TriDurLE

National Center for Transportation Infrastructure Durability & Life-Extension

Project ID: 2021-MST-01

Preparation of Pavement Infrastructure for Connected and Autonomous Vehicle Deployment – Phase I

Final Report

By

Xianbiao Hu, Chenxi Chen, and Yang Song

Pennsylvania State University

Jenny Liu and David Yizhuan Wang

Missouri University of Science and Technology

for

National University Transportation Center TriDurLE

Department of Civil & Environmental Engineering

405 Spokane Street PO Box 642910

Washington State University Pullman, WA 99164-2910

Date 8/31/2024

Table of Contents

Contents

List of Figures

List of Tables

Studies have been initiated to investigate the potential impact of Connected and Automated Vehicles (CAVs) on transportation infrastructure to support future CAV testing and deployment. However, most existing research only focuses on the different wandering patterns of CAV. To bridge this gap, an apple-to-apple comparison is first performed to systematically reveal the behavioral differences between the human-driven vehicle (HDV) and CAV trajectory patterns for the first time, with the data collected from camera-based Next Generation Simulation (NGSIM) dataset and autonomous driving co-simulation platform, CARLA and SUMO, respectively. A gradient boosting-based ensemble learning model for pavement performance (i.e., International Roughness Index, IRI) prediction is then developed with the input features including three driving pattern features, namely, lateral wandering deviation, longitudinal car-following distance, and driving speed, plus other twenty context variables. A total of 1,707 observations is extracted from the Long-Term Pavement Performance (LTPP) database for model training purposes. The result indicates that the trained model can accurately predict pavement deterioration, and that CAV deteriorates pavement faster than HDV by 8.1% on average. The results of sensitivity analysis show that CAV deployment will create a greater impact on the younger pavements, and the rate of pavement deterioration is found to be stable under light traffic, whereas it will increase under congested traffic.

Chapter 1. Introduction

1.1 Problem Background

Transportation infrastructure system networks are essential to sustain our economy, society, and quality of life. The Federal Highway Administration (FHWA) indicates the data-driven asset management (inspection, maintenance, and emergency response) of urban infrastructure (building, transportation, energy, and cyber) represents the future of work and identifies it as one of their strategic objectives (FHWA 2018). In the broad field of traffic operations, representative examples include the data-driven optimization of traffic signal infrastructure using reinforcement learning methods [1-5], the application of automated vehicles for highway infrastructure maintenance [6-9], and the development of electric vehicle charging infrastructure to promote sustainable transportation modes [10-12].

CAV technologies suggest great potentials for future transportation systems. When compared with HDVs, one fundamental difference is CAV's movements are controlled, and thus can be scheduled, by the computer algorithms with pre-defined objectives such as minimizing fuel consumptions or maximizing traffic throughputs [13-15]. Trajectory planning is the main task to achieve these goals [16]. It uses the on-board data (e.g., in-vehicle sensors, radar, camera, lidar, etc.) and the vehicle communication data (e.g. dedicated short-range communications, or Cellular, etc.) from the perception layer as the inputs for decision makings. Then, the computed CAV trajectory information is transferred to the control layer to move CAV forward [17]. While CAVs might be ten or fifteen years away from large-scale deployment in reality, considering the long-life span of transportation infrastructures, it might be beneficial for DOTs to start to prepare infrastructure for the testing and deployment of CAV.

1.2 Problem Statement

With the rapid advancement of connected and automated vehicle (CAV) technologies, the world is poised to witness a transformation in the way we travel. However, with this transformation, it is essential to ensure that the transportation infrastructure is prepared to accommodate the unique driving pattern of CAVs, which is shown to potentially impact a wide range of physical infrastructures from geometric design to pavement performance [18, 19]. Though a large-scale

deployment of CAV might be ten or fifteen years away, considering the long lifespan of physical transportation infrastructure like pavement and bridges, it might be beneficial to start preparing transportation infrastructure to support future CAV testing and deployment now. Motivated by these needs, this project focuses on modeling the impact of CAVs on the pavement with the goal of providing guidance for the management to understand deterioration patterns, the infrastructure maintenance and rehabilitation needs, and durability of road surfaces in the future for the new transportation era.

Existing research on the impact of CAVs on pavement performance, however, is limited and mostly focuses on those introduced by the different lateral wandering patterns. Research suggests that due to the latest vehicular technologies such as lane keeping and lane centering, the vehicle is usually placed at the centerline of a highway lane, and such centralized wheel wander distribution will inevitably lead to an acceleration of pavement deterioration [20-24]. In these research studies, the lateral wander is frequently assumed to be zero, meaning, CAVs will strictly follow the planned trajectory without any derivations. However, the assumption is unrealistic and may overestimate their impact on pavement distresses. Pavement design based on the overconservative assumptions can dramatically increase the cost of a paving project. While autonomous vehicles may be programmed to stay within their designated lanes and maintain a consistent trajectory, there are several factors that can cause lateral wandering, such as uneven road surfaces, strong crosswinds, and unexpected obstacles. Additionally, the sensors and algorithms used by autonomous vehicles to detect and respond to changes in the environment may not always be 100% accurate or reliable, which can also result in lateral wandering. Therefore, it is important to calibrate the degree of lateral wandering of autonomous vehicles and accommodate their patterns in the pavement deterioration model.

Additionally, some other driving features such as the inter-vehicle gap distance, as well as the driving speed may also impact pavement deterioration but received little attention in the literature. When vehicles follow too closely to one another, the reduced rest period between loads is likely to create additional ruts or permanent deformation in the pavement. A closer car-following distance may cause vehicles to brake more frequently. Heavy vehicles that are driving at a higher speed with more frequent braking and acceleration will cause higher pavement friction and can also contribute to pavement deterioration. As such, it becomes important to understand how the CAV technology impacts these driving behavior patterns, and subsequently, how they may impact

pavement deterioration.

To understand how CAV may impact pavement deterioration differently, an apple-to-apple comparison to systematically highlight the driving behavior patterns between CAV and HDV is a prerequisite. The "apple-to-apple comparison" in this situation means both CAV and HDV trajectories come from the same roadway segment with the same driving context, such as the number of lanes, lane width, and distance from upstream and downstream intersections or ramps. While some companies such as Waymo has made certain CAV dataset publicly available, the dataset is usually small in size and does not capture the overall traffic flow conditions (e.g., traffic volume). Further, in practice, it is challenging to have both datasets collected from the same roadway segment and with the same driving context.

To bridge these research gaps, the researchers in this study resort to Next Generation SIMulation (NGSIM) camera-based trajectory dataset from the Federal Highway Administration (FHWA) to extract HDV driving patterns, and develop an autonomous driving co-simulation platform, CARLA, coupled with SUMO, to build a CAV simulation scenario that mimics the NGSIM dataset environment (e.g., lane set up, traffic volume). Next, three features are defined to reflect the driving pattern differences between HDV and CAV, namely, lateral wandering deviation, longitudinal car-following distance, and driving speed, and use them to characterize the differences between the two datasets. Further, to assess the impact of CAV traffic patterns on pavement performance, a gradient boosting-based ensemble learning model is developed to predict the International Roughness Index (IRI) of asphalt pavements. In addition to the three above-mentioned features, twenty context variables are defined to describe the context environment of the surrounding area that the vehicle is in. The methodology of this study is summarized in the flow chart shown in Figure 1 below.

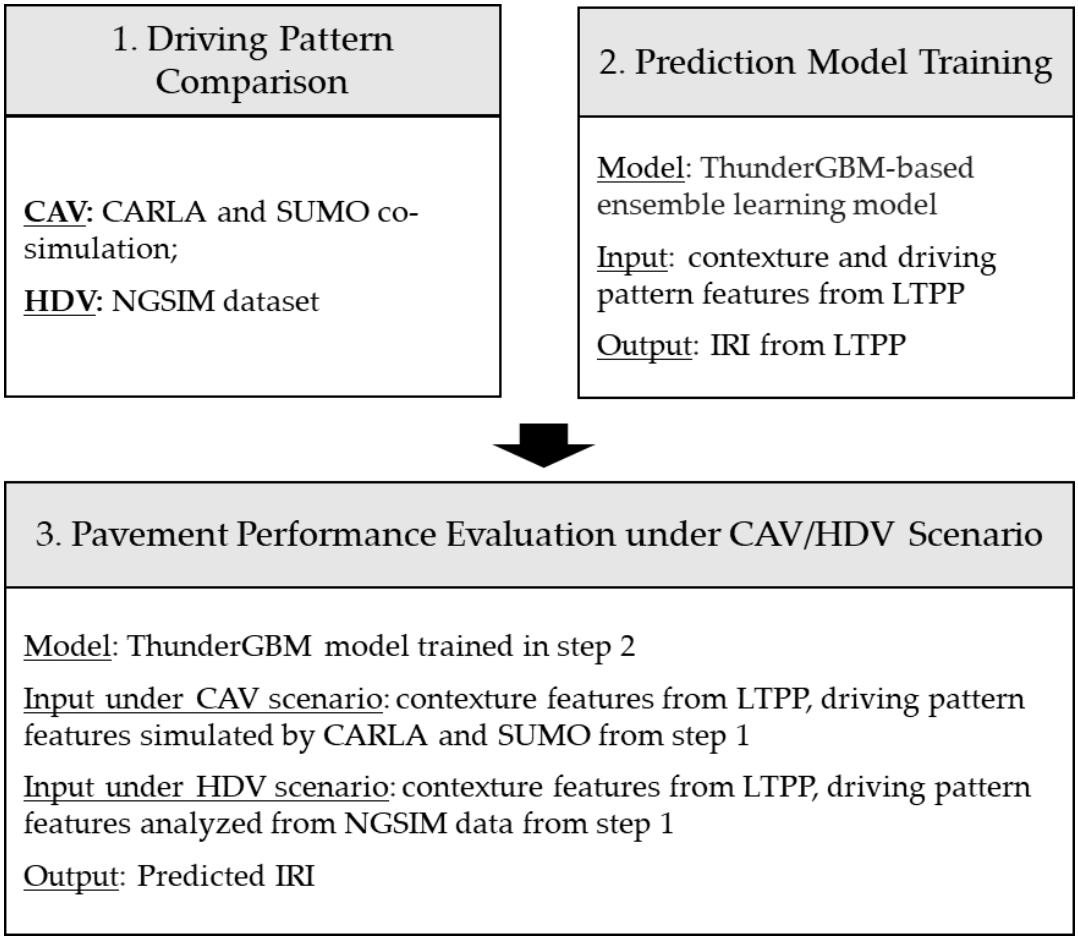


Figure 1. Flow chart of the methodology

The developed ensemble learning model is trained with the Long-Term Pavement Performance (LTPP) database, from which a total of 1,707 observations are extracted. 80% of the dataset is used for training and 20% dataset is used for testing purposes. Sensitivity analysis is also conducted and shows that the impact on pavement deterioration changes when the pavement age, truck volume, or CAV penetration rate changes. The results from this study will advance a fundamental understanding of how system-wide CAV deployment may impact pavement performance, and further, how to collect, analyze, maintain, and report pavement data in the current pavement management system to assist the decision-makers in finding optimum and cost-effective strategies for maintaining pavement infrastructure in serviceable condition.

