0
	26 – 30	43	6.9	478	6.8
	31 – 35	28	4.5	319	4.6
	36 – 40	25	4.0	267	3.8
	> 40	52	8.3	611	8.7
Driving frequency	Never	139	22.2	1565	22.3
	Half the times	115	18.4	1277	18.2
	Always	371	59.4	4162	59.4
Trip frequency	Do not travel	2	0.3	24	0.3
	Rarely	79	12.6	882	12.6
	Not frequent	160	25.6	1808	25.8
	Twice a month	109	17.4	1236	17.6
	More than twice a month	99	15.8	1116	15.9
	Twice a week	54	8.6	601	8.6
	More than twice a week	67	10.7	760	10.9
	Twice a day	44	7.0	464	6.6
	More than twice a day	11	1.8	113	1.6
Type of vehicle	Car	548	87.7	6150	87.8
	Bus	63	10.1	706	10.1
	Truck	2	0.3	20	0.3
	Motorcycle	12	1.9	128	1.8
Drive on freeways	Yes	525	84.0	5883	84.0
	No	99	15.8	1109	15.8
State of residence	São Paulo	357	57.1	4010	57.3
	Minas Gerais	69	11.0	786	11.2
	Paraná	39	6.2	433	6.2
	Rio de Janeiro	29	4.6	311	4.4
	Ceará	20	3.2	235	3.4
	Other	110	17.6	1217	17.4
Driving in congested conditions	Never	364	58.2	4080	58.3
	Half the times	192	30.7	2157	30.8
	Always	69	11.0	767	11.0

Figure 18 – Box plots of ratings and traffic factors – (A) Grade magnitude, (B) Number of traffic lanes, (C) Speed limit , (D) Truck percent, and (E) Traffic density

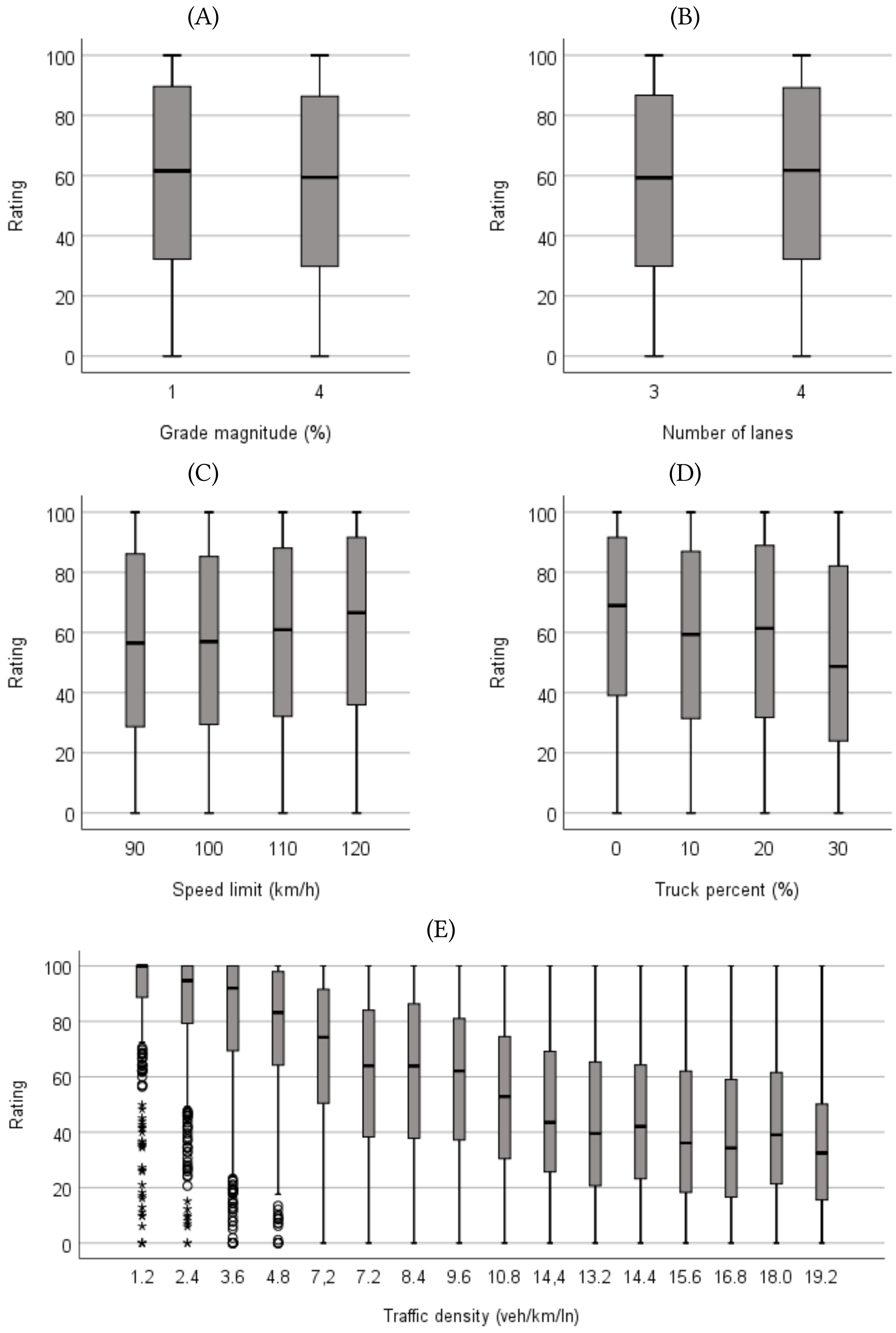


Table 12 – Sociodemographic factors gathered into two categories

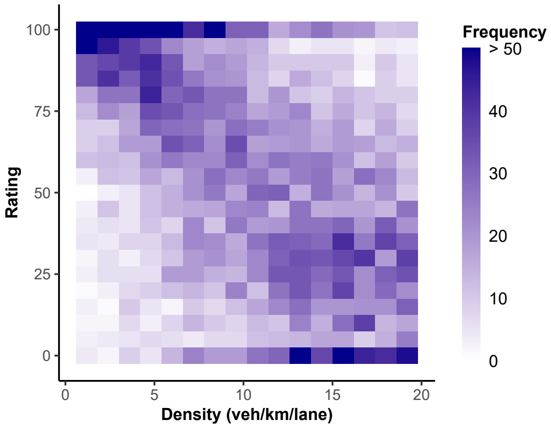
Variables	Category	Video clips watched	Percent
Sex	Male	4215	60.30
	Female	2775	39.70
Education level	Until complete undergraduate	4301	61.41
	Graduate (Master or Doctoral)	2703	38.59
Driving frequency	Never	1565	22.34
	Half the times or more	5439	77.66
Trip frequency	Twice a month or less	3950	56.40
	More than twice a month	3054	43.60
Type of vehicle	Car	6150	87.81
	Others	854	12.19
State of residence	São Paulo	4010	57.25
	Others	2994	42.75
Driving in congested conditions	Never	4080	58.25
	Halt the times or more	2924	41.75

ratings (> 75) for low traffic densities (< 5 veh/km/ln) and high frequency of low ratings (< 30) for high traffic densities (> 15 veh/km/ln).

The Kolmogorov-Smirnov test was used to investigate the normality of the dependent variable (rating). The test result [$D(7004) = 0.094$, $p < 0.001$] indicates that the rating distribution is significantly non-normal.

The relationships between the independent variables and the dependent variable were investigated using three statistical techniques: (1) bivariate correlation; (2) multiple linear regression; and (3) decision trees. The results of these analyses are shown in the following sections.

Figure 19 – Heat map of the frequency of ratings, according to the ratings and traffic densities



4.5.1 Bivariate correlation

Considering the Kolmogorov-Smirnov test results, the bivariate correlation between the dependent variable and the independent variables was performed using the Spearman's ρ correlation coefficient. Table 13 summarizes the results found in terms of ρ values and their significance level.

As shown in Table 13, traffic density is the only factor that shows a significant, strong and negative correlation with the rating ($\rho = -0.520$ and $p < 0.001$), which means that, as the traffic density increases, the respondents perceive a drop in the trip quality. This result confirms the hypothesis that traffic density is an adequate performance measure to reflect the freeway user perception of service quality.

The results obtained also show that five factors did not present a statistically significant correlation with the user perception of trip quality given by the ratings: frequency with which they are the driver ($\rho = -0.020$ and $p = 0.092 > 0.05$), having a driver's license ($\rho = -0.015$ and $p = 0.213$), type of road on which they travel ($\rho = -0.013$ and $p = 0.294$), type of vehicle used ($\rho = -0.011$ and $p = 0.367$), and sex ($\rho = 0.009$ and $p = 0.454$).

The other factors, despite showing a statistically significant correlation with the dependent variable ($p < 0.05$), are weakly correlated with the rating ($0.029 \leq |\rho| \leq 0.116$), which indicates that they did not significantly affect the respondents' perception of the trip quality.

