Leveraging LiDAR-Based Simulations to Quantify the Complexity of The Static Environment for Autonomous Vehicles in Rural Settings

by

MOHAMED ABOHASSAN

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Transportation Engineering

Department of Civil and Environmental Engineering
University of Alberta

ABSTRACT

The advent of autonomous vehicles promises to address key challenges faced by the transportation and traffic safety sectors, particularly concerning the high incidence of traffic fatalities and injuries. The full-scale deployment of autonomous vehicle fleets is contingent upon public acceptance, which is concerned about the available models' safety levels, evidenced by the multiple reported AV-related traffic incidents. Achieving the necessary safety standards for autonomous vehicles necessitates comprehensive and rigorous testing. Virtual simulations take precedence over road and track testing, as they expedite AV development by enabling safe, cost-effective, and large-scale testing.

AVs face a big data problem stemming from the vast volume of collected sensory information data surpassing real time processing capabilities. This mismatch often leads to driving errors and traffic collisions. Although different models to quantify the complexity of the environment have been explored, they often lacked objectivity, ignored the physics of the simulated sensors, and failed to utilize robust LiDAR data. Hence, this study aims to develop a framework that can transform the complexity of the environment into easy-to-understand values while accounting for the drawbacks of the existing frameworks.

This study conducted virtual simulations on 34 kilometers of LiDAR-based digital environments sourced from two-way-two-lane rural roads in Alberta, Canada. The complexity of the environment was assessed by calculating the required real time data rates essential for AVs to maintain their regular navigation functions. The data rate requirements encompass the sensor's specifications, the surrounding environment's dynamicity, and weather conditions in its calculations. Furthermore, this study focused on the static physical environment only, dividing the analyzed environments into RRFs and FE to isolate and study the impact of general road geometry features such as vertical curves, horizontal curves, and roadway

width on the complexity of the environment and quantify the extra computational burdens incurred from the analysis of the entire section to highlight the problem of WVC for AVs.

This study employed two distinct approaches to demonstrate the versatility of the implemented framework and explore diverse perspectives on environment perception for AVs: the primary occupancy method and the secondary volumetric method, where the advantages and drawbacks of each method are highlighted.

The study analyzed fluctuations in data rate requirements along the vehicle's trajectory by dividing the surveyed roadways into distinct frames and generating novel synthetic viewpoints to faithfully replicate real-world situations. This process was facilitated using the open-source Vista simulator. The analysis revealed that the environment could be deemed complex in two scenarios. Firstly, when data rate values spiked, it indicated a substantial volume of information that needed processing. Secondly, when data rates plummeted, it signified a significant loss of information, potentially jeopardizing the vehicle's operations.

The analysis results indicated that the roadside features are anticipated to escalate environmental complexity by 140-400% based on density. Additionally, widening the roadway by adding an extra lane was observed to raise processing requirements by 12.3-16.5%. Concerning road alignment, crest vertical curves were found to decrease data rates by up to 4% due to occlusion challenges at these points, while sag curves increased requirements by 7% due to enhanced visibility. In horizontal curves, roadside occlusion contributed to a decrease in data rate requirements by up to 19%. As for the weather conditions, heavy rain increased the AV's processing demands by a significant 240% when compared to normal weather conditions.

The developed framework and results, supported by statistical testing, can help AV developers make more informed decisions by understanding the impact of the different road elements. Moreover, government

agencies and IOOs can also exploit the findings of this study to accommodate AV requirements in the current human-tailored road designs and optimize future designs for AV deployment.

PREFACE

Part of the work presented in this thesis is being prepared for journal publication.

- **Abohassan, M.**, El-Basyouny, K. " *Utilizing Lidar-Based Digitized Infrastructures to Estimate the Influence of the Static Road Features on the Performance of Autonomous Vehicles* "

ACKNOWLEDGMENTS

I am deeply grateful to my supervisor, Dr. Karim El-Basyouny, for his unwavering support, invaluable guidance, and insightful advice throughout this research endeavor. His boundless knowledge and expertise have been instrumental in shaping this study. Without his trust, dedication, and continuous assistance, this research would not have come to fruition. I would also like to extend my thanks to Dr. Sharon Harper for her assistance in proofreading my work.