This report is organized as follows. Section 2 systematically reviews relevant literature. Section 3 compares the driving pattern differences between HDV and CAV through an apple-to-apple comparison with NGSIM and Carla-SUMO co-simulation dataset. Section 4 defines the features

that might contribute to pavement deterioration and present the gradient boosting-based ensemble learning model. The developed model is validated in Section 5, and the results are discussed. Section 6 concludes this paper.

Chapter 2. Literature Review

2.1 Impact of CAV driving patterns on pavement performance

Although many driving patterns of CAV may impact pavement deterioration, lateral vehicle wandering is the only one that has been widely researched, which is shown to have a significant impact on pavement rutting, fatigue, and hydroplaning potential. For example, Chen et al. [22] evaluated the potential consequences of large-scale CAV implementation on the long-term service performance of road infrastructure. A finite element modeling approach was used to investigate pavement performance based on pavement rutting. In another research, four lateral distribution control models for autonomous trucks, namely, zero wanders, double peak Gaussian, uniform and two-section uniform, were proposed to generate the distribution curves [21]. Then, finite element analysis was utilized to investigate the effect of different wandering lateral distributions on pavement behavior. The rutting under four control modes with respect to the mixed traffic of autonomous trucks and non-autonomous trucks was compared. Similar research by Noorvand et al. [20] investigated the lateral wandering of autonomous trucks and simulated their impact on interstates pavement infrastructure with AASHTOWare Pavement Mechanistic-Empirical Design software. They were the first to compare the repetitions to the failure of a normal distribution wheel wander with the uniform and zero wander distribution. It was found that the uniformly distributed autonomous trucks mitigated pavement performance deterioration, including both rutting and fatigue cracking; thus, with a normal distribution of CAV wheel wandering, the designed pavement thickness can be reduced. To consider the pavement damage acceleration caused by autonomous and connect trucks (ACTs), Gungor and Al-Qadi [23] modified the damage accumulation equations in the Mechanistic-Empirical Pavement Design Guide (MEPDG) and take the lateral position of loading as an explicit input.

Aside from lateral wandering, vehicle speed and car-following spacing have also been investigated for their impact on pavement deterioration, although few have taken CAV's specific pattern into consideration. For example, according to AASHTO [25], the speed and volume of trucks (which determines the spacing between vehicles) can change pavement life. Many researchers studied the effects of rest periods on pavement fatigue and rutting under HDV traffic scenarios. For example, Underwood and Zeiada [26] tested the effects of a rest period on pavement fatigue with the different loading patterns and found that not only the mean value of the rest period impacts the

pavement fatigue, but also the distribution pattern of the rest period exerts the effect. Motevalizadeh et al. [27] studied the impact of loading and resting times, stress levels, and temperatures on hot-mix asphalt rutting accumulation. The results suggested the resting period had a significant impact on the rutting accumulation rate, especially at the initial stage. The smaller car following distance between ACT platoon, may also accelerate the pavement damage, as a consequence of which, Gungor et al. [24] developed a ACT fleet control strategy to minimize the pavement damage by optimize the lateral position of trucks and spacing between them. Mollenhauer et al. [28] measured the asphalt strain under different heavy loading frequencies and evaluated the impacts of traffic speed on pavement fatigue. However, there has not been sufficient research exploring the impact of vehicle speed and car-following spacing with realistic CAV driving patterns on pavement performance.

2.2 Dataset for CAV and HDV pattern comparison

Some data collection efforts that reveal driving behavior patterns for HDV and CAV can be found in the literature, with most of them focusing on lateral wandering patterns. For example, Blab and Litzka [29] found that lateral wander of human-driven vehicles follows a Laplace distribution while most research considered a normal distribution. The standard deviation of the normal distribution ranges from 8 to 24 inches. MEPDG recommends 10 inches as the default value for the standard deviation. In terms of CAV wandering, many studies assume that CAV has a zero wandering [30, 31] or even-distributed wander [32], which is not realistic and may overestimate its negative impact on pavement damage. Instead of adopting unrealistic assumptions, some studies collected the vehicle wandering dataset and investigated the impact of different wandering pattern patterns on pavement deterioration. For example, Sinanmis and Woods [33] investigated the relationship between wheel wandering of HDV and the geometric characteristics of roads. They analyzed the vehicle positions data collected by the cameras on 100 sections of pavement in the United Kingdom and concluded that the standard deviation of lateral wandering had a linear association with the lane width and road width. Another study by Zhou et al. [34] studied the lane following and centering technology with Texas A&M University's autonomous vehicle and analyzed its position data. The result indicated that while the lateral wandering patterns for both CAV and HDV could be modeled as a normal distribution, CAVs had at least three times smaller standard deviation of lateral wandering. Besides lateral wandering, the vehicle speed is also

considered a reason for CAV to impact pavement performance differently. Bimbraw [35] concluded that the travel time of CAV will decrease, and the ride experience will be smoother, as autonomous vehicles can travel at higher speeds with minimum chances of error, but the degree of its impact on the pavement was not quantified. To the best of our knowledge, other than those summarized above that focus on lateral wandering differences, the research to systematically investigate the driving pattern of CAV and HDV and quantitatively compare the difference has not been available.

Chapter 3. CAV and HDV Pattern Comparison

The objective of this section is to systematically compare the behavioral differences between CAV and HDV that may deteriorate pavement differently. To this end, the experimental design to achieve an apple-to-apple comparison is first explained, including the generation of the HDV dataset and CAV dataset. This section also defines the important trajectory-based features related to pavement performance deterioration and presents the data acquisition procedure.

3.1 Scenario setup

The Next Generation SIMulation (NGSIM) program is released by FHWA to collect detailed high-quality traffic datasets, covering two freeways and two urban arterials. As traffic on the urban street is subject to traffic lights and other factors such as pedestrians and bicyclists, a freeway dataset collected on eastbound I-80 in the San Francisco Bay area, in Emeryville, CA, on April 13, 2005, is used to characterize the representative HDV driving patterns. The study area is approximately 500 meters (1,640 feet) in length and multiple synchronized digital video cameras are mounted from the top of a 30-story building adjacent to the freeway. These cameras record vehicles passing through the study area, and image processing software is used to extract the vehicle trajectory data from the video recordings. Figure 2 illustrates the scenario setup, with the left figure providing a bird's view of the site, and the right figure showing the detailed highway geometries. The studied freeway segment has six lanes, each with a 12 ft width, and two ramps (one on-ramp and one off-ramp) are included in this area as well. The collected trajectory dataset has 25 characteristics, including vehicle ID, time stamp, location, vehicle class, velocity, acceleration, lane ID, proceeding vehicle ID, following vehicle ID, space headway, and time headway. These vehicle trajectory data provide the precise location of each vehicle within the study area every one-tenth of a second. The data used in this study were collected on April 13, 2005, from 4:03:56 pm to 4:08:56 pm. A total of 303,847 trajectory points from 778 vehicles are collected.

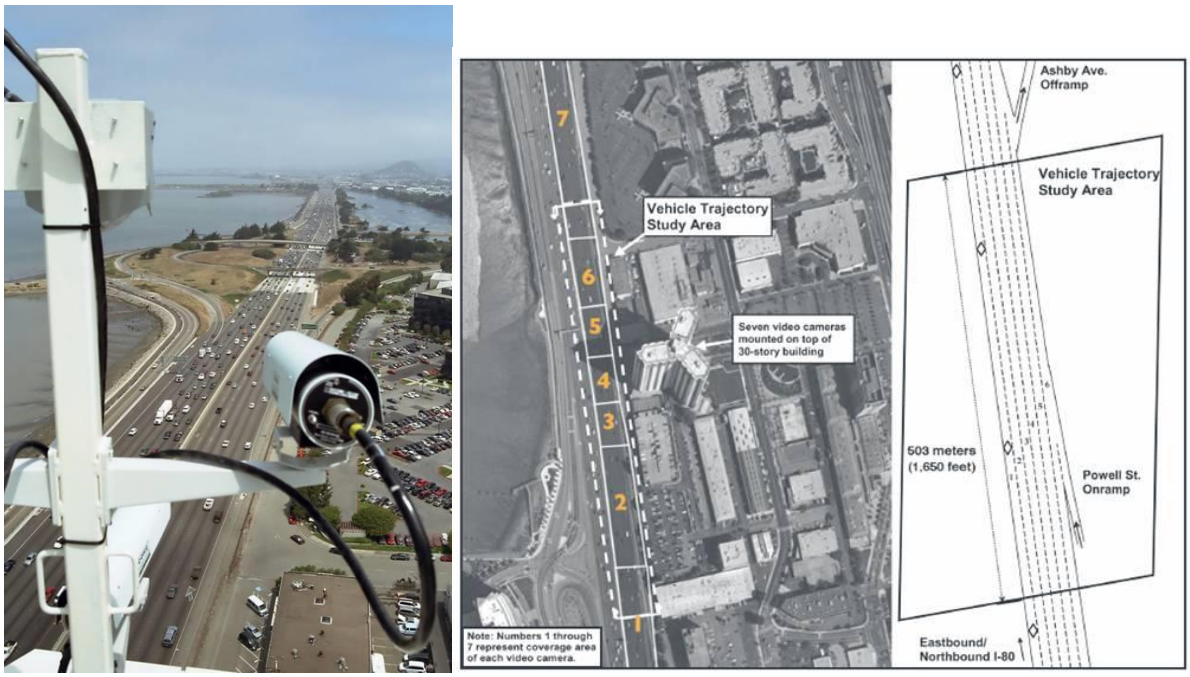


Figure 2. Illustration of NGSIM I-80 dataset collection in Emeryville, CA

To focus on the representative driving behaviors of the freeway vehicles, the collected trajectories are filtered with the following two criteria. First, the vehicles that are from the onramp or drive towards the offramp are removed so that the observed behavioral differences (such as speed change) can be mainly attributed to vehicle profile as opposed to the need for merging or diverging. For similar reasons, the vehicles that have changed lanes in this segment are also removed. After data cleaning, 601 vehicle trajectories remain and are utilized for driving pattern analysis. Samples of the vehicle trajectories are shown in Figure 3, with X-axis standing for the lateral distance from the left side of the roadway boundary, and Y-axis representing the longitudinal location. The blue lines divide the figure into 6 lanes, and each curve represents the trajectory of each vehicle trajectory.