Table 13 – Degree of correlation between ratings and traffic and sociodemographic factors, as measured by Spearman's ρ, and associated probabilities p

Factor	Spearmans's ρ	p-value
Traffic density	−0.520	< 0.001
Truck percent	−0.116	< 0.001
Posted speed limit	0.072	< 0.001
State of residence	−0.070	< 0.001
Age	−0.050	< 0.001
Grade magnitude	−0.040	0.001
Trip frequency	−0.037	0.002
Number of traffic lanes	0.037	0.002
Driving in congested conditions	−0.033	0.005
Education level	0.029	0.014
Driving frequency	−0.020	0.092
Driver's license	−0.015	0.213
Type of road	−0.013	0.294
Type of vehicle	−0.011	0.367
Sex	0.009	0.454

4.5.2 Linear regression

To quantify the importance of factors in user perception, a multiple linear regression model was calibrated using the rating as the dependent variable and the other 16 factors as independent variables. It is important to emphasize that this model is not intended to predict the ratings given by users through the other factors evaluated. The strategy chosen for model building was stepwise regression and the criteria adopted were: probability of $F \leq 0.05$ of being inserted and probability of $F \geq 0.10$ of being removed. The final model is based on ten factors and no independent variable showed collinearity with another. Table 14 presents the sequence of input factors in the model, the standardized coefficients, the t-test and the significance of the independent variables. The variables sex, driver's license, driving frequency, type of vehicle and type of freeway were not included in the model.

The standardized coefficients show that density ($\beta = -0.514$) has an impact almost four times greater than the second most important factor, the truck percent ($\beta = -0.143$), on the rating given by respondents for trip quality. Other traffic factors, such as the posted speed limit, grade magnitude and number of lanes entered the model and were considered significant, however, with importance 10 times smaller than traffic density – that is, they have a much smaller importance in perception of trip quality. None of the sociodemographic factors had a standardized coefficient module greater than 0.1, which suggests that they do not significantly influence the perception of trip quality.

4.5.3 Decision tree

To investigate the effect of the independent variables on the ratings, a decision tree model was trained with a CART algorithm [95]. CART can identify the most important independent variables and how they interact with one another [103], making it a good choice for this task.

SPSS was used to fit the CART, with the following parameters: maximum tree depth = 5;

Table 14 – Standardized coefficients, t-value, and significance for independent variables using linear regression stepwise model

Variable	Standardized coefficient β	t-value	p-value
(Constant)		18.423	< 0.001
Traffic density	−0.514	−51.105	< 0.001
Truck percent	−0.143	−14.231	< 0.001
State of residence	−0.051	−4.803	< 0.001
Posted speed limit	0.060	5.968	< 0.001
Grade magnitude	−0.047	−4.647	< 0.001
Number of traffic lanes	0.042	4.177	< 0.001
Education level	0.037	3.611	< 0.001
Driving in congested conditions	−0.027	−2.643	0.008
Trip frequency	−0.030	−2.926	0.003
Age	−0.023	−2.178	0.029

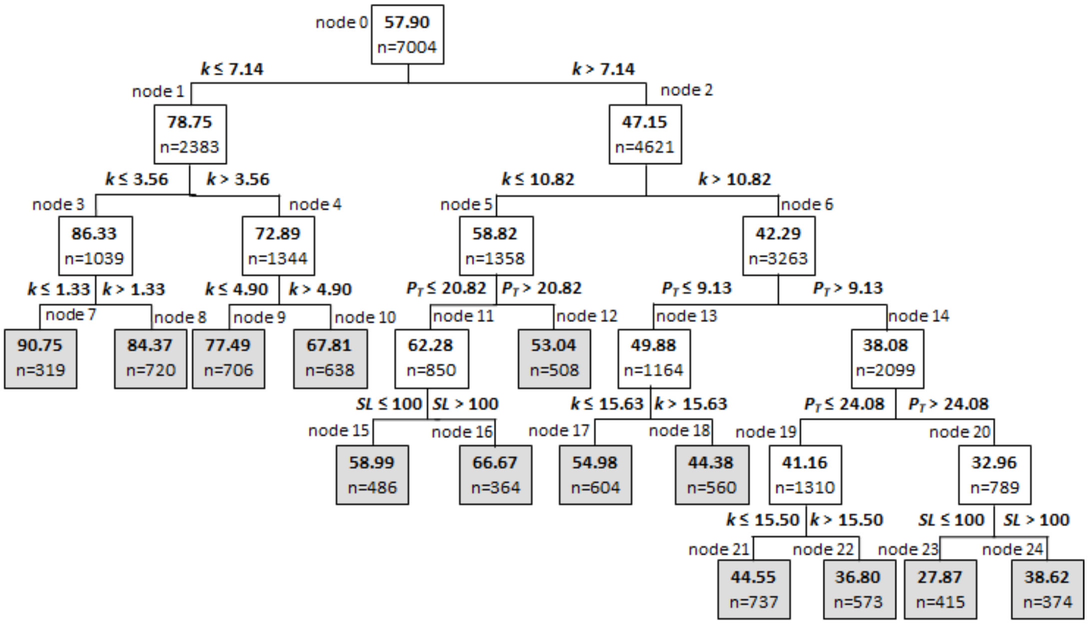
minimum number of cases in child nodes = 300; minimum number of cases in parent nodes = 600; and minimum improvement in the variance reduction = 0.0001. Figure 20 presents the decision tree that presented 5 levels, 24 nodes and 13 terminal nodes. It contains the averages of the ratings and the number of responses referring to each tree node.

The first division performed by the algorithm was performed by the traffic density variable. Node 1, formed by video clips that depicted densities lower than 7.14 veh/km/ln, shows that all the subsequent divisions are carried out based on traffic density. This suggests that when traffic density is low, other independent variables did not influence the user perception of trip quality.

Node 2 defines a branch with traffic density greater than 7.14 veh/km/ln, whose first division was also defined by the density ($k \leq 10.82$ or $k > 10.82$ veh/km/ln). In the division of the second level, both for the division of node 5 and node 6, the variable truck percent (P_T) was used. For node 5, which presented densities between 7.15 and 10.82 veh/km/ln, the division was made in $P_T = 20.82\%$, while in node 6, which represents cases with $k > 10.82$ veh/km/ln, the tree division occurs with $P_T = 9.13\%$. This indicates that the truck percent in the traffic stream has an impact on user perception when the traffic density is greater than 7.15 veh/km/ln.

If the traffic density is less than or equal to 10.82 veh/km/ln (node 5), the truck percent negatively affects the perception of trip quality if $P_T > 20.8\%$ (node 12). If $P_T \leq 20.82\%$ (node 11), SL, the posted speed limit becomes the factor that divides the respondents: the trip quality is better if the $SL > 100$ km/h (node 16, terminal node) and worse, if otherwise (node 15, terminal node).

Figure 20 – CART tree of participants' perception of trip quality in relation to the traffic variables. The values in the boxes are the average rating for each group and the number of responses



For traffic density $k > 10.82$ veh/km/ln (node 6), the truck percent P_T is the factor that defines the perception of the trip quality, for values $P_T > 9.13\%$ (node 14). For scenarios with $9.13\% < P_T \leq 24.08\%$ (node 19), the trip quality is perceived as better if $k \leq 15.50$ veh/km/ln (node 21, terminal node), or worse, if $k > 15.50$ veh/km/ln (node 22, terminal node). In scenarios with $P_T > 24.08\%$ (node 20), the posted speed limit becomes the factor that defines the perception of trip quality: the quality is perceived as worse, if $SL \leq 100$ km/h (node 23, terminal node) and better if $SL > 100$ km/h (node 24, terminal node).

Table 15 presents the relative importance of each of the independent variables obtained by the CART model. Traffic density was more important in the user perception. The truck percent was the second most important, but with a magnitude about 10 times smaller than the traffic density. The third most important variable was the posted speed limit, with a relative importance of 1.9%. It is important to note that only these three variables appeared in the decision tree and the number of lanes and grade magnitude variables were in the sixth and seventh position, respectively.

Table 15 – Relative importance of the independent variables using CART method

Independent variable	Relative importance (%)
Traffic density	100.000
Truck percent	8.858
Posted speed limit	1.925
Age	1.031
State of residence	0.727
Number of traffic lanes	0.651
Grade magnitude	0.390
Type of vehicle	0.113
Driving frequency	0.094
Driver's license	0.071
Driving in congested conditions	0.030
Sex	0.007
Education level	0.006
Type of road	0.000

4.6 Conclusions

The research concluded that the sociodemographic factors analyzed did not influence the respondents' perception of the trip quality. From the traffic factors, traffic density showed the greatest relevance in the ratings given by the respondents, as identified in other studies [7, 35]. The second highest importance was related to the truck percent and was more relevant for travels under more intense traffic, with traffic densities greater than 7 veh/km/ln.

Since the method presented consistent results compared to the literature, the use of traffic microsimulation, with animated video clips recorded from the driver's viewpoint, proved to be an efficient tool to identify user perception of the trip quality on freeways. In addition, although it was considered a convenience sample, the web survey proved to be an efficient tool to obtain the necessary sample (reaching more than 800 respondents) for this study.

It is recommended in future research that the study of sociodemographic factors in terms of user perception should be analyzed in more depth. A question that can be attempted to be answered is "Do people of different sexes, or different education levels, rate travel in the same way, or do they give different importance to traffic factors?" In addition, the survey can be disseminated once more to try to collect responses from groups that were not representative in this collection, such as a lower education levels, or bus, truck and motorcycle drivers. Additional aspects that would be interesting to study include the effect of truck percent in densities higher than 7 veh/km/ln, with more levels of truck percent and greater grade magnitudes.

Chapter

Concluding remarks 5

The research reported in this dissertation had the main objective of developing a method to incorporate the travelers' perception of the quality of trips into estimating LOS boundaries for freeway segments and investigating the effect of sociodemographic, traffic and roadway factors on the perceived quality of service. The three previous chapters contain their conclusions and recommendations, having been written as self-contained articles. Instead of repeating them, these final remarks focus on the research questions proposed to reach the main objective of the research.