I want to thank the professors at the Department of Transportation Engineering at the University of Alberta for providing me with the technical knowledge and tools that helped realize this research. Special thanks go to Maged Gouda for his assistance throughout my studies and research.

I am extremely indebted to my family for their unconditional support throughout my journey. In particular, I am immensely grateful to my brother, Ahmed Abohassan, whose steadfast encouragement and presence were my pillars of strength during this formidable challenge, a debt I can never fully repay. I can not also forget the role of my friends in this success, especially Marwan Ayman and Mostafa Abd El-Bary, for providing me with a safe and understanding space whenever I needed to vent and express my frustrations.

Acknowledgments are extended to Dr. Tae Kwon and Dr. Stephen Wong for being part of my examination committee.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 BACKGROUND

Recent technological advancements have promised to revolutionize the world and accelerate the ongoing transition toward a more Artificial Intelligence (AI)-driven future. From a transportation and traffic safety standpoint, AI technology continues to evolve and become more intricate, leading to its successful integration into various transportation fields. That is why it is expected, in the coming years, to play a leading role in addressing some of the most pressing challenges in the transport sector, such as public health and safety concerns, the increase in travel demand, and environmental problems [1].

Traffic injuries have long been a significant public health issue, which has been the case ever since the advent of automobiles and the very first recorded traffic death in the late 19th century [2]. In 2018, the World Health Organization reported that 1.35 million people die on the road prematurely every year, which means that there are over 3,500 traffic-related deaths in the world daily [3], not to mention the tens of millions who suffer from debilitating injuries caused by road collisions. What is even more concerning about these figures is that the prime subjects for deaths from road traffic injuries are mainly people aged between 5 and 29, meaning that they are primarily children and economically productive young adults [3]. In the U.S., for example, the 2021 toll of vehicular accidents is a staggering $498 billion [4]. Furthermore, the government of Canada revealed that 1768 road fatalities occurred in the year 2021 [5].

Engineers have managed to attenuate the cost of traffic accidents by adopting a paradigm shift in how they view traffic collisions. This shift has changed the definition of collisions from random and inevitable to predictable and preventable. It has also shifted the responsibility of traffic safety from the transportation sector to become multidisciplinary, requiring different stakeholders' expertise. Accommodating human errors and providing road systems that are forgiving in nature was also a vital part of the paradigm shift and has helped to change many policies in multiple countries [6]. Despite the success achieved in decreasing the number of collisions through the paradigm shift in recent years, human driving errors remain a persistent problem that pushed researchers to look elsewhere for more innovative solutions that could provide safer road environments [7].

Numerous elements play a role in road accidents, as highlighted in [8]. Among these, vehicle malfunctions, environmental conditions, and human factors are the primary instigators. To further emphasize the harmful role of human driving errors on traffic safety, it has been well documented that human errors were the leading cause of traffic collisions, with an overwhelming 94% of errors leading to collisions through actions like speeding, driving under the influence, and distracted driving [7]. The number of fatalities and injuries is staggering, reflecting the gravity of the situation and the necessity to find solutions.

Autonomous Vehicles (AVs) offer a solution to dissociate humans from vehicles, thus eliminating human driving errors by operating independently without any human intervention [8] [9] [10]. Also, with the tremendous improvement in both research and industry that the AVs have witnessed in the past decade, they present themselves as a solid alternative that can potentially change the future of transportation. That is why this topic has captured the attention of experts from both the public and private sectors and infrastructure owners/operators (IOO) worldwide [11].