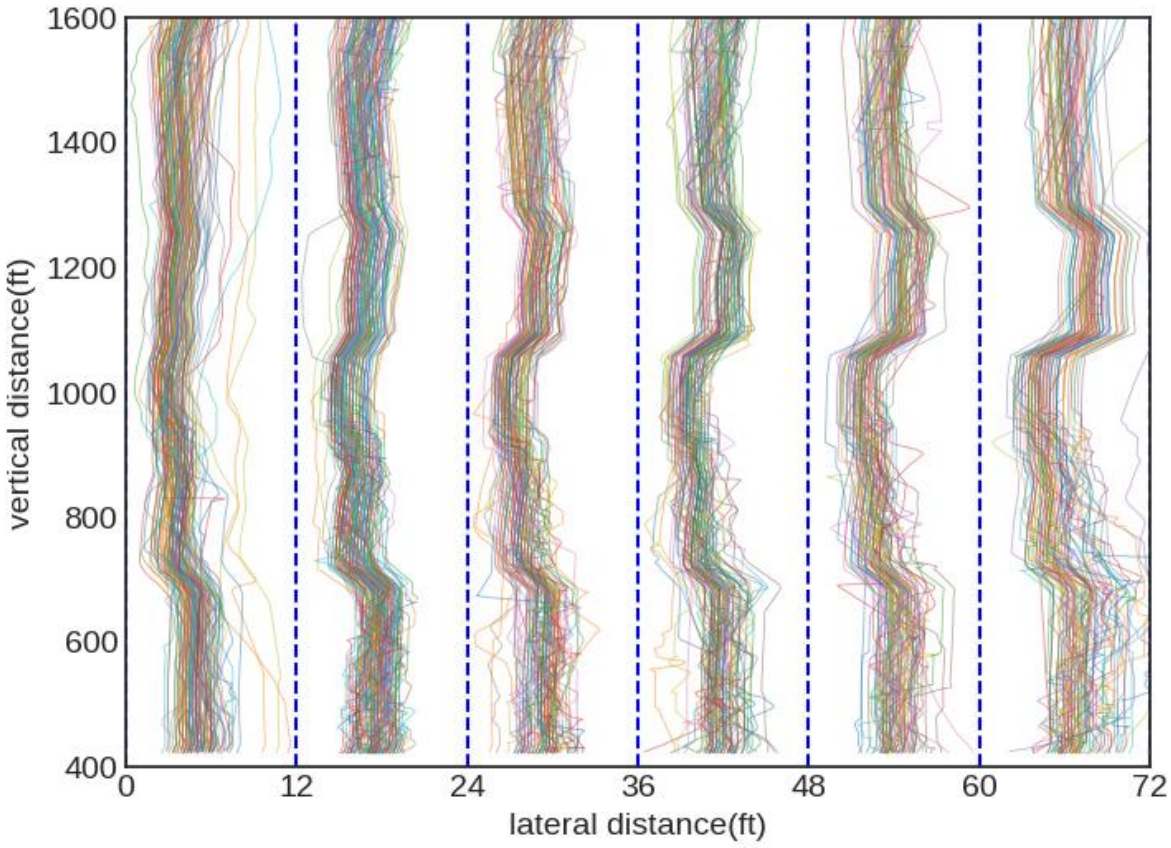


Figure 3. NGSIM vehicle trajectories after data cleaning

For the CAV trajectories generation, the researchers resort to a self-developed co-simulation framework that integrates CARLA, an open-source simulator for autonomous driving research, and SUMO, a microscopic multi-modal traffic simulation package. With the development of autonomous driving system, it is possible to collect data with the sensors from one single CAV, such as its traveling speed and acceleration. However, the capability of performing experiments with all vehicles being CAV on a roadway segment is currently impossible, not to mention that the researchers may have to test scenarios with mixed CAV and HDV with varying market penetration rates. As such, the co-simulation software becomes the best option that is currently available. As a popular CAV simulation tool, CARLA provides a 3D driving environment and supports flexible specifications of sensor suites, environmental conditions, full control of all static and dynamic actors, and map generation. In addition, SUMO, as a traffic simulation tool, has been commonly used for the simulation and analysis of road traffic and traffic management systems.

In this research, the simulation environment is created from the real-world data of the research area in the same segment of eastbound I-80 in the San Francisco Bay area or an apple-to-apple comparison purpose, as shown in Figure 4. The environment is created from the real-world OpenStreetMap data of the research area in I-80. The OpenStreetMap for the surrounding area is downloaded, and then the I-80 network dataset is extracted and imported into CARLA-SUMO co-simulation tool. CARLA 0.9.11 and SUMO 1.10.0 are used to set up the simulation environment. The number of lanes, lane width, traffic volume, and other simulation inputs are set to be exactly the same as those in the NGSIM dataset. The default autopilot mode in CARLA, as well as the lane-centering, adaptive cruise control, and other autonomous driving technologies, are enabled to simulate connected and autonomous driving. Simulation is then performed to generate the trajectories of all vehicles, and record and export them for research purposes. The generated dataset includes the vehicle ID, time, vehicle type, location (both longitudinal and lateral), speed, and acceleration information of each vehicle in the traffic flow. Then, the driving features, such as inter-vehicle gap distance and others, can be extracted from these continuous vehicle trajectories for later analysis.

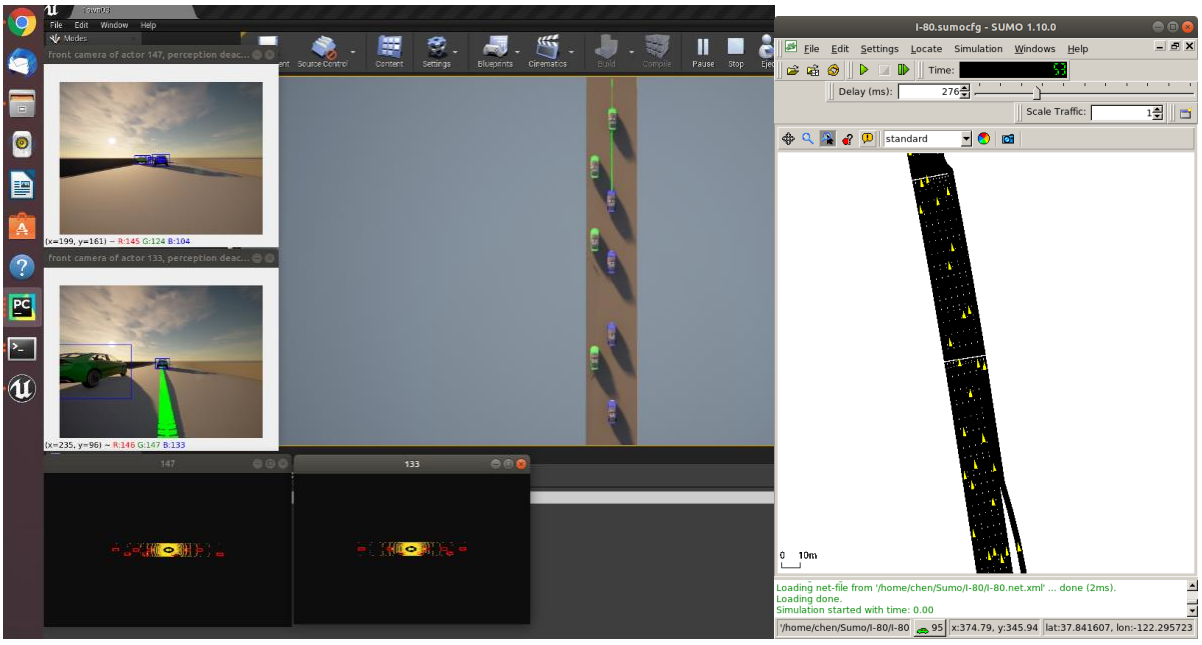


Figure 4. CARLA and SUMO co-simulation tool for Eastbound I-80

3.2 Pattern comparison

After extracting the raw driving data from NGSIM and CARLA-SUMO co-simulation tool, different traffic patterns of CAV and HDV are examined. The comparison of statistical characteristics is performed in three aspects, i.e., lateral wandering, longitudinal car-following distance, and vehicle speed.

3.2.1 Lateral wandering

The lateral position distributions of all HDVs (in orange color) and CAVs (in blue color) are presented in Figure 5. The X-axis represents the lateral position to the left edge of each lane, in feet, i.e., 0 means the most left side and 12 means the right side of the 12ft wide lane. The Y-axis represents the percentage of vehicles with different lateral positions. It can be observed that the lateral wandering distribution of CAV is concentrated between 3 to 4 feet. More than 70% of autonomous vehicles have a lateral wandering of less than 0.3 feet, meaning the CAV trajectories are mostly located along the center of the lane accurately. On the other hand, the lateral wandering distribution of HDV follows normal distribution approximately. It indicates that the wandering of HDV is more spread out compared with that of CAV. The standard deviations of CAV and HDV equal 0.46 ft and 1.21 ft, respectively, indicating a significant difference.

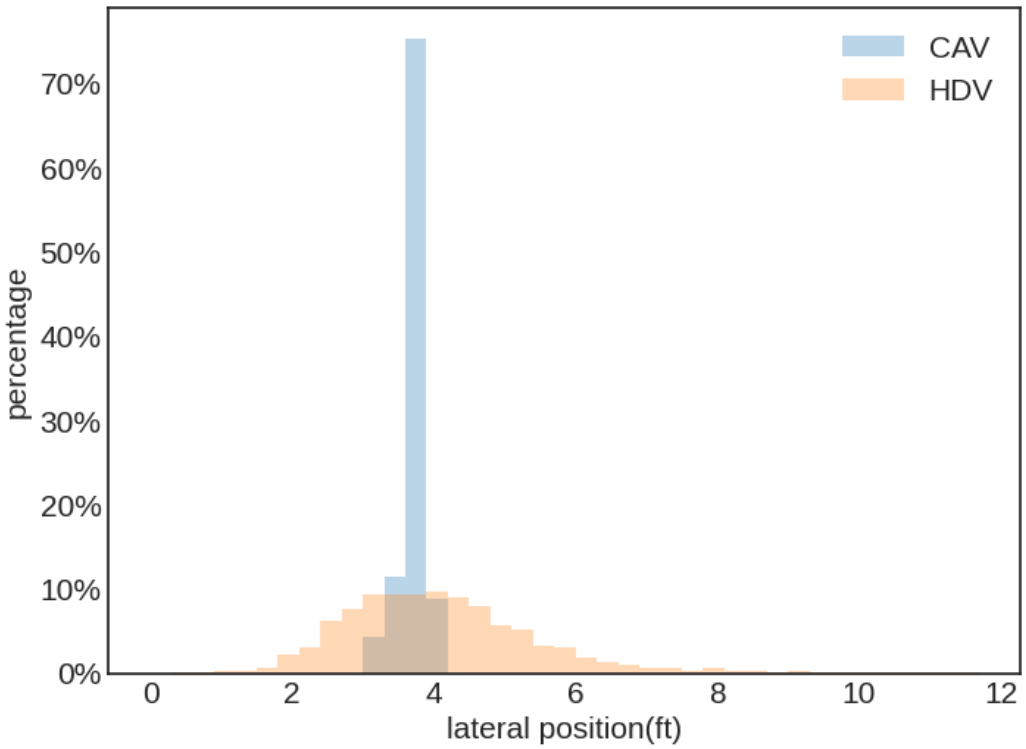


Figure 5. Lateral wondering comparison between CAV and HDV

3.2.2 Longitudinal car-following spacing

Longitudinal car-following distance, reflecting the rest period of pavement, influences pavement performance. Different car-following distances of HDV and CAV will result in different influences on the pavement. Without uncertainties brought by human driving behaviors, the car-following distance of CAV is designed narrower to improve traffic capacity. It will lead to a different distribution of car-following distance.