The first question was how to measure the travelers' perception of the quality of service on freeways. The results show that using animated video clips associated with a visual analog scale proved efficient. The proposed approach has demonstrated that, with modest costs, it was possible to create 417 realistic video clips with the desired traffic and freeway conditions and, using a web-based survey, to recruit almost a thousand respondents who provided a sample of over ten thousand ratings of freeway trips. Even though samples obtained from web surveys are almost always convenience samples, this study's sample is, by far, the largest in the extant literature.

The second question focused on verifying the existence of a correlation between the quality of service as perceived by drivers and a service measure. The results have shown a clear correlation between traffic stream density (one of the service measures used by the HCM for freeways) and the participants' ratings. Furthermore, traffic density was the factor that most influenced the participants' perception of trip quality. This result was consistent with other references found in the literature. The second factor that most affected the participants' perception of the trip quality was the percentage of trucks, and the decision tree has shown that this factor is more relevant as the traffic gets heavy. Posted speed limit, grade magnitude, and the number of lanes were also correlated to the ratings, albeit at a much lower level.

The third question was how to incorporate the travelers' perceptions of trip quality into the estimation of density-based freeway LOS criteria. The proposed approach, which is flexible enough to be applied to a wide range of highway components, proved equal to that task, as described in Chapter 3. Logistic regression was used to find the density level at which the probability of users rating the trip as LOS A is equal to the probability of users rating the trip not being in LOS B; the process is then repeated for the next LOS and so on until the last LOS. Two of the advantages of proposed approach are that it provides the confidence interval for the estimated LOS thresholds and it does not require any assumption about the number of service levels before the data collection. These two characteristics of the proposed approach have shown that participants in the case study reported in this dissertation do not distinguish LOS D from LOS E.

The two final research questions were related to the investigation of possible effects of traffic, roadway, and sociodemographic characteristics on the participants' perception of the trip quality. A fractional factorial design was used to investigate this issue, as described in Chapter 4. The results indicate that, at least for the sample used in this study, sociodemographic factors do not have any significant influence on the perceived quality of the trips. In contrast, traffic density is the main factor that affect such perception. Interestingly, truck percentage has been shown to have an impact that is significant only when density is above a certain threshold, which suggests that this finding deserves further investigation.

Considering that all research questions have been answered, one can thus conclude that the main goal of the study has been successfully reached.

References

1 ANTT. *Relatório anual de atividades 2019*. Brasília, DF, 2020. Available at: https://portal.antt.gov.br/relatorios-anuais.

2 IBGE. *Produto Interno Bruto dos municípios 2019*. Rio de Janeiro, RJ, 2021. Available at: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101896_informativo.pdf.

3 ARTESP. *Relatório da administração e das demonstrações contábeis do exercício de 2021*. São Paulo, SP, 2022. Available at: http://www.artesp.sp.gov.br/RelatorioAnual/RELATORIOANUAL2021.pdf.

4 DER-SP. *Malha rodoviária do estado de São Paulo*. São Paulo, SP, 2021. Available at: http://www.der.sp.gov.br/WebSite/Arquivos/MalhaRodoviaria/ExtensaoMalha2021.pdf.

5 ANDRADE, G. R.; RODRIGUES-SILVA, K. C.; PUTY FILHO, S. A. Panorama normativo e tecnológico da avaliação operacional das concessões rodoviárias. *In*: CBR&C. *Anais do 7o. Congresso Brasileiro de Rodovias e Concessões*. Foz do Iguaçú, PR, 2011. p. 14.

6 DNIT. *Manual de estudos de tráfego: publicação IPR-723*. Rio de Janeiro, RJ, 2006. Available at: http://ipr.dnit.gov.br/normas-e-manuais/manuais/documentos/723_manual_estudos_trafego.pdf.

7 TRB. *Highway Capacity Manual 7th edition: A guide for multimodal mobility analysis*. Washington, DC, 2022. Available at: https://doi.org/10.17226/26432.

8 LEMKE, K. The new German Highway Capacity Manual (HBS 2015). *Transportation Research Procedia*, v. 15, p. 26–35, 2016. Available at: https://doi.org/10.1016/j.trpro.2016.06.003.

9 HEIKOOP, H.; HENKENS, N. Recent developments and distory of the Dutch HCM. *Transportation Research Procedia*, v. 15, p. 51–62, 2016. Available at: https://doi.org/10.1016/j.trpro.2016.06.005.

10 UTIMURA, F. S.; SETTI, J. R.; EGAMI, C. Y.; MON-MA, M. L. Verificação da acurácia de estimativas do nível de serviço em rodovias de pista simples no estado de São Paulo. *In*: *Anais*

do XXI Congresso de Pesquisa e Ensino em Transportes - ANPET. Rio de Janeiro, RJ: Associação Nacional de Pesquisa e Ensino em Transportes (ANPET), 2007.

11 CUNHA, A. L. B. N.; SETTI, J. R. Fatores de equivalência para caminhões em rodovias de pista dupla. *In*: CBR&C. *Congresso Brasileiro de Rodovias & Concessões*. Florianópolis, SC, 2009.

12 PIVA, F. J.; SETTI, J. R. Avaliação do impacto de veículos pesados na qualidade de serviço de rodovias de pista dupla usando dados empíricos. *Transportes*, v. 23, n. 4, p. 51–59, 2015. Available at: https://doi.org/10.14295/transportes.v23i4.949.

13 CARVALHO, L. G. S.; SETTI, J. R. Calibration of the VISSIM truck performance model using GPS data. *Transportes*, v. 27, n. 3, p. 131–143, 2019. Available at: https://doi.org/10.14295/transportes.v27i3.2042.

14 TRB. *Highway Capacity Manual, Special Report 87*. 2nd edition. Washington, DC, 1965.

15 TRB. *Highway Capacity Manual, Special Report 209*. 3rd edition. Washington, DC, 1985.

16 TRB. *Highway Capacity Manual*. 5th edition. Washington, DC, 2010.

17 ROESS, R. P.; PRASSAS, E. S. *The Highway Capacity Manual: A conceptual and research history*. Springer International Publishing, 2014. v. 5. 491 p. (Springer Tracts on Transportation and Traffic, v. 5). Available at: https://doi.org/10.1007/978-3-319-05786-6.

18 DOWLING, R.; REINKE, D.; FLANNERY, A.; RYUS, P.; VANDEHEY, M.; PETRITSCH, T.; LANDIS, B.; ROUPHAIL, N.; BONNESON, J. *Multimodal level of service analysis for urban streets – NCHRP Report 616*. Washington, DC: Transportation Research Board, 2008. Available at: https://dx.doi.org/10.17226/23086.

19 CHOOCHARUKUL, K.; SINHA, K. C.; MANNERING, F. L. User perceptions and engineering definitions of highway level of service: an exploratory statistical comparison. *Transportation Research Part A: Policy and Practice*, v. 38, n. 9-10, p. 677–689, 2004. Available at: https://doi.org/10.1016/j.tra.2004.08.001.

20 WASHBURN, S.; KIRSCHNER, D. Rural freeway level of service based on traveler perception. *Transportation Research Record*, v. 1988, n. 1988, p. 31–37, 2006. Available at: https://doi.org/10.3141/1988-06.

21 PAPADIMITRIOU, E.; MYLONA, V.; GOLIAS, J. Perceived level of service, driver, and traffic characteristics: Piecewise linear model. *Journal of Transportation Engineering*, v. 136, n. 10, p. 887–894, 2010. Available at: https://doi.org/10.1061/(ASCE)TE.1943-5436.0000154.

22 PÉCHEUX, K. K.; FLANNERY, A.; WOCHINGER, K.; REPHLO, J.; LAPPIN, J. Automobile drivers' perceptions of service quality on urban streets. *Transportation Research Record*, v. 1883, p. 167–175, 2004. Available at: https://doi.org/10.3141/1883-19.

23 FLANNERY, A.; WOCHINGER, K.; MARTIN, A. Driver assessment of service quality on urban streets. *Transportation Research Record*, v. 1920, p. 25–31, 2005. Available at: https://doi.org/10.1177/0361198105192000103.

24 FANG, F. C.; ELEFTERIADOU, L.; PECHEUX, K. K.; PIETRUCHA, M. T. Using fuzzy clustering of user perception to define levels of service at signalized intersections. *Journal of Transportation Engineering*, v. 129, n. 6, p. 657–663, 2003. Available at: https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(657).

25 OTHAYOTH, D.; RAO, K. Statistical analysis of user perceived level of service at signalized intersection. *In*: TRB. *96th Annual Meeting of the TRB*. Washington, D.C., 2017. Available at: https://www.researchgate.net/publication/318645997_Statistical_Analysis_of_User_Perceived_Level_of_Service_at_Signalized_Intersection.

26 SARKAR, P. P.; MALLIKARJUNA, C. Effect of perception and attitudinal variables on mode choice behavior: A case study of indian city, agartala. *Travel Behaviour and Society*, v. 12, p. 108–114, 2018. Available at: https://doi.org/10.1016/j.tbs.2017.04.003.

27 KO, J.; LEE, S.; BYUN, M. Exploring factors associated with commute mode choice: An application of city-level general social survey data. *Transport Policy*, v. 75, p. 36–46, 2019. Available at: https://doi.org/10.1016/j.tranpol.2018.12.007.

28 LI, X.; TANG, J.; HU, X.; WANG, W. Assessing intercity multimodal choice behavior in a touristy city: a factor analysis. *Journal of Transport Geography*, v. 86, 2020. Available at: https://doi.org/10.1016/j.jtrangeo.2020.102776.