AVs have many advantages, particularly because they offer the promise to provide safer roads. A recent study has shown that with the elimination of human error, the current traffic collision figures would go down by 70% in 25 years [12]. The adoption of AVs will also contribute to increased road capacity, reduced parking costs, reduced energy consumption, reduced driver stress, and increased productivity [13]. Additionally, AVs would provide people with disabilities, like visual impairments, the opportunity to travel safely [14]. Though there might be potential drawbacks to the deployment of AVs like increased vehicle and infrastructure costs, reduced security and privacy, and concerns about system failures, which might be detrimental to the whole premise of AVs [13], researchers are closing these gaps and addressing the concerns with AVs at rapid speeds. Therefore, prominent automakers such as Ford and General Motors, leading IT companies like Apple and Google, and popular ride-hailing platforms such as Uber and DiDi have dedicated substantial resources to develop advanced autonomous vehicle (AV) models in order to secure a strong foothold in the market [6].

AVs commonly rely on answering three questions for their navigation operations: first, where is the vehicle located? Second, what objects surround the vehicle? And third, what is the safe action to take in its next move? [15]. AVs draw upon the capabilities of the equipped sensors

to answer these questions by continuously gathering information about the environment in a manner that is analogous to human drivers, then relaying it to the onboard computer, which processes the assimilated data and executes the required manoeuvres, including, steering, swerving, speeding and stopping [15]. Ergo, high-quality sensory information is the linchpin for robust and reliable AV operations. AVs usually carry multiple sensors for real time data acquisition and environment perception. These sensors typically include LiDAR, Radar, and cameras, among others.

Light Detection and Ranging (LiDAR) is an advanced optical remote sensing technology that collects information about the surrounding environment. LiDAR uses laser scanners, sensors, Global Navigation Satellite System (GNSS) receivers, and Inertial Measurement Units (IMU) to collect their data. The basic working principle for LiDAR is the constant emission of laser beams from the laser scanners that are reflected off objects to measure their distance. Thousands of beams are emitted per second, resulting in millions of distance measurements, ultimately creating a 3D point cloud map of the surrounding environment [16]. LiDAR has been in high demand in recent years owing to its accurate and robust measurements and high scanning speed, enabling it to collect more than 200,000 points per second, depending on the unit model [17]. Hence, the revenues from LiDAR sales are anticipated to skyrocket, reaching figures close to $6910 million by 2025 [17].

Radio Detection and Ranging (Radar) relies on emitting electromagnetic waves with specific frequencies. The scattered waves are then collected to collect range information about the targets in range. The Doppler property of EM waves is exploited to determine the relative speed and position of the detected targets around the radar [18]. On the other hand, cameras are one of the most used and widely available technologies for perceiving surroundings. The basic principle of camera operations is detecting light from the surroundings. They provide AVs with high-resolution coloured images [17].

Each of these sensors possesses strengths and weaknesses. For example, radar is unaffected by lighting conditions, whereas cameras are sensitive to external lighting conditions [19]. LiDAR and cameras are susceptible to bad weather, unlike radar, which is not [19]. Radar outputs are usually characterized by their low resolution [19]. LiDAR, on the other hand, is superior in range accuracy and angular resolutions [20]. Finally, LiDAR is the most expensive option, whereas cameras are the cheapest [19]. Sensor redundancy occurs when multiple sensors address a task.

Thus, state-of-the-art AV models maximize the robustness and reliability of their performance by sporting different sensors to complement each other and achieve high levels of sensor redundancy [21].

The onboard computer is the cornerstone of autonomous systems. It plays a pivotal role in their operations as it makes sense of the surroundings and executes the appropriate actions to maintain safe operations. Typically, the computer must perform four core, interconnected tasks to achieve safe autonomous driving. These tasks can be summarized as follows: environment perception, localization and mapping, decision-making and planning, and vehicle control [22].

The initial task, known as environment perception, involves collecting information about the surroundings and extracting the relevant features. These features might be static, like road obstacles and signs, or they can be dynamic, like the movements of the other road users. This task is done through the assimilation and fusion of data provided by the equipped sensors of the AVs [23]. For highway environments, AVs generally prioritize detecting obstacles within the road itself rather than all the potential obstacles in the entire scene [24].