Figure 6 illustrates the longitudinal car-following spacing comparison between CAV and HDV. The X-axis represents the car-following spacing, in feet, and the Y-axis represents the percentage of different car-following spacing. It can be found that the spacing of CAV is concentrated between 25 and 50 feet; whereas the spacing of HDV trajectories is distributed from 0 to 150 feet. The mean spacings between vehicles under the CAV scenario and HDV scenario are equal to 31.5ft and 66.4ft, respectively. The standard deviations of spacing under CAV and HDV scenarios equal to 3.8 feet and 46.6 feet, respectively. These results show that with the help of cooperative driving automation technology, the CAV has a much smaller and even-distributed car-following distance.

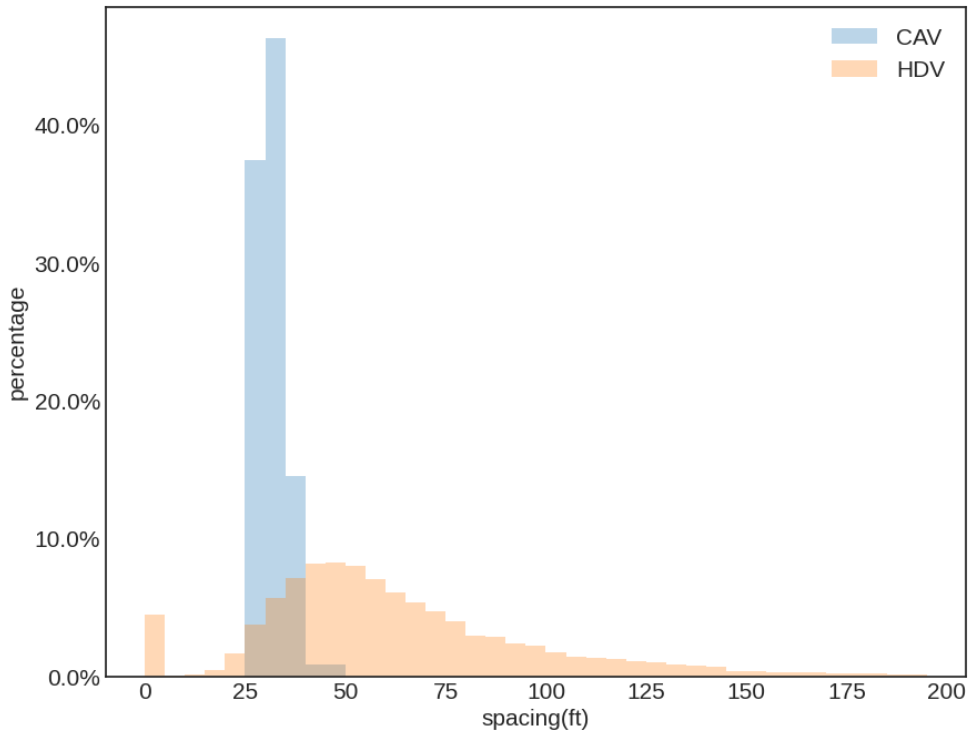


Figure 6. Car-following distance comparison between CAV and HDV

3.2.3 Vehicle speed

In terms of vehicle speed, there also exist significant differences between HDV and CAV. Figure 7 illustrates the vehicle speed comparison between CAV and HDV. The X-axis represents the speed in mph, and the Y-axis represents the percentage of different speeds. It can be found that the CAV speed distribution curve skews to the left when compared with that of HDV. The traffic flow in the studied freeway segment was extremely congested and the average speeds of CAV and HDV are equal to 14.4 mph and 18.4 mph, respectively, meaning, CAVs are in general, driving slower than HDVs. CAV distribution is more concentrated and narrower than HDV because the vehicles can communicate and share information with each other. In other words, the CAV operating conditions are more homogenous, whereas the HDVs may exhibit very different driving behaviors.

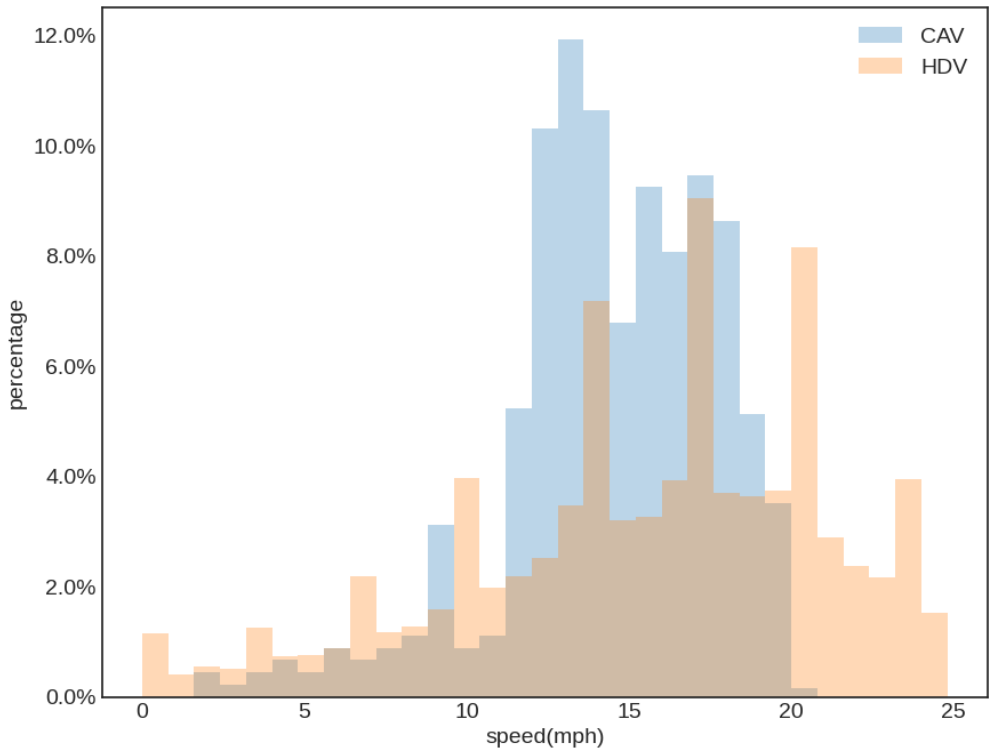


Figure 7. Speed comparison between CAV and HDV

Chapter 4. Pavement Deterioration Prediction Model Development

In this study, a boosting-based ensemble learning model to predict the pavement IRI based on the LTPP dataset is adopted from the team's previous research [36]. The response variable for the model is the IRI measurement at a specific age after the traffic open date. While the mathematical formulation is similar to the previous study, three new driving pattern-related features are added to the machine learning model, and supplement the context-based features that describe the environment that roadway segment is in, and the model is re-trained with the updated dataset with the new features. For convenient reference, the mathematical notations used in this paper are summarized below.

i: Index of observation from LTPP dataset

$AADT$: Annual average daily traffic

K: The proportion of $AADT$ occurring in the peak hour

D: The proportion of peak-hour volume in the major direction

PHV: Peak hour traffic volume

$LTWD$: Lateral wandering

LW: Lane width

LS: Longitudinal spacing

PHV_t: Peak hour traffic volume of the lane where the test sections are located.

v: Vehicle speed

v_f: Free flow speed

v_m: Speed limit

CAP: Actual capacity of a road (pcu/h)

α: A coefficient of the Bureau of Public Roads (BPR) function, with recommended values of 0.15

β: Another coefficient of the BPR function, with recommended values of 4

4.1 Driving feature definition

4.1.1 Driving pattern-related features

To evaluate the impact of CAV on pavement performance, features to calibrate different traffic patterns need to be defined. As mentioned before, three behavioral factors are highlighted, i.e., the standard deviation of vehicle wheel lateral wandering, longitudinal car-following space, and traffic speed, to illustrate the different traffic patterns between CAV and HDV. As no direct data can be used to assess the driving pattern impact on pavement deterioration, the researchers utilized the LTPP dataset to extract these three driving features. It should be noted that the purpose of doing this is very different from those described in the section above. While Section 3 generates the different traffic patterns between CAV and HDV, the dataset is not suitable for machine learning model training because the corresponding pavement performance is not available. As such, the researchers resort to the LTPP dataset, which provides not only the context information (such as traffic volume and weather) but also the pavement deterioration conditions, for machine learning model training. The definitions of these three features and data extraction process from the LTPP database are presented below.

Lateral wandering ($LTWD$) is the standard deviation of the transverse wander of the vehicle tire location across the lane width. Since the lateral wander values are not explicitly available in the LTPP dataset, a model proposed by Sinanmis and Woods [33] to estimate the standard deviation of vehicle position is adopted. The mathematical calculation is summarized by Eqs. (1) and (2) with an assumption that a human-driven vehicle's lateral wandering pattern is in a linear relationship with the lane width and road width.

$$LTWD = -0.006 + 0.033LW + 0.017RW \quad (1)$$

$$RW = LW * n + ISW + OSW \quad (2)$$

where LW is the lane width and RW is the road width. The road width can be derived by Eq. (2), n is lane number, ISW and OSW denote inside and outside shoulder widths, respectively.