29 ARELLANA, J.; FUENTES, L.; CANTILLO, J.; ALVAREZ, V. Multivariate analysis of user perceptions about the serviceability of urban roads: case of barranquilla. *International Journal of Pavement Engineering*, v. 22, n. 1, p. 54–63, 2021. Available at: https://doi.org/10.1080/10298436.2019.1577420.

30 RODRIGUEZ-VALENCIA, A.; BARRERO, G. A.; ORTIZ-RAMIREZ, H. A.; VALLEJO-BORDA, J. A. Power of user perception on pedestrian quality of service. *Transportation Research Record*, v. 2674, n. 5, p. 250–258, 2020. Available at: https://doi.org/10.1177/0361198120914611.

31 VALLEJO-BORDA, J. A.; CANTILLO, V.; RODRIGUEZ-VALENCIA, A. A perception-based cognitive map of the pedestrian perceived quality of service on urban sidewalks. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 73, p. 107–118, 2020. Available at: https://doi.org/10.1016/j.trf.2020.06.013.

32 OBELHEIRO, M. R.; CYBIS, H. B.; RIBEIRO, J. L. Level of service method for brazilian toll plazas. *Procedia – Social and Behavioral Sciences*, v. 16, p. 120–130, 2011. Available at: https://doi.org/10.1016/j.sbspro.2011.04.435.

33 NAVANDAR, Y. V.; DHAMANIYA, A.; PATEL, D. Empirical analysis of level of service based on users perception at manual tollbooth operation in india. *Transportation Research Procedia*, v. 37, p. 314–321, 2019. Available at: https://doi.org/10.1016/j.trpro.2018.12.198.

34 XU, Z.; ZOU, X.; OH, T.; VU, H. L. Studying freeway merging conflicts using virtual reality technology. *Journal of Safety Research*, v. 76, p. 16–29, 2021. Available at: https://doi.org/10.1016/j.jsr.2020.11.002.

35 WASHBURN, S.; RAMLACKHAN, K.; MCLEOD, D. Quality-of-service perceptions by rural freeway travelers: exploratory analysis. *Transportation Research Record*, v. 1883, n. 1, p. 132–139, 2004. Available at: https://doi.org/10.3141/1883-15.

36 PAIVA, A. P. O.; SETTI, J. R. Um método de delimitação de niveis de serviço com base na percepção dos usuários. *In*: *Anais do XXIX Congresso Nacional de Pesquisa em Transporte - ANPET*. Ouro Preto, MG: Associação Nacional de Pesquisa e Ensino em Transportes (ANPET), 2015.

37 JENSEN, S. U. Car drivers' experienced level of service on freeways. *Transportation Research Record*, v. 2615, p. 132–139, jan 2017. Available at: https://doi.org/10.3141/2615-15.

38 CAMPOS, C. I. de; MARCOMINI, L. A.; PANICE, N. R.; PIVA, F. J.; LAROCCA, A. P. C. Perception analysis of highway quality of service using a driving simulator and eye tracking system. *Transportes*, v. 28, n. 3, p. 165–179, 2020. Available at: https://doi.org/10.14295/transportes.v28i3.2015.

39 KITA, H. Level-of-service measure of road traffic based on the driver's perception. *In*: *Transportation Research Circular E-C018: Proceedings of the Fourth International Symposium on Highway Capacity*. Maui, Hawaii, USA: TRB, 2000. p. 53–62. Available at: https://onlinepubs.trb.org/onlinepubs/circulars/ec018/05_22.pdf.

40 WANG, Z.; ZHANG, M.; LIU, H. A utility-based method for urban transportation system multi-modal level of service evaluation. *In*: *CICTP 2014*. Reston, VA: American Society of Civil Engineers, 2014. p. 3052–3065. Available at: https://doi.org/10.1061/9780784413623.292.

41 HALL, F.; WAKEFIELD, S.; AL-KAISY, A. Freeway quality of service: What really matters to drivers and passengers? *Transportation Research Record*, v. 1776, n. 1776, p. 17–23, 2001. Available at: https://doi.org/10.3141/1776-03.

42 HOSTOVSKY, C.; HALL, F. Freeway quality of service: Perceptions from tractor-trailer drivers. *Transportation Research Record*, v. 1852, n. 1852, p. 19–25, 2003. Available at: https://doi.org/10.3141/1852-03.

43 HOSTOVSKY, C.; WAKEFIELD, S.; HALL, F. Freeway users' perceptions of quality of service: Comparison of three groups. *Transportation Research Record*, v. 1883, p. 150–157, 2004. Available at: https://doi.org/10.3141/1883-17.

44 KO, B.; WASHBURN, S.; MCLEOD, D. Performance measures for truck level of service. *Transportation Research Record*, v. 2130, p. 120–128, 2009. Available at: https://doi.org/10.3141/2130-15.

45 NAKAMURA, H.; SUZUKI, K.; RYU, S. Analysis of the interrelationship among traffic flow conditions, driving behavior, and degree of driver's satisfaction on rural motorways. *In*: *Transportation Research Circular E-C018: Proceedings of the Fourth International Symposium on Highway Capacity*. Washington, D.C., EUA: Transportation Research Board, 2000. p. 42–52.

46 MORRISS, J. L. *Identification of preferred performance measures for the assessment of level of service on two-lane highways*. 2005. 134 p. Dissertation (PhD) — University of Florida, 2005. Available at: https://ufl-flvc.primo.exlibrisgroup.com/permalink/01FALSC_UFL/175ga98/alma990280920850306597.

47 FANG, F. C.; PECHEUX, K. K. Fuzzy data mining approach for quantifying signalized intersection level of services based on user perceptions. *Journal of Transportation Engineering*, v. 135, n. 6, p. 349–358, 2009. Available at: https://doi.org/10.1061/(ASCE)0733-947X(2009)135:6(349).

48 PIVA, F. J.; SETTI, J. R.; WASHBURN, S. Using traffic simulation for level of service traveller perception studies. *Promet – Traffic&Transportation*, v. 34, n. 2, p. 297–308, 2022. Available at: https://doi.org/10.7307/ptt.v34i2.3965.

49 DE VOS, J.; WITLOX, F. Travel satisfaction revisited. On the pivotal role of travel satisfaction in conceptualising a travel behaviour process. *Transportation Research Part A: Policy and Practice*, v. 106, p. 364–373, 2017. Available at: https://doi.org/10.1016/j.tra.2017.10.009.

50 ETTEMA, D.; GÄRLING, T.; ERIKSSON, L.; FRIMAN, M.; OLSSON, L. E.; FUJII, S. Satisfaction with travel and subjective well-being: Development and test of a measurement tool. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 14, n. 3, p. 167–175, 2011. Available at: https://doi.org/10.1016/j.trf.2010.11.002.

51 SINGLETON, P. A. Validating the satisfaction with travel scale as a measure of hedonic subjective well-being for commuting in a U.S. city. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 60, p. 399–414, 2019. Available at: https://doi.org/10.1016/j.trf.2018.10.029.

52 KITA, H.; KOUCHI, A. Quantifying perceived quality of traffic service and its aggregation structure. *Transportation Research Part C: Emerging Technologies*, v. 19, n. 2, p. 296–306, 2011. Available at: https://doi.org/10.1016/j.trc.2010.05.015.

53 JOU, R.-C.; KOU, C.-C.; CHEN, Y.-W. Drivers' perception of LOSs at signalised intersections. *Transportation Research Part A: Policy and Practice*, v. 54, p. 141–154, 2013. Available at: https://doi.org/10.1016/j.tra.2013.07.013.

54 JOU, R.-C.; CHEN, Y.-W. Drivers' acceptance of delay time at different levels of service at signalised intersections. *Transportation Research Part A: Policy and Practice*, v. 58, p. 54–66, 2013. Available at: https://doi.org/10.1016/j.tra.2013.10.009.

55 KANG, L.; XIONG, Y.; MANNERING, F. L. Statistical analysis of pedestrian perceptions of sidewalk level of service in the presence of bicycles. *Transportation Research Part A: Policy and Practice*, v. 53, p. 10–21, 2013. Available at: https://doi.org/10.1016/j.tra.2013.05.002.

56 GRISWOLD, J. B.; YU, M.; FILINGERI, V.; GREMBEK, O.; WALKER, J. L. A behavioral modeling approach to bicycle level of service. *Transportation Research Part A: Policy and Practice*, v. 116, p. 166–177, 2018. Available at: https://doi.org/10.1016/j.tra.2018.06.006.

57 WANG, X. Integrating GIS, simulation models, and visualization in traffic impact analysis. *Computers, Environment and Urban Systems*, v. 29, n. 4, p. 471–496, 2005. Available at: https://doi.org/10.1016/j.compenvurbsys.2004.01.002.

58 KULIGA, S.; THRASH, T.; DALTON, R.; HöLSCHER, C. Virtual reality as an empirical research tool — exploring user experience in a real building and a corresponding virtual model. *Computers, Environment and Urban Systems*, v. 54, p. 363–375, 2015. Available at: https://doi.org/10.1016/j.compenvurbsys.2015.09.006.

59 FUNKE, F.; REIPS, U.-D. Why semantic differentials in web-based research should be made from visual analogue scales and not from 5-point scales. *Field Methods*, v. 24, n. 3, p. 310–327, 2012. Available at: https://doi.org/10.1177/1525822X12444061.