This task entails two primary assignments. First, the object detection assignment is responsible for accurately detecting the different road objects, even under challenging road and weather conditions [25]. The second assignment is related to lane detection, an essential feature in all ADAS since it allows AVs to stay within the lane boundaries and avoid different obstacles [26]. Currently, deep-learning techniques have dominated this task as much research has been directed toward developing algorithms that have faster and more accurate results [27] [28] [29] [30].

Localization and mapping are the second core tasks in autonomous driving. Localization involves finding the accurate location of the ego vehicle relative to a map [31]. As for the mapping, the AV builds multilayer HD maps for path planning [32]. The onboard computer ensures effective navigation and manoeuvring by maintaining an accurate position on the road. GPS-IMU-based localization methods are the most common. However, the accuracy of this system does not fulfill complex environmental requirements [33]. More recently, prebuilt HD maps have been proven to be more practical and more accurate than traditional maps, offering more functionalities to the AV by supporting its perception and localization modules through the extraction of the critical static properties such as roads, traffic lights, road markings, buildings and trees [34].

Furthermore, SLAM (simultaneous localization and mapping) has received much attention recently, especially in AVs [35]. SLAM attempts to construct maps and localize the ego vehicle simultaneously [36]. This algorithm can be LiDAR-based, like LOAM [37], which can be executed in real time, or IMLS-SLAM, which prioritizes the reduction of the accumulated drift [38]. Otherwise, SLAM can be camera-based by implementing frame-to-frame matching [39], achieved through feature-based matching [40] or direct matching [41].

The third task involves decision-making and planning. To optimize the decision-making task, the AV is expected to predict the driving patterns of the surrounding road users and assess the corresponding risks [42]. The HMM model (hidden Markov Model) is commonly used by autonomous systems in anticipating a target vehicle's actions on the road [43]. Planning an expedient route from an origin to a destination on the map is done through local planners seeking to balance the sampling size with the computation efficiency to refine the shortest and safest path to the ego vehicle [44].

Upon collecting all the necessary information and planning the vehicle's route, the final task is vehicle control. This task is done through lateral and longitudinal controllers expected to handle different road surfaces. Control laws dictate the output commands given to the AV based on the state of the vehicle and its trajectory. Various models of control laws attempt to optimize different parameters, such as fuzzy control, PID control, Stanley control, and model predictive control (MPC) [45] [46] [47].

AVs, in terms of operation, can be classified into two systems: modular-based and end-to-end-based [48]. The modular-based system treats each module individually. As such, localization, perception, control modules, etc., are decoupled, which facilitates bringing people from different backgrounds together [49]. Additionally, this model is more interpretable as each module is well-defined. Nevertheless, they are susceptible to wasting computations on unnecessary tasks like identifying objects unrelated to driving tasks and error propagation [49].

On the other hand, the end-to-end-based system highly relies on artificial intelligence, where machine learning is used to process the sensor data and then generate the necessary commands. End-to-end systems offer more efficient computations as they are self-optimizing, a sought-after trait for autonomous systems. However, the black-box nature of deep learning renders their results unverifiable [49].

The perception of the environmental task is of utmost importance in autonomous driving. It serves as a fundamental step because the onboard computer must comprehensively understand the surroundings through environment perception to achieve optimal results in subsequent tasks, such as decision-making and planning. AVs can make informed decisions and plan actions by accurately perceiving the environment, leading to safe and efficient autonomous driving [22]. The main goal of this task is to extract the relevant features necessary for navigation operations due to the big data problem caused by the vast amount of data fed by the different sensors, increasing the computational complexity in processing this volume of data [50] [51] [52] [39]. This task is typically done by pruning the irrelevant features of the road through preprocessing techniques as part of intelligent data prioritization mechanisms [51]. The pertinent features mainly comprise road surfaces like lanes and boundaries, traffic signs and signals, vehicles, and pedestrians [22] [23].