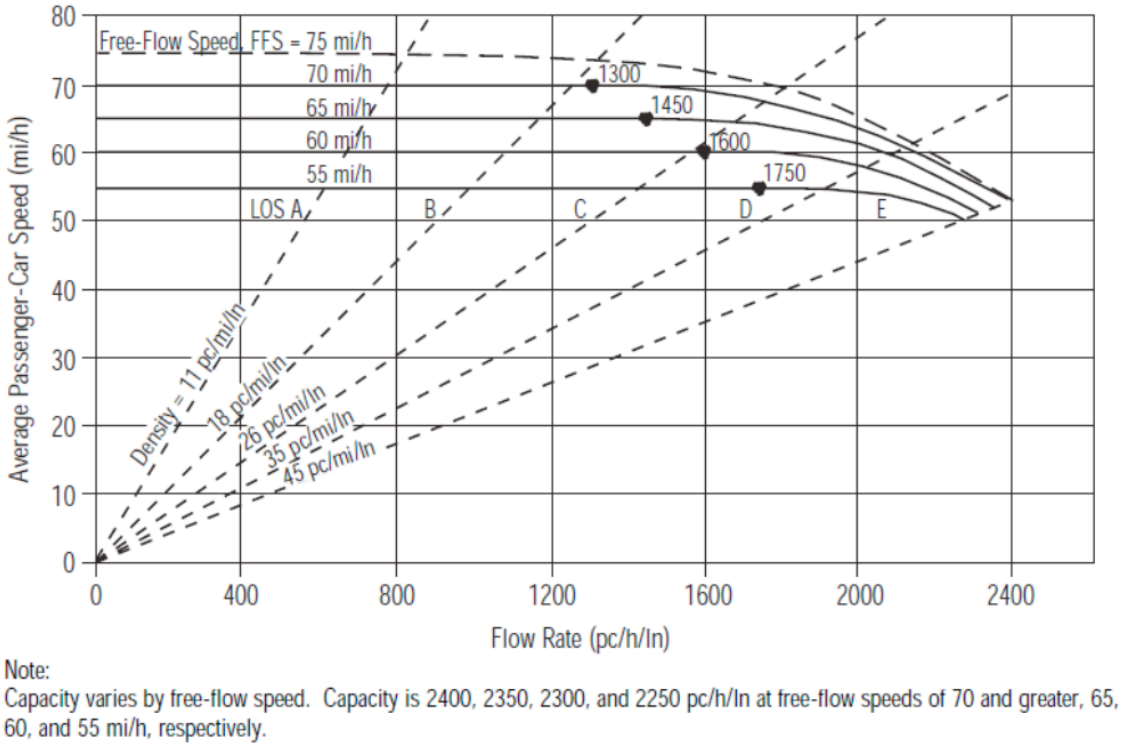
Traffic speed (v) is the second driving-related feature, which is estimated using the Bureau of Public Roads (BPR) function [37]. The mathematical calculation is summarized in Eqs. (3) and(4), where parameters α and β are two parameters of the BPR function, PHV is the peak hour volume, and CAP is the capacity. PHV is calculated by Eq. (4), in which AADT stands for the average

annual daily traffic volume, K-factor is the proportion of AADT occurring in the peak hour, and D-factor is the peak-hour volume proportion in the major direction.

$$v = \frac{v_f}{1 + \alpha \left(\frac{PHV}{CAP}\right)^{\beta}} \quad (3)$$

$$PHV = AADT * K * D \quad (4)$$

The two parameters, α and β, are set to be 0.15 and 4 which are standard values in practice. The roadway capacity CAP is estimated according to the Highway Capacity Manual (HCM), based on Exhibit 23-2 which is shown in Figure 8 below, with an explicit relationship between traffic flow and speed.



Note:
Capacity varies by free-flow speed. Capacity is 2400, 2350, 2300, and 2250 pc/h/ln at free-flow speeds of 70 and greater, 65, 60, and 55 mi/h, respectively.

Figure 8. Speed-flow curves and LOS for basic freeway segments (HCM Exhibit 23-2)

The third driving-related feature is longitudinal car-following spacing (LTS), which represents the average spacing between two consecutive vehicles. Learning from classic traffic flow theory [38], spacing is equal to the inverse of traffic flow density k, i.e., $LTS = 1/k$, whereas as the

fundamental diagram suggests, the density k is equal to the flow divided by traffic speed. In other words, $k = PHTV/v$. With some simple reorganization, it can be found that $LTS = \frac{v}{PHTV}$.

4.1.2 Context-based features

As the pavement deterioration is affected by many context-based factors other than traffic condition (e.g., climate, pavement material, pavement structure and pavement initial performance), the pavement performance can't be predicted accurately with only the traffic information mentioned in 0. Therefore, in addition to the three driving pattern features that are presented above, 20 more context-related features are also defined to describe the context conditions that may contribute to the pavement deterioration. Consistent with the previous research [36], these context-based features are divided into six categories, including pavement performance (multiple types of distresses, rutting, and friction), pavement structure, drainage, climate conditions, traffic history and material properties. The notation, definition, and source of the response variable are provided in Table 1 below.

Table 1 Definition of context-related features

Category	Notation	Definition and Unit	LTPP Module
Pavement performance	IRI_0	The IRI value measured when age was 0. (m/km)	MON
	Cr_{Gator}	Area of alligator cracking in square meters. (m^2)	
	Cr_{Lwp}	Length of longitudinal cracks within the defined wheel paths in meters. (m)	
	Cr_{Lnwp}	Length of longitudinal cracks not in the defined wheel paths in meters. (m)	
	Pt_A	Area of patches in square meters. (m^2)	
	Pt_N	Number of patches in square meters. (m^2)	
	Cr_{Wp}	Length of wheelpath cracks in meters. (m)	
	Cr_{Gt183}	Total length of transverse cracks greater than 1.83. (m)	

	Rt	The depth of rutting in millimeters. (mm)	
	Fr	Friction number between the vehicle wheel tire and the pavement	
Pavement structure	Tk_{Sb}	Layer thickness measurement for surface coarse and binder course. (in)	SPS
	Md_s	Average backcalculated elastic modulus of the surface layer.(psi)	FWD
Drainage	$Hydr$	Average measured hydraulic conductivity of the specimen. (cm/sec)	DRAIN
Climate	$Prcp$	Average monthly precipitation in millimeters. (mm)	AWS
	Fz	Average freeze index. (°C/day)	
Traffic	$Esal$	Annual average ESAL ($kESAL$)	TRF
	$Esal_q$	quadratic form of Kesal ($kESAL^2$)	
	Age	Time duration between new construction to roughness survey date. ($year$)	
Material properties	Gr	Mean specific gravity of asphalt cement	TST
	Pt_{Ca}	Coarse aggregate amount percent by total weight of aggregate in percentage. (%)	

4.2 Boosting-based ensemble machine learning model development

Machine learning models have been increasingly used to forecast pavement conditions. Currently, many pavement performance prediction models have been developed using linear regression or simple machine learning techniques. However, as pointed out in the team's previous study [36], challenges remain in the model assumptions, training efficiency, interpretability, and prediction

accuracy. In this study, the research team builds on the previous modeling approach that utilizes an boosting-based ensemble learning model, couple with the newly defined machine learning features, to predict the International Roughness Index (IRI) of asphalt pavements.

The adopted model utilizes a boosting-based learning algorithm that reduces bias by training new weak learners based on the outputs of previous learners [39]. It is developed using the gradient boosting decision tree (GBDT) algorithm [40]. However, GBDT has a drawback in that its efficiency decreases when working with large datasets with multiple dimensions as it uses the first derivative of the loss function during base model training. The ThunderGBM modeling approach makes improvement upon this by incorporating second-order gradient statistics into its optimization process, leading to improved prediction accuracy and efficiency. Additionally, as a more recent implementation of the gradient boosting algorithm, ThunderGBM is designed to be graphics processing unit (GPU)-oriented and includes optimization techniques such as high dimensional data histogram building and feature-based data partitioning to maximize GPU usage. This makes ThunderGBM a faster option compared to other boosting algorithms such as XGBoost [41], LightGBM [42], and CatBoost [43], while still producing comparable models.

The training dataset is designed as $O = [(x_1, IRI_1), (x_2, IRI_2), \dots, (x_n, IRI_n)\ ]$, where x_i is an array representing 23 defined features, i.e., three driving pattern-related features plus the 20 context-based features. IRI_i is the pavement IRI value for the i^{th} observation out of total N_e training samples. The target of learning process is to find the function that minimizes the objective function $L(\Theta)$, which is defined as the summation of the loss function l and the regularization term Ω as presented in Eq. (5).

$$Min\ L(\Theta) = \sum_{i=1}^{N_e} l(IRI_i, \widehat{IRI}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{5}$$

$$\Theta = [f_1, f_2, \dots f_K\} \tag{6}$$

$$l\left(IRI_i, \widehat{IRI}_i\right) = \left(IRI_i - \widehat{IRI}_i\right)^2 \tag{7}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{8}$$

$$\widehat{IRI}_i = \sum_{k=1}^{K} f_k(x_i) \tag{9}$$

$$x_i = \begin{bmatrix} Performance_i, Structure_i, Traffic_i, Climate_i, \\ LTWD_i, v_i, LS_i \end{bmatrix} \tag{10}$$

where l is the loss function and Ω is a regularization term. $\widehat{IRI}_i$ is the combination of a total number of K base regression trees with each being $f_k(x_i)$, as shown in Eq. (6) and Eq. (9). The loss function is defined in Eq. (7), in which IRI_i is the ground truth value, and $\widehat{IRI}_i$ is the predicted value from the developed machine learning model. The regularization term is defined in Eq. (8), and Eq. (10) illustrates the 23 features to be used in this machine learning model.

To train the machine learning model, the dataset is randomly split into two parts: 80% for training and 20% for testing. The training subset is used to determine the model parameters, and the testing subset is reserved for evaluating the model's performance. A performance metric, Root Mean Square Error (RMSE), is defined by Eq. (11) below and used to validate the model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{11}$$

Chapter 5. Numerical Analysis

In this section, experiments are conducted to quantitively analyze the different impacts of CAV and HDV traffic patterns on pavement performance. The LTPP dataset is used to train the proposed ensemble machine learning model and evaluate the prediction accuracy. Then, the trained machine learning model is applied to predict the pavement IRI, under HDV and CAV scenarios, respectively. The predicted results are analyzed to understand the different impacts of HDV and CAV traffic patterns. Sensitivity analysis is then conducted to understand how the pavement IRI changes as other factors vary, including pavement age, traffic volume, and CAV market penetration rate.

5.1 Validation of the boosting-based prediction model

The LTPP data used in this research is the same as the team's former research [36]. Specifically, 246 road sections with 1,707 observations that are released in July 2021 are retrieved from Structural Factors for Flexible Pavements (SPS-1) experiment. The extracted sections cover all four climate regions (i.e., dry freeze, dry no-freeze, wet freeze, and wet no-freeze) across the U.S. As mentioned above, 80% of the dataset is used for the training of the proposed machine learning model, and the rest 20% is used for testing purposes. The results are summarized below for the training and testing dataset, respectively.