60 PTV. *VISSIM version 11*. 2020.

61 BETHONICO, F. C.; PIVA, J. F.; SETTI, J. R. Calibração de simuladores microscópicos de tráfego através de medidas macroscópicas. *In*: *Anais do XXX Congresso Nacional de Pesquisa em Transporte*. Rio de Janeiro, RJ: Associação Nacional de Pesquisa e Ensino em Transportes (ANPET), 2016.

62 AASHTO. *A policy on geometric design of highways and streets*. Washington, DC: American Association of State Highway and Transportation Officials, 2018.

63 PIVA, F. J. *Supplemental data video S1*. Zenodo, 2021. Available at: https://doi.org/10.5281/zenodo.4499031.

64 HOHMANN, S.; GEISTEFELDT, J. Traffic flow quality from the user's perspective. *Transportation Research Procedia*, Elsevier, v. 15, p. 721–731, 2016. Available at: https://doi.org/10.1016/j.trpro.2016.06.060.

65 TORFS, K.; MEESMANN, U.; BERGHE, W. Van den; TROTTA, M. *ESRA 2015 - The results. Synthesis of the main findings from the ESRA survey in 17 countries*. Brussels, Belgium, 2016. Available at: https://www.esranet.eu/storage/minisites/esra2015results.pdf.

66 XU, X.; FAN, C.-K. Autonomous vehicles, risk perceptions and insurance demand: An individual survey in China. *Transportation Research Part A: Policy and Practice*, v. 124, p. 549–556, 2019. Available at: https://doi.org/10.1016/j.tra.2018.04.009.

67 CAMPOS, C. I. de; PITOMBO, C. S.; DELHOMME, P.; QUINTANILHA, J. A. Comparative analysis of data reduction techniques for questionnaire validation using self-reported driver behaviors. *Journal of Safety Research*, Elsevier, v. 73, p. 133–142, 2020. Available at: https://doi.org/10.1016/j.jsr.2020.02.004.

68 STEINBAKK, R. T.; ULLEBERG, P.; SAGBERG, F.; FOSTERVOLD, K. I. Analysing the influence of visible roadwork activity on drivers' speed choice at work zones using a video-based experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 44, p. 53–62, 2017. Available at: https://doi.org/10.1016/j.trf.2016.10.003.

69 KNOOP, V. L.; KEYVAN-EKBATANI, M.; BAAT, M. de; TAALE, H.; HOOGENDOORN, S. P. Lane change behavior on freeways: an online survey using video clips. *Journal of Advanced Transportation*, v. 2018, 2018. Available at: https://doi.org/10.1155/2018/9236028.

70 LI, J.; CURRANO, R.; SIRKIN, D.; GOEDICKE, D.; TENNENT, H.; LEVINE, A.; EVERS, V.; JU, W. On-road and online studies to investigate beliefs and behaviors of Netherlands, US and Mexico pedestrians encountering hidden-driver vehicles. *In*: *HRI 20: ACM/IEEE International Conference on Human-Robot Interaction*. Cambridge, UK: Association for Computing Machinery, 2020. p. 141–149. Available at: https://doi.org/10.1145/3319502.3374790.

71 YE, X.; CHEN, J.; JIANG, G.; YAN, X. Modeling pedestrian level of service at signalized intersection crosswalks under mixed traffic conditions. *Transportation Research Record*, v. 2512, n. 1, p. 46–55, 2015. Available at: https://doi.org/10.3141/2512-06.

72 ZHANG, W.; WANG, K.; WANG, L.; FENG, Z.; DU, Y. Exploring factors affecting pedestrians' red-light running behaviors at intersections in China. *Accident Analysis & Prevention*, v. 96, p. 71–78, 2016. Available at: https://doi.org/10.1016/j.aap.2016.07.038.

73 PARK, M.; JANG, K.; LEE, J.; YEO, H. Logistic regression model for discretionary lane changing under congested traffic. *Transportmetrica A: Transport Science*, v. 11, n. 4, p. 333–344, 2015. Available at: https://doi.org/10.1080/23249935.2014.994686.

74 NG, C.; SUSILAWATI, S.; KAMAL, M. A. S.; CHEW, I. M. L. Development of a binary logistic lane change model and its validation using empirical freeway data. *Transportmetrica B: Transport Dynamics*, v. 8, n. 1, p. 49–71, 2020. Available at: https://doi.org/10.1080/21680566.2020.1715309.

75 PAPAIOANNOU, D.; MARTÍNEZ, L. M. Measuring satisfaction with transit and car trips with use of one logistic regression model. *Transportation Research Record*, v. 2543, n. 1, p. 101–107, 2016. Available at: https://doi.org/10.3141/2543-11.

76 TAN, L.; MA, C. Choice behavior of commuters' rail transit mode during the covid-19 pandemic based on logistic model. *Journal of Traffic and Transportation Engineering (English Edition)*, v. 8, n. 2, p. 186–195, 2021. Available at: https://doi.org/10.1016/j.jtte.2020.07.002.

77 WU, C. J.; HAMADA, M. S. *Experiments: planning, analysis, and optimization.* 3rd edition. Hoboken, NJ, USA: John Wiley & Sons, 2021.

78 CRAWFORD, F.; WATLING, D.; CONNORS, R. Identifying road user classes based on repeated trip behaviour using Bluetooth data. *Transportation Research Part A: Policy and Practice*, v. 113, p. 55–74, 2018. Available at: https://doi.org/10.1016/j.tra.2018.03.027.

79 ALSALEH, N.; FAROOQ, B. Interpretable data-driven demand modelling for on-demand transit services. *Transportation Research Part A: Policy and Practice*, v. 154, p. 1–22, 2021. Available at: https://doi.org/10.1016/j.tra.2021.10.001.

80 LI, X.; TANG, J.; HU, X.; WANG, W. Assessing intercity multimodal choice behavior in a touristy city: a factor analysis. *Journal of Transport Geography*, v. 86, 2020. Available at: https://doi.org/10.1016/j.jtrangeo.2020.102776.

81 JIA, X.; FELICIANI, C.; MURAKAMI, H.; NAGAHAMA, A.; YANAGISAWA, D.; NISHINARI, K. Revisiting the level-of-service framework for pedestrian comfortability: velocity depicts more accurate perceived congestion than local density. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 87, p. 403–425, 2022. Available at: https://doi.org/10.1016/j.trf.2022.04.007.

82 BIRAGO, D.; MENSAH, S. O.; SHARMA, S. Level of service delivery of public transport and mode choice in Accra, Ghana. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 46, p. 284–300, 2017. Available at: https://doi.org/10.1016/j.trf.2016.09.033.

83 ABENOZA, R. F.; CATS, O.; SUSILO, Y. O. Determinants of traveler satisfaction: Evidence for non-linear and asymmetric effects. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 66, p. 339–356, 2019. Available at: https://doi.org/10.1016/j.trf.2019.09.009.

84 RODGERS, J. L.; NICEWANDER, W. A. Thirteen ways to look at the correlation coefficient. *The American Statistician*, v. 42, n. 1, p. 59–66, 1988. Available at: https://doi.org/10.1080/00031305.1988.10475524.

85 FENG, M.; WANG, X.; LEE, J.; ABDEL-ATY, M.; MAO, S. Transferability of safety performance functions and hotspot identification for freeways of the United States and China. *Accident Analysis & Prevention*, v. 139, p. 105493, 2020. Available at: https://doi.org/10.1016/j.aap.2020.105493.

86 KYRIAKIDIS, M.; HAPPEE, R.; WINTER, J. C. de. Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 32, p. 127–140, 2015. Available at: https://doi.org/10.1016/j.trf.2015.04.014.

87 COHEN, J.; COHEN, P.; WEST, S. G.; AIKEN, L. S. *Applied multiple regression/correlation analysis for the behavioral sciences.* 3rd ed. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2002. Available at: https://doi.org/10.4324/9780203774441.

88 YI, B.; CAO, H.; SONG, X.; ZHAO, S.; GUO, W.; LI, M. How to identify the take-over criticality in conditionally automated driving? an examination using drivers' physiological parameters and situational factors. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 85, p. 161–178, 2022. Available at: https://doi.org/10.1016/j.trf.2021.12.007.

89 ELHADIDY, A. A.; EL-BADAWY, S. M.; ELBELTAGI, E. E. A simplified pavement condition index regression model for pavement evaluation. *International Journal of Pavement Engineering*, v. 22, n. 5, p. 643–652, 2021. Available at: https://doi.org/10.1080/10298436.2019.1633579.

90 YEBOAH, A. S.; CODJOE, J.; THAPA, R. Estimating average daily traffic on low-volume roadways in Louisiana. *Transportation Research Record*, p. 03611981221106166, 2022. Available at: https://doi.org/10.1177/03611981221106166.

91 WU, J.; XU, H. Annual average daily traffic prediction model for minor roads at intersections. *Journal of Transportation Engineering, Part A: Systems*, v. 145, n. 10, p. 04019041–1–04019041–9, 2019. Available at: https://doi.org/10.1061/JTEPBS.0000262.

92 ASGHARZADEH, M.; GUBBALA, P. S.; KONDYLI, A.; SCHROCK, S. D. Effect of on-ramp demand and flow distribution on capacity at merge bottleneck locations. *Transportation Letters*, v. 12, n. 8, p. 550–558, 2020. Available at: https://doi.org/10.1080/19427867.2019.1665774.

93 NAVIDI, W. C. *Statistics for engineers and scientists*. 5th edition. New York, NY, USA: McGraw-Hill Higher Education, 2020.