According to the recently published ITF report, experts agreed that having a network that only serves AVs would solve various problems by providing highly controlled environments. However, this option was deemed infeasible due to the lack of available space, particularly within urban areas, to build new networks solely dedicated to AVs [53]. Furthermore, it is believed that having parallel networks for AVs would disincentivize their use of shared road networks, precluding the possibility of achieving many benefits promised by AVs [53]. Accordingly, AVs will almost certainly use the available road networks and will have to circumvent their limitations [53].

What's more, AVs' operations, at present, have been limited to testing and piloting initiatives; their activities have been confined to specific regions where road conditions are known and the environments are reasonably predictable [53]. Developers have embraced this approach to ensure repetitive experiences lead to continual enhancements. This iterative process holds the potential to fully capitalize on the advantages of automation [53]. Nevertheless, this approach has also limited the areas where AVs can be confidently deployed. The current data acquired from laboratory-based approaches and typical test tracks are insufficient to assess AVs' safety. Some estimates believe that, for AVs to achieve a 95% confidence that their failure rates are better than human driver failures, they must drive over 275 million miles [54]. The most optimistic projections believe this feat can be achieved in 84 years [55].

1.2 MOTIVATION

While the development of level 5 full automation has not progressed at the anticipated pace [56], there has been significant progress in achieving advanced technical improvement. Thus, the vision of attaining level 4 automation on the roads became a reality, although the journey proved arduous. On the other hand, the dreams of deploying level 5 autonomous vehicles, for now, still seem to be far-fetched [57].

Nevertheless, significant setbacks have also been encountered along the way. System failures for levels four and five driving have been reported to cause collisions on multiple occasions. One occasion involved an ADS in a crash due to rain during the Hyundai competition [58]. Another collision occurred when Google's ADS failed to estimate the speed of a passing bus when performing a lane change [59]. A failure of onboard sensors was reported to be a key cause of a collision for another automated vehicle [60]. In 2018, a fatal collision was reported in Tempe, Arizona, between a human and an automated vehicle run by Uber [61]. And a Tesla autopilot failed to recognize a truck, killing the truck driver in the collision [62]. In California, where AV testing is permitted on public roads, there have been 303 reported AV collisions between 2015 and April 2021 for different AV models operating in autonomous mode or right after disengaging from conventional mode [63]. Therefore, it is paramount to understand how AVs perceive their surroundings and identify their targets to execute their manoeuvres safely since the public acceptance of AVs is intricately linked to their ability to prevent traffic collisions [64].

One category of collisions that demands the attention of AV developers is Wildlife-Vehicle Collisions (WVCs). These collisions have far-reaching consequences as they seriously threaten the safety of drivers and animals involved. Every year, disconcerting statistics emerge from the US concerning the repercussions of WVCs. These incidents contribute to a troubling toll, with over 59,000 injuries and more than 440 human fatalities attributed to this type of collision [65]. The average cost of such collisions was reported as $885 per collision [66], and the associated costs amounted to around $6 to 12 billion [67]. In the year 2018, the state of California documented a staggering material damage cost of $76 million resulting from significant WVCs [68]. Similarly, in Sweden, the annual expenses attributed to WVCs amounted to a substantial estimate of around $406 million [69]. In Alberta, it was reported that WVCs contributed to 60% of all reported vehicle accidents transpiring on its rural highways [70].

Predominantly, the object detection algorithms integrated into AVs prioritize the road lanes, road signs and signals, vehicles, pedestrians, and cyclists [71] [72]. However, a notable imbalance exists in the attention given to developing methodologies for animal detection [73] [74]. AVs deployed in busy urban areas will experience complex scenarios that require advanced decision-making, such as crossing crowded intersections [75]. Likewise, AVs are expected to encounter complicated situations to handle WVC in rural areas safely [78].