The residual plot and the residual distribution of the model on training and testing data are presented in Figure 9. The X-axis stands for the predicted IRI, whereas Y-axis is the residuals. The blue dots represent the training sample, whereas the green dots are for the testing dataset. It can be clearly observed from Figure 9 that the training residuals are all very close to 0, whereas the testing residuals deviate more significantly. Overall, R square values of 1.0, and 0.887 are obtained for the training and testing dataset, respectively, both indicating satisfactory prediction results. In addition, it can be seen from the residual distribution of training data that the residuals are generally evenly distributed on the negative and positive sides. In other words, the proposed model achieves a good balance between underestimation and overestimation in both the training (80%) and testing (20%) dataset.

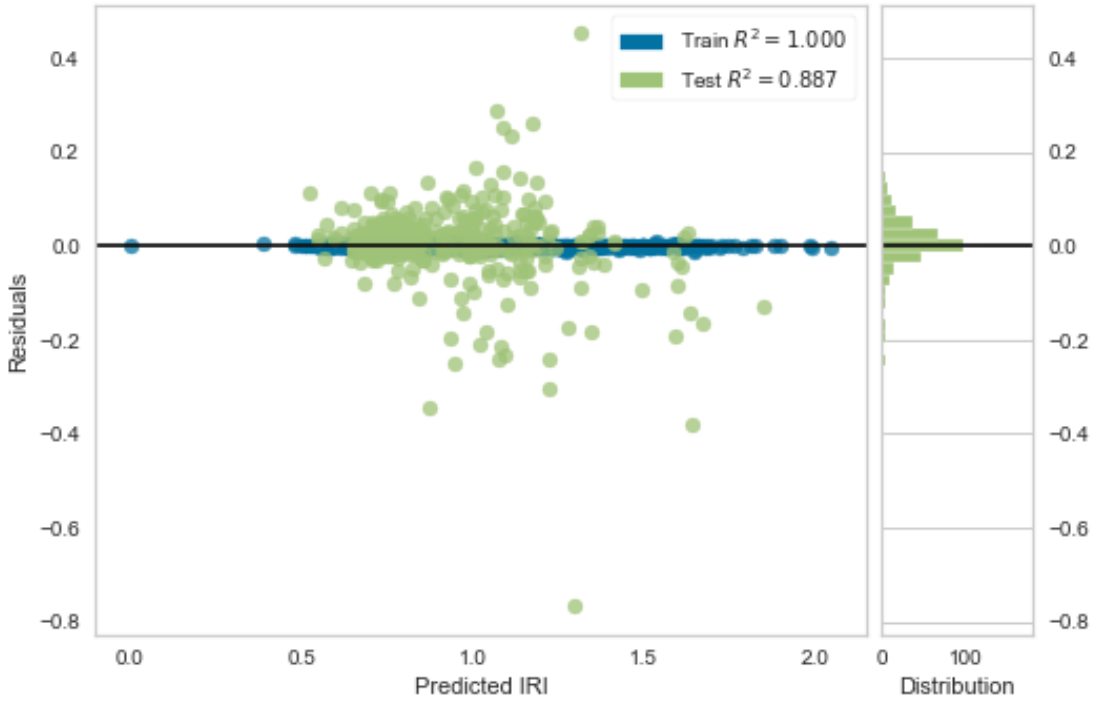


Figure 9 Residual plot of the proposed ensemble learning model

The developed model is also compared with five benchmark models, i.e., the multivariate linear regression analysis model, MEPDG model, artificial neural network (ANN) model, random forest model (RF) model, and the ThunderGBM model with only 20 features but not the three newly developed driving pattern-related features. The R square values for these five benchmark models are shown to be 0.577, 0.560, 0.792, 0.862, and 0.881, respectively. Compared with the proposed model with an R square of 0.887, the proposed ensemble learning model with the 23 features is shown to have the highest prediction accuracy.

5.2 Pavement deterioration comparison between CAV and HDV

In this section, the trained machine learning model is applied to understand the different impacts of HDV and CAV traffic patterns on pavement deterioration. The input dataset includes two parts: (1) twenty features related to pavement performance, pavement structure, traffic, climate, and other properties; and (2) three driving pattern-related features including lateral wandering, longitudinal car-following distance, and vehicle speed. To perform the experiment under a realistic

setting, the values of the twenty features are extracted from the LTPP dataset. For the remaining three driving pattern-related features, the NGSIM dataset and CARLA-SUMO co-simulation results are utilized to indicate the HDV and CAV driving patterns, respectively. It should be noted that although the context features are set to be the same for the CAV and HDV, it is very important to include these features as they provide the context information that the pavement segment is in, and the same driving pattern may create different pavement damage under different settings (for example, the same inter-vehicle spacing may deteriorate pavement differently between a dry warm condition versus a wet cold condition). These two sets of trajectory-related features are denoted as x_{HDV} and x_{CAV}, respectively. To summarize, x_{HDV} and x_{CAV} are defined in the following manner, with the first portion of $[Performance_i, Structure_i, Traffic_i, Climate_i]$ from the LTPP dataset and the second portion $[LTWD_i, v_i, LS_i]$ from NGSIM or CARLA-SUMO simulation.

$$x_{HDV} = \begin{bmatrix} Performance_i, Structure_i, Traffic_i, Climate_i, \\ LTWD_{hdv}, v_{hdv}, LS_{hdv} \end{bmatrix} \quad (12)$$

$$x_{CAV} = \begin{bmatrix} Performance_i, Structure_i, Traffic_i, Climate_i, \\ LTWD_{cav}, v_{cav}, LS_{cav} \end{bmatrix} \quad (13)$$

The results are summarized in Table 2 below. The first row shows the results of HDV, and the second row is for CAV. As there are a total of 1,707 data records, the average values are computed to compare the overall differences. The initial pavement IRI is the same for both CAV and HDV, at 0.839 m/km. Under the HDV scenario, the final IRI becomes 0.9719 m/km, increasing by 0.1329, or equivalently, 15.84%, compared to the initial IRI. However, under the CAV scenario, the final IRI becomes 0.9826 m/km, or in other words, increases by 0.1436 on average, or 17.11%. Comparing these numbers, it can be found that on average, CAV leads to an IRI increase that is 0.0107 higher than that of HDV, or equivalently, 8.1%. Since the first twenty context-based features are the same for each data point, the differences can be contributed to the different driving patterns between CAV and HDV. In other words, everything else being same, the deployment of CAV will deteriorate pavement at a speed that is 8.1% faster than HDVs, calling for more frequent maintenance in the future.

Table 2 Characteristics of Pavement IRI under HDV and CAV scenarios

	Initial mean	**Final mean**	**Increased mean**	**CAV**

	IRI	IRI	IRI	Impact
HDV	0.8390	0.9719	0.1329 (15.84%)	8.1%
CAV	0.8390	0.9826	0.1436 (17.11%)	

5.3 Sensitivity analysis

Following the overall pavement deterioration comparison between CAV and HDV, in this section, the researchers investigate how different ages of pavement, varied traffic volume, and CAV market penetration rate will impact pavement performance.

5.3.1 Pavement IRI with different ages

In this first subsection, the impact of different ages of pavement on IRI under CAV and HDV traffic patterns are analyzed, respectively. The age of pavements ranges from 0 to 20 years, with a mean value of 8.60. 75% of pavements have an age of over 11.27 years. pavements are divided into five groups as shown in Table 3 and perform analysis for each group.

Table 3 Statistical characteristics of increased mean IRI with different pavement ages

Age (Year)	Number of data records	Initial mean IRI	Final mean IRI for CAV	Final mean IRI for HDV	IRI increase for CAV	IRI increase for HDV	CAV impact
[0,4]	225	0.8895	0.9404	0.9291	0.0509	0.0396	28.5%
[4,8]	624	0.8452	0.9994	0.9895	0.1542	0.1443	6.86%
[8,12]	504	0.8164	0.9871	0.9729	0.1707	0.1564	9.14%
[12,16]	253	0.8135	0.9616	0.9557	0.1481	0.1422	4.15%
[16,20]	101	0.8639	1.0025	0.9943	0.1386	0.1303	6.37%

Table 3 also summarizes the statistical characteristics of the increased mean IRI with different pavement ages, including the number of data records, initial mean IRIs, final mean IRIs for CAV

and HDV, IRI increases for CAV and HDV, and the impact of CAV special driving patterns. It can be observed that for both CAV and HDV scenarios, the pavement IRI increases significantly. The increased IRI under the HDV traffic pattern ranges from 0.0396 m/km to 0.1564 m/km; whereas it ranges from 0.0509 m/km to 0.1707 m/km for the CAV scenario. The results also show that the rate of pavement deterioration increases sharply as the pavement age increase until 12 years, and after that, the deterioration rate decreases slowly.

To compare the impact of CAV and HDV on pavement more clearly, Figure 10 is generated to visualize the IRI increases for CAV and HDV, respectively. The CAV impact, defined as the IRI increase for CAV minus the IRI increase for CAV HDV, is also shown in the same figure. It can be found that in general, the IRI under the CAV scenario increases more than that under the HDV scenario, and this is the same for each group, which is consistent with the findings from the last section. Further, the CAV impact differently on pavements with varying ages, but in general, it creates higher impacts on a new pavement than an aging pavement. The CAV impact on a new pavement (i.e., 0~4 years) is found to be most significant, at 28.5% (i.e., CAVs deteriorate pavement at a speed that is 28.5% faster than HDVs do). The impact becomes 6.86% for pavements with age of 4-8 years, 9.14% for pavements with age of 8-12 years, 4.15% for pavements with age of 12-16 years, and 6.37% for pavement with age greater than 16 years. While fluctuations are observed, in general, as suggested by the blue curve in Figure 10, the CAV impact is less significant when pavement age increases.