94 GHASEMZADEH, A.; HAMMIT, B. E.; AHMED, M. M.; YOUNG, R. K. Parametric ordinal logistic regression and non-parametric decision tree approaches for assessing the impact of weather conditions on driver speed selection using naturalistic driving data. *Transportation Research Record*, v. 2672, n. 12, p. 137–147, 2018. Available at: https://doi.org/10.1177/0361198118758035.

95 BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. *Classification and regression trees*. Boca Raton: Chapman & Hall/CRC, 1984. Available at: https://doi.org/10.1201/9781315139470.

96 OÑA, J. de; OÑA, R. de; CALVO, F. J. A classification tree approach to identify key factors of transit service quality. *Expert Systems with Applications*, v. 39, n. 12, p. 11164–11171, 2012. Available at: https://doi.org/10.1016/j.eswa.2012.03.037.

97 FIGUEIRA, A. C.; LAROCCA, A. P. C. Proposal of a driver profile classification in relation to risk level in overtaking maneuvers. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 74, p. 375–385, 2020. Available at: https://doi.org/10.1016/j.trf.2020.08.012.

98 LIU, M.; WU, Z.; CHEN, Y.; ZHANG, X. Utilizing decision tree method and ANFIS to explore real-time crash risk for urban freeways. *In*: *CICTP 2020*. Xi'an, China: ASCE, 2020. p. 2495–2508. Available at: https://doi.org/10.1061/9780784483053.211.

99 ABELLÁN, J.; LÓPEZ, G.; OÑA, J. D. Analysis of traffic accident severity using decision rules via decision trees. *Expert Systems with Applications*, v. 40, n. 15, p. 6047–6054, 2013. Available at: https://doi.org/10.1016/j.eswa.2013.05.027.

100 OÑA, J. de; LÓPEZ, G.; ABELLÁN, J. Extracting decision rules from police accident reports through decision trees. *Accident Analysis & Prevention*, v. 50, p. 1151–1160, 2013. Available at: https://doi.org/10.1016/j.aap.2012.09.006.

101 BAZILINSKYY, P.; KYRIAKIDIS, M.; DODOU, D.; WINTER, J. de. When will most cars be able to drive fully automatically? projections of 18,970 survey respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, v. 64, p. 184–195, 2019. Available at: https://doi.org/10.1016/j.trf.2019.05.008.

102 BOX, G. E.; HUNTER, W. G.; HUNTER, J. S. *Statistics for experimenters: design, discovery, and innovation*. 2. ed. Hoboken, NJ, USA: John Wiley & Sons, 2005.

103 MA, X. *Using classification and regression trees – A practical primer*. Charlotte, NC: Information Age Publishing, 2018.

APPENDIX

Fitted logit models

A

Appendix (A) shows the coefficients from SPSS for the fitted logit functions, including Nagelkerke's ρ^2, and the coefficients are significant at $\alpha = 5\%$. Table 16 presents the coefficients calculated for five levels of service and Table 17 included the results for four levels of services.

Table 16 – SPSS results for fitted binary logit models considering five levels of service (A to E)

$P[\ell = A](k)$								
Nagelkerke's $\rho^2 = 0.551$							CI$_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	exp(B)	Lower	Upper
β_1	−0.3362	0.0099	1158.50	1	< 0.001	0.7145	0.7008	0.7284
β_0	2.2823	0.0901	642.02	1	< 0.001	9.7991		
$P[\ell = A \text{ or } B](k)$								
Nagelkerke's $\rho^2 = 0.527$							CI$_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	exp(B)	Lower	Upper
β_1	−0.2621	0.0068	1488.01	1	< 0.001	0.7694	0.7593	0.7798
β_0	3.0467	0.0888	1177.61	1	< 0.001	21.0462		
$P[\ell = A \text{ or } B \text{ or } C](k)$								
Nagelkerke's $\rho^2 = 0.394$							CI$_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	exp(B)	Lower	Upper
β_1	−0.1916	0.0053	1297.84	1	< 0.001	0.8257	0.8171	0.8343
β_0	3.3057	0.0887	1387.55	1	< 0.001	27.2681		
$P[\ell = A \text{ or } B \text{ or } C \text{ or } D](k)$								
Nagelkerke's $\rho^2 = 0.216$							CI$_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	exp(B)	Lower	Upper
β_1	−0.1464	0.0060	593.87	1	< 0.001	0.8638	0.8537	0.8741
β_0	4.0769	0.1194	1166.39	1	< 0.001	58.9620		

Table 17 – SPSS results for fitted binary logit models considering four levels of service (A', B', C' or D')

$P[\ell = A'](k)$							
Nagelkerke's $\rho^2 = 0.570$						$\mathrm{CI}_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
β_1	−0.329	0.009	1286.600	1	< 0.001	0.720	0.707	0.733
β_0	2.634	0.092	823.363	1	< 0.001	13.931		

$P[\ell = A' \text{ or } B'](k)$								
Nagelkerke's $\rho^2 = 0.483$							$\mathrm{CI}_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
β_1	−0.231	0.006	1481.328	1	< 0.001	0.794	0.785	0.803
β_0	3.393	0.091	1383.742	1	< 0.001	29.768		

$P[\ell = A' \text{ or } B' \text{ or } C'](k)$								
Nagelkerke's $\rho^2 = 0.267$							$\mathrm{CI}_{95\%}$ for exp(B)	
	B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
β_1	−0.159	0.006	795.170	1	< 0.001	0.853	0.844	0.863
β_0	3.923	0.109	1293.770	1	< 0.001	50.551		

APPENDIX B

Design and planning matrices

The present appendix presents the complete matrices used to create the fractional factorial design. The design matrix used in the fractional factorial experiment is shown in Table 18. Table 19 provides the design matrix combining the 7 columns for 128 runs and pseudo-factors $C2$, D and E and Table 20 contains the planning matrix for the 128 scenarios used for the creation of the video clips.

Table 18 – Complete design matrix for 128 runs, created with $2^{(column)}$ blocks of $2^{(column-1)}$ of −1 and +1 in each column

	Columns						
Runs	1	2	3	4	5	6	7
1	−1	−1	−1	−1	−1	−1	−1
2	−1	−1	−1	−1	−1	−1	+1
3	−1	−1	−1	−1	−1	+1	−1
4	−1	−1	−1	−1	−1	+1	+1
5	−1	−1	−1	−1	+1	−1	−1
6	−1	−1	−1	−1	+1	−1	+1
7	−1	−1	−1	−1	+1	+1	−1
8	−1	−1	−1	−1	+1	+1	+1
9	−1	−1	−1	+1	−1	−1	−1
10	−1	−1	−1	+1	−1	−1	+1
11	−1	−1	−1	+1	−1	+1	−1
12	−1	−1	−1	+1	−1	+1	+1
13	−1	−1	−1	+1	+1	−1	−1
14	−1	−1	−1	+1	+1	−1	+1
15	−1	−1	−1	+1	+1	+1	−1

(Continued on next page)

Table 18 – continued from previous page

	Columns						
Runs	1	2	3	4	5	6	7
16	−1	−1	−1	+1	+1	+1	+1
17	−1	−1	+1	−1	−1	−1	−1
18	−1	−1	+1	−1	−1	−1	+1
19	−1	−1	+1	−1	−1	+1	−1
20	−1	−1	+1	−1	−1	+1	+1
21	−1	−1	+1	−1	+1	−1	−1
22	−1	−1	+1	−1	+1	−1	+1
23	−1	−1	+1	−1	+1	+1	−1
24	−1	−1	+1	−1	+1	+1	+1
25	−1	−1	+1	+1	−1	−1	−1
26	−1	−1	+1	+1	−1	−1	+1
27	−1	−1	+1	+1	−1	+1	−1
28	−1	−1	+1	+1	−1	+1	+1
29	−1	−1	+1	+1	+1	−1	−1
30	−1	−1	+1	+1	+1	−1	+1
31	−1	−1	+1	+1	+1	+1	−1
32	−1	−1	+1	+1	+1	+1	+1
33	−1	+1	−1	−1	−1	−1	−1
34	−1	+1	−1	−1	−1	−1	+1
35	−1	+1	−1	−1	−1	+1	−1
36	−1	+1	−1	−1	−1	+1	+1
37	−1	+1	−1	−1	+1	−1	−1
38	−1	+1	−1	−1	+1	−1	+1
39	−1	+1	−1	−1	+1	+1	−1
40	−1	+1	−1	−1	+1	+1	+1
41	−1	+1	−1	+1	−1	−1	−1
42	−1	+1	−1	+1	−1	−1	+1
43	−1	+1	−1	+1	−1	+1	−1
44	−1	+1	−1	+1	−1	+1	+1
45	−1	+1	−1	+1	+1	−1	−1
46	−1	+1	−1	+1	+1	−1	+1
47	−1	+1	−1	+1	+1	+1	−1
48	−1	+1	−1	+1	+1	+1	+1
49	−1	+1	+1	−1	−1	−1	−1
50	−1	+1	+1	−1	−1	−1	+1
51	−1	+1	+1	−1	−1	+1	−1
52	−1	+1	+1	−1	−1	+1	+1
53	−1	+1	+1	−1	+1	−1	−1
54	−1	+1	+1	−1	+1	−1	+1
55	−1	+1	+1	−1	+1	+1	−1
56	−1	+1	+1	−1	+1	+1	+1
57	−1	+1	+1	+1	−1	−1	−1
58	−1	+1	+1	+1	−1	−1	+1
59	−1	+1	+1	+1	−1	+1	−1
60	−1	+1	+1	+1	−1	+1	+1
61	−1	+1	+1	+1	+1	−1	−1
62	−1	+1	+1	+1	+1	−1	+1
63	−1	+1	+1	+1	+1	+1	−1
64	−1	+1	+1	+1	+1	+1	+1
65	+1	−1	−1	−1	−1	−1	−1
66	+1	−1	−1	−1	−1	−1	+1