Animal detection necessitates high detection precision and rapid response time to avert potential collisions with wildlife effectively. First, The AV has to discern the presence of an animal near the driving lanes. Subsequently, vigilant monitoring of the detected animal's movements ensues, and predictive analysis is applied to anticipate its trajectory, thus gauging the risk of collision. This predictive analysis will ultimately allow the ego vehicle to make informed decisions that prevent impending collisions [76]. Hence, in environments with roadside vegetation, the AV has to be able to spot animals lurking in the bushes before they enter the road by processing this region of the road.

For domains where safety is of the utmost importance, such as AVs, simulation-based approaches are indispensable for developers and researchers [77]. Such approaches can offer a safe context to conduct necessary large-scale testing and validations through the digitization of the environment [78]. In addition, since physical testing for AVs is generally banned inside most cities, the only real-life substitute would be time-consuming and expensive field tests, which are also hard to reproduce and often seen as inconvenient [72]. Simulations are often considered the better approach to off-road and inaccessible environment tests [79].

Moreover, virtual simulations are both labour-saving and cost-effective [80]. The high-fidelity 3D point cloud data from LiDAR is particularly effective due to its millimeter-level accuracy [11]. This LiDAR data can be used to create Digital Twins [81] of the physical environment and opens the door for risk-free simulation environments, especially since the sought-after safety levels cannot be achieved until billions of kilometres of testing by the AVs under different weather conditions [82]. An analysis conducted by Kalra and Paddock [82] has statistically shown that achieving a reliability level that can guarantee a fatality rate of 1.09 per 160 million kilometres, with a confidence level of 95% along with an 80% power that the AV failure rates are at least 20% better than humans, will require approximately an extra 500 years.

There are two types of LiDAR data with respect to their mode of collection: first, aerial collection using drones or helicopters, and second, terrestrial collection, which can either be static by mounting the LiDAR on a tripod or mobile by setting up the system on a moving platform. Mobile laser scanning (MLS) has been widely recognized to be the more expedient form of LiDAR data as it is known for the high level of detail for its outputs, cheap cost of operation, high point density, and the ability to scan below bridges and inside tunnels [83].

The tremendous amount of information that is relayed to the AV through its equipped sensors poses a great challenge for their real time processing capabilities which are not yet advanced enough to handle this surge of data [73] [39]. Hence, it is essential to understand the factors that influence the processing demands and how will they vary under different road and weeather conditions.

In an attempt to accelerate the progress of autonomous driving, a recently published white paper by Neural Propulsion Systems [84] has explored the feasibility of developing a scanning system to eliminate preventable roadway deaths for level 4 AVs. The vehicle would have to see well enough to enable zero roadway deaths to have this system in place. In other words, this system must handle a massive surge of sensor data to assimilate and process in real time. Principles of physics and information theory were implemented to develop an equation that estimates the amount of data that would allow the vehicle to see well enough. This equation delineates the amount of data required to recreate the scene around the AV with enough frequency and fidelity to provide the vehicle with sufficient stopping sight distance, thus preventing the possible collision with any obstacles along its path. The required data rate values are contingent upon the density of the points around the vehicle or, in other words, the complexity of the surroundings. The data requirements are expected to be dynamic as they would vary with the change of the scene around the vehicle at any instant.

This research fundamentally attempts to bridge the existing gap within the literature by presenting a framework to facilitate an understanding of the surrounding environment's influence on AVs' performance. To accomplish this, the study implements a simulation-based, disaggregated approach that breaks down the virtual environment made of LiDAR point cloud data into individual frames. The reconstruction of the scenes around the vehicle at any instant on the road allows for the thorough analysis of each scene separately to determine the necessary data rates

required to guarantee the safe operation of the AVs. This microscopic and detailed analysis offers a powerful tool to see through the vehicle's sensors and understand the intricacies of the surroundings. Within this study, LiDAR is used as the sole visionary sensor in the simulations since it's the preferred sensor in most driverless vehicles [85] [86] [87].