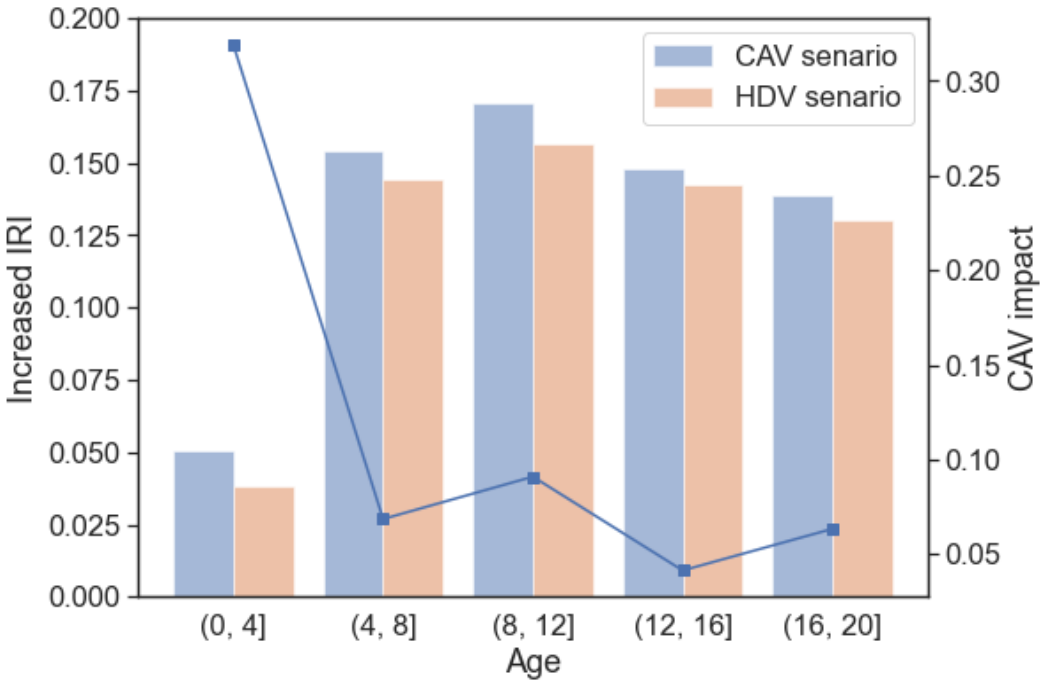


Figure 10 Impact of different pavement age on IRI

5.3.2 Pavement IRI with Varied Volumes

In this subsection, how varying traffic volume impact pavement performance is investigated. Different peak-hour truck volumes (PHTV) ranging from 0 to 150 vehicles per hour (vph) are adopted to reflect traffic volume. Figure 11 presents the influence of traffic volumes on pavement performance under the CAV scenario. The X-axis represents the PHTV in five groups, and Y-axis shows the increase in IRI. It can be observed that the increased IRI is relatively small, at around 0.13 m/km, when the PHTV is less than 90 vph. After that, the increased IRI increases obviously from 0.1341 m/km to 0.1974 m/km as the truck volume increases from 90 vph to 150 vph. The results suggest that the impact of CAV traffic pattern on pavement IRI is firstly stable at a relatively lower level when PHTV is lower than 90 vph. After that, the pavement IRI increases significantly as the truck volume increases, or in other words, the pavement degradation speeds up when the truck volume is greater than 90 vph.

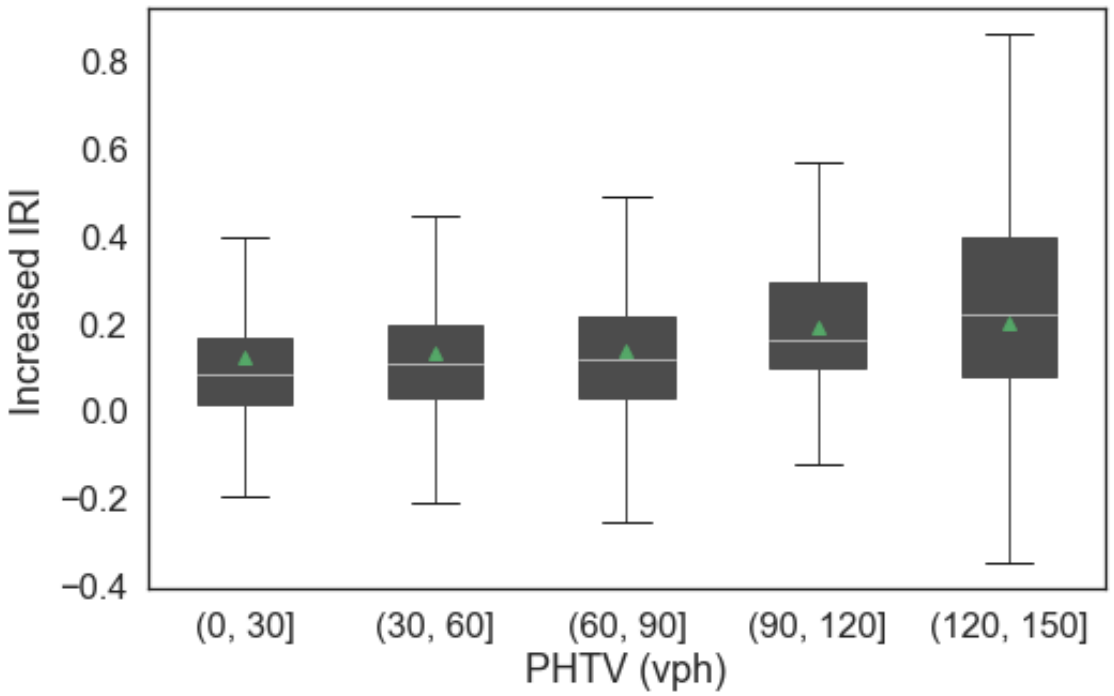


Figure 11 Relationship between PHTV and increased IRI

5.3.3 Pavement IRI Different Penetration Rate

Apart from the pavement age and traffic volume, the impact of different CAV market penetration rates on pavement performance is also analyzed. The driving features for the mixed traffic with both CAV and HDV are calculated by the summation of two normal distributions. Specifically, if the CAV penetration rate is p, the lateral wandering standard deviation $LTWD_{mix}$ is calculated by $p * LTWD_{CAV} + (1-p) * LTWD_{HDV}$, the average speed of mix traffic $SPEED_{mix}$ is calculated by $p * SPEED_{CAV} + (1-p) * SPEED_{HDV}$, and the average car-following spacing between vehicles of mix traffic $SPACING_{mix}$ is calculated by the weighted harmonic mean: $1/(p/SPACING_{CAV} + (1-p)/SPACING_{HDV})$.

The relationship between increased mean IRI and CAV penetration rate is shown in Figure 12. The X-axis is the CAV penetration rate, increasing from 20% to 100%, and Y-axis is the increased mean IRI. It can be observed that in general, the increase in IRI increases when the CAV market penetration goes up. The mean IRI increases from 0.08 m/km to 0.14 m/km as the CAV penetration rate increases from 20% to 100%. In particular, the mean IRI increases significantly from 0.08

m/km to 0.13 m/km when the penetration rate increases to 0.4. After that, the increasing trend becomes less significant and the increased mean IRI remains stable at around 0.14. The data suggests that the pavement performance will be affected more significantly at the beginning stage of CAV deployment when the CAV market penetration rate increases. After the CAV accounts for half of the vehicles in the traffic flow, the impact of the CAV traffic pattern will become stable.

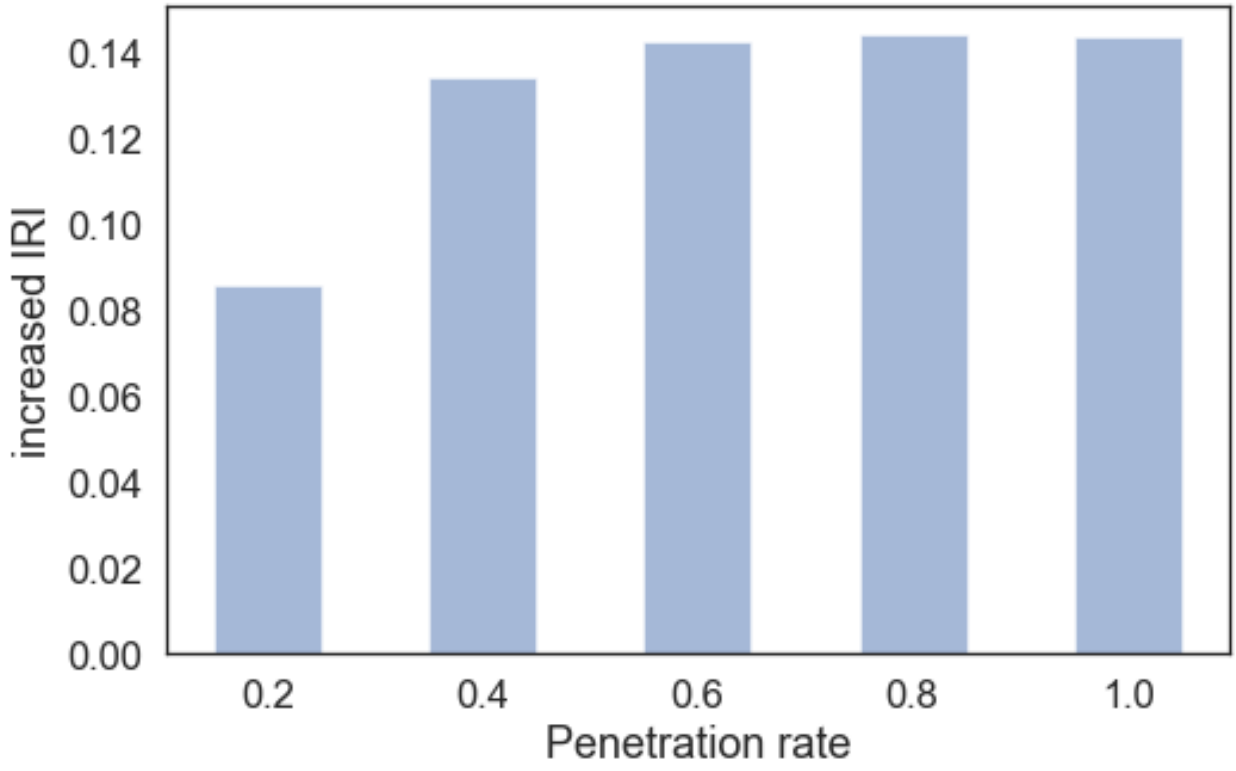


Figure 12 Relationship between increased mean IRI and CAV penetration rate

Chapter 6. Conclusion

In this study, the researchers systematically and quantitatively model the impact of CAV traffic patterns on pavement performance and compare it with HDV. Three trajectory-related features, i.e., lateral wandering, longitudinal car-following distance, and driving speed, are selected to calibrate the traffic pattern of CAV and HDV. To compare their different traffic patterns, those trajectory-related features are extracted from the NGSIM dataset for the HDV traffic pattern, and from the CARLA-SUMO co-simulation platform for the CAV traffic pattern. Through statistical comparison, it can be clearly observed that the traffic pattern of CAV is different from that of HDV.

A ThunderGBM-based ensemble learning model is developed to learn the relationship between pavement IRI and input features. The input dataset includes twenty features, related to performance, structure, climate, and traffic collected from the LTPP database, as well as three trajectory-related features. The trained model is then applied to predict the pavement IRI using different driving pattern features of HDV and CAV. The modeling results show that the final mean pavement IRI under the HDV driving pattern is 0.9719 m/km, while the mean IRI under the CAV driving pattern reaches 0.9826 m/km. In other words, that special CAV traffic pattern will deteriorate pavement faster than HDV by an average of 8.1%.