(Continued on next page)

Table 18 – continued from previous page

	Columns						
Runs	1	2	3	4	5	6	7
67	+1	−1	−1	−1	−1	+1	−1
68	+1	−1	−1	−1	−1	+1	+1
69	+1	−1	−1	−1	+1	−1	−1
70	+1	−1	−1	−1	+1	−1	+1
71	+1	−1	−1	−1	+1	+1	−1
72	+1	−1	−1	−1	+1	+1	+1
73	+1	−1	−1	+1	−1	−1	−1
74	+1	−1	−1	+1	−1	−1	+1
75	+1	−1	−1	+1	−1	+1	−1
76	+1	−1	−1	+1	−1	+1	+1
77	+1	−1	−1	+1	+1	−1	−1
78	+1	−1	−1	+1	+1	−1	+1
79	+1	−1	−1	+1	+1	+1	−1
80	+1	−1	−1	+1	+1	+1	+1
81	+1	−1	+1	−1	−1	−1	−1
82	+1	−1	+1	−1	−1	−1	+1
83	+1	−1	+1	−1	−1	+1	−1
84	+1	−1	+1	−1	−1	+1	+1
85	+1	−1	+1	−1	+1	−1	−1
86	+1	−1	+1	−1	+1	−1	+1
87	+1	−1	+1	−1	+1	+1	−1
88	+1	−1	+1	−1	+1	+1	+1
89	+1	−1	+1	+1	−1	−1	−1
90	+1	−1	+1	+1	−1	−1	+1
91	+1	−1	+1	+1	−1	+1	−1
92	+1	−1	+1	+1	−1	+1	+1
93	+1	−1	+1	+1	+1	−1	−1
94	+1	−1	+1	+1	+1	−1	+1
95	+1	−1	+1	+1	+1	+1	−1
96	+1	−1	+1	+1	+1	+1	+1
97	+1	+1	−1	−1	−1	−1	−1
98	+1	+1	−1	−1	−1	−1	+1
99	+1	+1	−1	−1	−1	+1	−1
100	+1	+1	−1	−1	−1	+1	+1
101	+1	+1	−1	−1	+1	−1	−1
102	+1	+1	−1	−1	+1	−1	+1
103	+1	+1	−1	−1	+1	+1	−1
104	+1	+1	−1	−1	+1	+1	+1
105	+1	+1	−1	+1	−1	−1	−1
106	+1	+1	−1	+1	−1	−1	+1
107	+1	+1	−1	+1	−1	+1	−1
108	+1	+1	−1	+1	−1	+1	+1
109	+1	+1	−1	+1	+1	−1	−1
110	+1	+1	−1	+1	+1	−1	+1
111	+1	+1	−1	+1	+1	+1	−1
112	+1	+1	−1	+1	+1	+1	+1
113	+1	+1	+1	−1	−1	−1	−1
114	+1	+1	+1	−1	−1	−1	+1
115	+1	+1	+1	−1	−1	+1	−1
116	+1	+1	+1	−1	−1	+1	+1
117	+1	+1	+1	−1	+1	−1	−1

(Continued on next page)

Table 18 – continued from previous page

	Columns						
Runs	1	2	3	4	5	6	7
118	+1	+1	+1	−1	+1	−1	+1
119	+1	+1	+1	−1	+1	+1	−1
120	+1	+1	+1	−1	+1	+1	+1
121	+1	+1	+1	+1	−1	−1	−1
122	+1	+1	+1	+1	−1	−1	+1
123	+1	+1	+1	+1	−1	+1	−1
124	+1	+1	+1	+1	−1	+1	+1
125	+1	+1	+1	+1	+1	−1	−1
126	+1	+1	+1	+1	+1	−1	+1
127	+1	+1	+1	+1	+1	+1	−1
128	+1	+1	+1	+1	+1	+1	+1

Table 19 – Complete design matrix combining the 7 columns for 128 runs and pseudofactors $C2$, D and E

Runs	$A1$	$A2$	$A3$	$A4$	$B1$	$B2$	$C1$	$C2$	D	E
1	−1	−1	−1	−1	−1	−1	−1	+1	+1	+1
2	−1	−1	−1	−1	−1	−1	+1	−1	−1	+1
3	−1	−1	−1	−1	−1	+1	−1	+1	−1	−1
4	−1	−1	−1	−1	−1	+1	+1	−1	+1	−1
5	−1	−1	−1	−1	+1	−1	−1	−1	+1	+1
6	−1	−1	−1	−1	+1	−1	+1	+1	−1	+1
7	−1	−1	−1	−1	+1	+1	−1	−1	−1	−1
8	−1	−1	−1	−1	+1	+1	+1	+1	+1	−1
9	−1	−1	−1	+1	−1	−1	−1	−1	−1	−1
10	−1	−1	−1	+1	−1	−1	+1	+1	+1	−1
11	−1	−1	−1	+1	−1	+1	−1	−1	+1	+1
12	−1	−1	−1	+1	−1	+1	+1	+1	−1	+1
13	−1	−1	−1	+1	+1	−1	−1	+1	−1	−1
14	−1	−1	−1	+1	+1	−1	+1	−1	+1	−1
15	−1	−1	−1	+1	+1	+1	−1	+1	+1	+1
16	−1	−1	−1	+1	+1	+1	+1	−1	−1	+1
17	−1	−1	+1	−1	−1	−1	−1	+1	+1	−1
18	−1	−1	+1	−1	−1	−1	+1	−1	−1	−1
19	−1	−1	+1	−1	−1	+1	−1	+1	−1	+1
20	−1	−1	+1	−1	−1	+1	+1	−1	+1	+1
21	−1	−1	+1	−1	+1	−1	−1	−1	+1	−1
22	−1	−1	+1	−1	+1	−1	+1	+1	−1	−1
23	−1	−1	+1	−1	+1	+1	−1	−1	−1	+1
24	−1	−1	+1	−1	+1	+1	+1	+1	+1	+1
25	−1	−1	+1	+1	−1	−1	−1	−1	−1	+1
26	−1	−1	+1	+1	−1	−1	+1	+1	+1	+1
27	−1	−1	+1	+1	−1	+1	−1	−1	+1	−1
28	−1	−1	+1	+1	−1	+1	+1	+1	−1	−1
29	−1	−1	+1	+1	+1	−1	−1	+1	−1	+1
30	−1	−1	+1	+1	+1	−1	+1	−1	+1	+1
31	−1	−1	+1	+1	+1	+1	−1	+1	+1	−1

(Continued on next page)

Table 19 – continued from previous page

Runs	$A1$	$A2$	$A3$	$A4$	$B1$	$B2$	$C1$	$C2$	D	E
32	−1	−1	+1	+1	+1	+1	+1	−1	−1	−1
33	−1	+1	−1	−1	−1	−1	−1	+1	−1	+1
34	−1	+1	−1	−1	−1	−1	+1	−1	+1	+1
35	−1	+1	−1	−1	−1	+1	−1	+1	+1	−1
36	−1	+1	−1	−1	−1	+1	+1	−1	−1	−1
37	−1	+1	−1	−1	+1	−1	−1	−1	−1	+1
38	−1	+1	−1	−1	+1	−1	+1	+1	+1	+1
39	−1	+1	−1	−1	+1	+1	−1	−1	+1	−1
40	−1	+1	−1	−1	+1	+1	+1	+1	−1	−1
41	−1	+1	−1	+1	−1	−1	−1	−1	+1	−1
42	−1	+1	−1	+1	−1	−1	+1	+1	−1	−1
43	−1	+1	−1	+1	−1	+1	−1	−1	−1	+1
44	−1	+1	−1	+1	−1	+1	+1	+1	+1	+1
45	−1	+1	−1	+1	+1	−1	−1	+1	+1	−1
46	−1	+1	−1	+1	+1	−1	+1	−1	−1	−1
47	−1	+1	−1	+1	+1	+1	−1	+1	−1	+1
48	−1	+1	−1	+1	+1	+1	+1	−1	+1	+1
49	−1	+1	+1	−1	−1	−1	−1	+1	−1	−1
50	−1	+1	+1	−1	−1	−1	+1	−1	+1	−1
51	−1	+1	+1	−1	−1	+1	−1	+1	+1	+1
52	−1	+1	+1	−1	−1	+1	+1	−1	−1	+1
53	−1	+1	+1	−1	+1	−1	−1	−1	−1	−1
54	−1	+1	+1	−1	+1	−1	+1	+1	+1	−1
55	−1	+1	+1	−1	+1	+1	−1	−1	+1	+1
56	−1	+1	+1	−1	+1	+1	+1	+1	−1	+1
57	−1	+1	+1	+1	−1	−1	−1	−1	+1	+1
58	−1	+1	+1	+1	−1	−1	+1	+1	−1	+1
59	−1	+1	+1	+1	−1	+1	−1	−1	−1	−1
60	−1	+1	+1	+1	−1	+1	+1	+1	+1	−1
61	−1	+1	+1	+1	+1	−1	−1	+1	+1	+1
62	−1	+1	+1	+1	+1	−1	+1	−1	−1	+1
63	−1	+1	+1	+1	+1	+1	−1	+1	−1	−1
64	−1	+1	+1	+1	+1	+1	+1	−1	+1	−1
65	+1	−1	−1	−1	−1	−1	−1	−1	+1	−1
66	+1	−1	−1	−1	−1	−1	+1	+1	−1	−1
67	+1	−1	−1	−1	−1	+1	−1	−1	−1	+1
68	+1	−1	−1	−1	−1	+1	+1	+1	+1	+1
69	+1	−1	−1	−1	+1	−1	−1	+1	+1	−1
70	+1	−1	−1	−1	+1	−1	+1	−1	−1	−1
71	+1	−1	−1	−1	+1	+1	−1	+1	−1	+1
72	+1	−1	−1	−1	+1	+1	+1	−1	+1	+1
73	+1	−1	−1	+1	−1	−1	−1	+1	−1	+1
74	+1	−1	−1	+1	−1	−1	+1	−1	+1	+1
75	+1	−1	−1	+1	−1	+1	−1	+1	+1	−1
76	+1	−1	−1	+1	−1	+1	+1	−1	−1	−1
77	+1	−1	−1	+1	+1	−1	−1	−1	−1	+1
78	+1	−1	−1	+1	+1	−1	+1	+1	+1	+1
79	+1	−1	−1	+1	+1	+1	−1	−1	+1	−1
80	+1	−1	−1	+1	+1	+1	+1	+1	−1	−1
81	+1	−1	+1	−1	−1	−1	−1	−1	+1	+1
82	+1	−1	+1	−1	−1	−1	+1	+1	−1	+1
83	+1	−1	+1	−1	−1	+1	−1	−1	−1	−1