Sensitivity analysis is also conducted to investigate how the pavement performance changes if the pavement age, traffic volume, and CAV penetration vary. The results show that CAV deployment will create a greater impact on the younger pavements than on the aged pavements. The rate of pavement deterioration is found to be stable under light traffic, whereas it will increase under congested traffic. At the beginning stage of CAV deployment with a lower penetration rate, the pavement performance will be affected more significantly. After the CAV accounts for half of the vehicles in the traffic flow, the impact of the CAV traffic pattern will become stable. These results from this study reveal how system-wide CAV deployment may impact pavement performance and may be further used to support the decision-makers to identify optimum and cost-effective strategies for maintaining pavement infrastructure in serviceable condition.

While our study has focused on assessing the impact of autonomous vehicles based on different lateral wondering, longitudinal spacing and vehicle speed profiles, there are several avenues for future research that can build upon our findings and expand the understanding of this topic. Firstly,

beside these three driving pattern features, CAVs may have different flow distribution patterns among different traffic lanes when compared with HDVs, which will also significantly impact pavement deterioration. Additionally, to better assess the impact of CAV on pavement deterioration, performance of CAV, such as the accuracy of perception sensors and vehicle control modules, needs to be investigated and updated with the evolvement of CAV technologies. Last but not the least, the resulted pavement deteriorations may in turn impact CAV's driving patterns, which may deserve more attentions. For example, the new pavement roughness profile will impact CAV's lateral wandering as well as their decisions on which route to travel on. A new feedback loop may be added to capture such interactions.

Chapter 7. Reference

1. Fan, W. and T. Yang, *Impact of Connected and Autonomous Vehicles on Signalized Intersections With Transit Signal Priority*. 2022, University of North Carolina at Charlotte. Center for Advanced Multimodal ….
2. Yang, T. and W. Fan, *Evaluation of transit signal priority at signalized intersections under connected vehicle environment.* Transportation Planning and Technology, 2023. **46**(2): p. 145-159.
3. Yang, T. and W. Fan, *Transit Signal Priority under Connected Vehicle Environment: Deep Reinforcement Learning Approach.* Journal of Intelligent Transportation Systems, 2024: p. 1-13.
4. Yang, T. and W. Fan, *Enhancing Robustness of Deep Reinforcement Learning Based Adaptive Traffic Signal Controllers in Mixed Traffic Environments Through Data Fusion and Multi-Discrete Actions.* IEEE Transactions on Intelligent Transportation Systems, 2024: p. 1-13.
5. Yang, T., *Transit Signal Priority Control with Connected Vehicle Technology: Deep Reinforcement Learning Approach*. 2024, The University of North Carolina at Charlotte.
6. Tang, Q., Y. Cheng, X. Hu, C. Chen, Y. Song, and R. Qin, *Evaluation Methodology of Leader-Follower Autonomous Vehicle System for Work Zone Maintenance.* Transportation Research Record: Journal of the Transportation Research Board, 2020.
7. Tang, Q., X. Hu, and R. Qin, *Development of Operation Guidelines for Leader-Follower Autonomous Maintenance Vehicles at Work Zone Locations.* Transportation Research Record Journal of the Transportation Research Board, 2021.
8. Tang, Q., X. Hu, and H. Yang, *Identification of Operational Design Domain for Autonomous Truck Mounted Attenuator System on Multilane Highways.* Transportation Research Record, 2021: p. 03611981211061555.
9. Tang, Q. and X. Hu, *A multi-state merging based analytical model for an operation design domain of autonomous vehicles in work zones on two-lane highways.* Journal of Intelligent Transportation Systems, 2022: p. 1-14.
10. Chen, C., Y. Song, X. Hu, and I.G. Guardiola, *Analysis of Electric Vehicle Charging Behavior Patterns with Function Principal Component Analysis Approach.* Journal of Advanced Transportation, 2020. **2020**: p. 8850654.
11. Song, Y. and X. Hu, *Learning-based demand-supply-coupled charging station location problem for electric vehicle demand management.* Transportation Research Part D: Transport and Environment, 2023. **125**: p. 103975.
12. Song, Y. and X. Hu, *Learning electric vehicle driver range anxiety with an initial state of charge-oriented gradient boosting approach.* Journal of Intelligent Transportation Systems, 2023. **27**(2): p. 238-256.
13. Cheng, Y., X. Hu, Q. Tang, H. Qi, and H. Yang, *Monte Carlo Tree Search-Based Mixed Traffic Flow Control Algorithm for Arterial Intersections.* Transportation Research Record, 2020. **2674**(8): p. 167-178.

14. Cheng, Y., C. Chen, X. Hu, K. Chen, Q. Tang, and Y. Song, *Enhancing Mixed Traffic Flow Safety via Connected and Autonomous Vehicle Trajectory Planning with a Reinforcement Learning Approach.* Journal of Advanced Transportation, 2021. **2021**: p. 6117890.

15. Cheng, Y., X. Hu, K. Chen, X. Yu, and Y. Luo, *Online longitudinal trajectory planning for connected and autonomous vehicles in mixed traffic flow with deep reinforcement learning approach.* Journal of Intelligent Transportation Systems, 2022: p. 1-15.

16. Eskandarian, A., C. Wu, and C. Sun, *Research Advances and Challenges of Autonomous and Connected Ground Vehicles.* IEEE Transactions on Intelligent Transportation Systems, 2019: p. 1-29.

17. Guanetti, J., Y. Kim, and F. Borrelli, *Control of connected and automated vehicles: State of the art and future challenges.* Annual Reviews in Control, 2018. **45**: p. 18-40.

18. Khoury, J., K. Amine, and R. Abi Saad, *An initial investigation of the effects of a fully automated vehicle fleet on geometric design.* Journal of Advanced Transportation, 2019. **2019**.

19. McDonald, D. *How might connected vehicles and autonomous vehicles influence 34 geometric design?* In *Proceedings of the Transportation Research Board 97th Annual Meeting, Washington, DC, USA*. 2018.

20. Noorvand, H., G. Karnati, and B.S. Underwood, *Autonomous Vehicles:Assessment of the Implications of Truck Positioning on Flexible Pavement Performance and Design.* Transportation Research Record, 2017. **2640**(1): p. 21-28.

21. Chen, F., M. Song, X. Ma, and X. Zhu, *Assess the impacts of different autonomous trucks' lateral control modes on asphalt pavement performance.* Transportation Research Part C: Emerging Technologies, 2019. **103**: p. 17-29.

22. Chen, F., R. Balieu, and N. Kringos, *Potential Influences on Long-Term Service Performance of Road Infrastructure by Automated Vehicles.* Transportation Research Record, 2016. **2550**(1): p. 72-79.

23. Gungor, O.E. and I.L. Al-Qadi, *Wander 2D: a flexible pavement design framework for autonomous and connected trucks.* International Journal of Pavement Engineering, 2022. **23**(1): p. 121-136.

24. Gungor, O.E., R. She, I.L. Al-Qadi, and Y. Ouyang, *One for all: Decentralized optimization of lateral position of autonomous trucks in a platoon to improve roadway infrastructure sustainability.* Transportation Research Part C: Emerging Technologies, 2020. **120**: p. 102783.

25. AASHTO, *Mechanistic-empirical pavement design guide: a manual of practice*. 2008: AASHTO.

26. Underwood, B.S. and W.A. Zeiada, *Characterization of microdamage healing in asphalt concrete with a smeared continuum damage approach.* Transportation Research Record, 2014. **2447**(1): p. 126-135.

27. Motevalizadeh, S.M., P. Ayar, S.H. Motevalizadeh, S. Yeganeh, and M. Ameri, *Investigating the impact of different loading patterns on the permanent deformation behaviour in hot mix asphalt.* Construction and Building Materials, 2018. **167**: p. 707-715.

28. Mollenhauer, K., M. Wistuba, and R. Rabe. *Loading frequency and fatigue: In situ conditions & impact on test results*. In *2nd Workshop on four point bending*. 2009. University of Minho Guimarães, Portugal.

29. Blab, R. and J. Litzka. *Measurements of the lateral distribution of heavy vehicles and its effects on the design of road pavements*. In *Proceedings of the international symposium on heavy vehicle weights and dimensions, road transport technology, University of Michigan*. 1995.

30. Noorvand, H., G. Karnati, and B.S.J.T.R.R. Underwood, *Autonomous vehicles: assessment of the implications of truck positioning on flexible pavement performance and design.* 2017. **2640**(1): p. 21-28.

31. Maria Coca, A., S.A. Romanoschi, M. Talebsafa, and C.J.J.o.T.E. Popescu, Part B: Pavements, *Simple Method for Incorporating Lateral Wheel Wander in the Computation of Fatigue Damage in New Flexible Pavement Structures.* 2020. **146**(4): p. 04020074.

32. Wu, R. and J.T. Harvey. *Evaluation of the effect of wander on rutting performance in HVS tests*. In *Proceedings of the 3rd international conference on accelerated pavement testing*. 2008.

33. Sinanmis, R. and L. Woods, *Relationship between channelisation and geometric characteristics of road pavements.* International Journal of Pavement Engineering, 2021. **22**(11): p. pp 1446-1453.

34. Zhou, F., S. Hu, S.T. Chrysler, Y. Kim, I. Damnjanovic, A. Talebpour, and A.J.T.r.r. Espejo, *Optimization of lateral wandering of automated vehicles to reduce hydroplaning potential and to improve pavement life.* 2019. **2673**(11): p. 81-89.

35. Bimbraw, K. *Autonomous cars: Past, present and future a review of the developments in the last century, the present scenario and the expected future of autonomous vehicle technology*. In *2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*. 2015. IEEE.

36. Song, Y., Y.D. Wang, X. Hu, and J. Liu, *An efficient and explainable ensemble learning model for asphalt pavement condition prediction based on LTPP dataset.* IEEE Transactions on Intelligent Transportation Systems, 2022. **23**(11): p. 22084-22093.

37. *Traffic Assignment Manual for Application with a Large, High Speed Computer*, U.B.o.P. Roads, Editor. 1964, US Department of Commerce.

38. Gerlough, D.L. and M.J. Huber, *Traffic flow theory*. 1976.

39. Wen, Z., H. Liu, J. Shi, Q. Li, B. He, and J. Chen, *ThunderGBM: Fast GBDTs and Random Forests on GPUs.* J. Mach. Learn. Res., 2020. **21**(108): p. 1-5.

40. Friedman, J.H., *Greedy function approximation: a gradient boosting machine.* Annals of statistics, 2001: p. 1189-1232.

41. Chen, T. and C. Guestrin. *Xgboost: A scalable tree boosting system*. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016.

42. Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, *Lightgbm: A highly efficient gradient boosting decision tree.* Advances in neural information processing systems, 2017. **30**: p. 3146-3154.

43. Prokhorenkova, L., G. Gusev, A. Vorobev, A.V. Dorogush, and A. Gulin, *CatBoost: unbiased boosting with categorical features.* arXiv preprint arXiv:1706.09516, 2017.