(Continued on next page)

Table 19 – continued from previous page

Runs	$A1$	$A2$	$A3$	$A4$	$B1$	$B2$	$C1$	$C2$	D	E
84	+1	−1	+1	−1	−1	+1	+1	+1	+1	−1
85	+1	−1	+1	−1	+1	−1	−1	+1	+1	+1
86	+1	−1	+1	−1	+1	−1	+1	−1	−1	+1
87	+1	−1	+1	−1	+1	+1	−1	+1	−1	−1
88	+1	−1	+1	−1	+1	+1	+1	−1	+1	−1
89	+1	−1	+1	+1	−1	−1	−1	+1	−1	−1
90	+1	−1	+1	+1	−1	−1	+1	−1	+1	−1
91	+1	−1	+1	+1	−1	+1	−1	+1	+1	+1
92	+1	−1	+1	+1	−1	+1	+1	−1	−1	+1
93	+1	−1	+1	+1	+1	−1	−1	−1	−1	−1
94	+1	−1	+1	+1	+1	−1	+1	+1	+1	−1
95	+1	−1	+1	+1	+1	+1	−1	−1	+1	+1
96	+1	−1	+1	+1	+1	+1	+1	+1	−1	+1
97	+1	+1	−1	−1	−1	−1	−1	−1	−1	−1
98	+1	+1	−1	−1	−1	−1	+1	+1	+1	−1
99	+1	+1	−1	−1	−1	+1	−1	−1	+1	+1
100	+1	+1	−1	−1	−1	+1	+1	+1	−1	+1
101	+1	+1	−1	−1	+1	−1	−1	+1	−1	−1
102	+1	+1	−1	−1	+1	−1	+1	−1	+1	−1
103	+1	+1	−1	−1	+1	+1	−1	+1	+1	+1
104	+1	+1	−1	−1	+1	+1	+1	−1	−1	+1
105	+1	+1	−1	+1	−1	−1	−1	+1	+1	+1
106	+1	+1	−1	+1	−1	−1	+1	−1	−1	+1
107	+1	+1	−1	+1	−1	+1	−1	+1	−1	−1
108	+1	+1	−1	+1	−1	+1	+1	−1	+1	−1
109	+1	+1	−1	+1	+1	−1	−1	−1	+1	+1
110	+1	+1	−1	+1	+1	−1	+1	+1	−1	+1
111	+1	+1	−1	+1	+1	+1	−1	−1	−1	−1
112	+1	+1	−1	+1	+1	+1	+1	+1	+1	−1
113	+1	+1	+1	−1	−1	−1	−1	−1	−1	+1
114	+1	+1	+1	−1	−1	−1	+1	+1	+1	+1
115	+1	+1	+1	−1	−1	+1	−1	−1	+1	−1
116	+1	+1	+1	−1	−1	+1	+1	+1	−1	−1
117	+1	+1	+1	−1	+1	−1	−1	+1	−1	+1
118	+1	+1	+1	−1	+1	−1	+1	−1	+1	+1
119	+1	+1	+1	−1	+1	+1	−1	+1	+1	−1
120	+1	+1	+1	−1	+1	+1	+1	−1	−1	−1
121	+1	+1	+1	+1	−1	−1	−1	+1	+1	−1
122	+1	+1	+1	+1	−1	−1	+1	−1	−1	−1
123	+1	+1	+1	+1	−1	+1	−1	+1	−1	+1
124	+1	+1	+1	+1	−1	+1	+1	−1	+1	+1
125	+1	+1	+1	+1	+1	−1	−1	−1	+1	−1
126	+1	+1	+1	+1	+1	−1	+1	+1	−1	−1
127	+1	+1	+1	+1	+1	+1	−1	−1	−1	+1
128	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1

Table 20 – Complete planning matrix for the 128 scenarios used for the experiment

Scenario	Traffic density	Truck percent	Speed limit	Grade magnitude	Number of traffic lanes
1	1	1	2	2	2
2	1	1	3	1	2
3	1	2	2	1	1
4	1	2	3	2	1
5	1	3	1	2	2
6	1	3	4	1	2
7	1	4	1	1	1
8	1	4	4	2	1
9	2	1	1	1	1
10	2	1	4	2	1
11	2	2	1	2	2
12	2	2	4	1	2
13	2	3	2	1	1
14	2	3	3	2	1
15	2	4	2	2	2
16	2	4	3	1	2
17	3	1	2	2	1
18	3	1	3	1	1
19	3	2	2	1	2
20	3	2	3	2	2
21	3	3	1	2	1
22	3	3	4	1	1
23	3	4	1	1	2
24	3	4	4	2	2
25	4	1	1	1	2
26	4	1	4	2	2
27	4	2	1	2	1
28	4	2	4	1	1
29	4	3	2	1	2
30	4	3	3	2	2
31	4	4	2	2	1
32	4	4	3	1	1
33	5	1	2	1	2
34	5	1	3	2	2
35	5	2	2	2	1
36	5	2	3	1	1
37	5	3	1	1	2
38	5	3	4	2	2
39	5	4	1	2	1
40	5	4	4	1	1
41	6	1	1	2	1
42	6	1	4	1	1
43	6	2	1	1	2
44	6	2	4	2	2
45	6	3	2	2	1
46	6	3	3	1	1
47	6	4	2	1	2
48	6	4	3	2	2
49	7	1	2	1	1
50	7	1	3	2	1

(Continued on next page)

Table 20 – continued from previous page

Scenario	Traffic density	Truck percent	Speed limit	Grade magnitude	Number of traffic lanes
51	7	2	2	2	2
52	7	2	3	1	2
53	7	3	1	1	1
54	7	3	4	2	1
55	7	4	1	2	2
56	7	4	4	1	2
57	8	1	1	2	2
58	8	1	4	1	2
59	8	2	1	1	1
60	8	2	4	2	1
61	8	3	2	2	2
62	8	3	3	1	2
63	8	4	2	1	1
64	8	4	3	2	1
65	9	1	1	2	1
66	9	1	4	1	1
67	9	2	1	1	2
68	9	2	4	2	2
69	9	3	2	2	1
70	9	3	3	1	1
71	9	4	2	1	2
72	9	4	3	2	2
73	10	1	2	1	2
74	10	1	3	2	2
75	10	2	2	2	1
76	10	2	3	1	1
77	10	3	1	1	2
78	10	3	4	2	2
79	10	4	1	2	1
80	10	4	4	1	1
81	11	1	1	2	2
82	11	1	4	1	2
83	11	2	1	1	1
84	11	2	4	2	1
85	11	3	2	2	2
86	11	3	3	1	2
87	11	4	2	1	1
88	11	4	3	2	1
89	12	1	2	1	1
90	12	1	3	2	1
91	12	2	2	2	2
92	12	2	3	1	2
93	12	3	1	1	1
94	12	3	4	2	1
95	12	4	1	2	2
96	12	4	4	1	2
97	13	1	1	1	1
98	13	1	4	2	1
99	13	2	1	2	2
100	13	2	4	1	2
101	13	3	2	1	1
102	13	3	3	2	1

(Continued on next page)

Table 20 – continued from previous page

Scenario	Traffic density	Truck percent	Speed limit	Grade magnitude	Number of traffic lanes
103	13	4	2	2	2
104	13	4	3	1	2
105	14	1	2	2	2
106	14	1	3	1	2
107	14	2	2	1	1
108	14	2	3	2	1
109	14	3	1	2	2
110	14	3	4	1	2
111	14	4	1	1	1
112	14	4	4	2	1
113	15	1	1	1	2
114	15	1	4	2	2
115	15	2	1	2	1
116	15	2	4	1	1
117	15	3	2	1	2
118	15	3	3	2	2
119	15	4	2	2	1
120	15	4	3	1	1
121	16	1	2	2	1
122	16	1	3	1	1
123	16	2	2	1	2
124	16	2	3	2	2
125	16	3	1	2	1
126	16	3	4	1	1
127	16	4	1	1	2
128	16	4	4	2	2