This algorithm uniquely captures environmental complexity, something not previously implemented in prior studies. By basing its calculations on real-world conditions and commercially available sensors, this approach transforms what was once a generic value into a specific, deterministic representation. The nuanced approach offered here not only differentiates this study from past work but also enables an explanation of the intricacies within the environment, setting a new standard in understanding complexity in this domain.

1.3 OBJECTIVES

This thesis attempts to gain a better and deeper understanding of the processing demands for AVs under different road and weather conditions by examining the dynamic nature of data rate requirements. More specifically, this research takes a unique approach by quantifying the impact of different road components, including vertical and horizontal curves, roadway width, and roadside features. It then translates them directly into processing requirements for AVs. This method simplifies the identification of critical sections, facilitating a more strategic approach to AV deployment.

The primary objective of this research is to develop a framework that explains the interaction between the static environment and the performance of AVs by utilizing a transformation function that captures the complexity of the surrounding environment and weather conditions, translating them into interpretable data rate requirements. In fulfilling the main objective, several secondary objectives regarding the possible factors that might affect the complexity of the environment will be addressed such as:

- **Investigating the effect of occlusion on AV performance and data rate requirements.** The occlusion challenge can be investigated by utilizing the proposed methodology. Occlusions occur when an object in the foreground blocks the vision of a target in the background [88] and AVs occlusion handling is, to date, one of the most challenging tasks [89]. Such a problem is bound to have implications on the data rate requirements and the

performance of the AVs in general. The proposed method can be used to investigate this problem and identify its potential causes, like horizontal curves, vertical curves, and vehicles on the road or roadside elements like vegetation.

- **Evaluating the impact of altering roadway width on the performance of AVs.** Grasping the implications of varying roadway widths on the complexity of the environment is paramount for ensuring the seamless integration of AVs into diverse urban landscapes. Quantifying this effect is crucial as it provides tangible data, enabling researchers and developers to design AV systems that can adeptly traverse different road environments where the number of driving lanes is bound to change occasionally.
- **Analyzing how roadside features affect the environmental complexity.** Roadside elements, particularly vegetation, can be notably inconsistent in their distribution, creating challenges for AVs in different scenarios. Understanding this variability is crucial as it directly affects the AVs' ability to handle wildlife collisions. Uneven distribution of vegetation can obstruct visibility and sensor signals, making it challenging for AVs to detect and respond to wildlife effectively. By delving into the intricacies of these roadside features, researchers and developers can gain valuable insights into how AVs can adapt to diverse environments and overcome the excess computational requirements of processing them in real time.
- **Comparing simulation outcomes based on different LiDAR sensors.** The presented framework can demonstrate the performance of various LiDAR sensors available within the virtual environment simulations. As such, the influence of the LiDAR sensors and their field of vision is simulated through the visual environment, and a direct comparison is made possible. Consequently, users can gain insights into the effectiveness and reliability of different LiDAR.
- **Investigating the implications of different weather conditions on the necessary processing power for AVs.** Exploiting the virtual environment can offer unparalleled versatility in simulating different weather conditions by manipulating the signal-to-noise ratio variable. Under favourable weather conditions, a high SNR value is expected. While in adverse weather situations, this value dramatically decreases to simulate the challenges of processing incoming data from the environment. This effect further highlights the

importance of employing a simulated approach since the on-field testing is often limited to favourable weather conditions for obvious safety concerns.

Adopting a simulation-based approach that utilizes LiDAR as the primary sensor for data acquisition and environment perception makes it possible to model AVs' performance under varying driving conditions. This ability, in turn, has the potential to predict the performance of AVs prior to deployment. Furthermore, achieving this research's primary and secondary objectives could yield invaluable insights for IOOs regarding infrastructure readiness for AVs.

1.4 RESEARCH CONTEXT

This investigation used LiDAR point cloud data sourced from highways in Alberta, Canada, employing the advanced PSP-7000 scanning vehicle equipped with double RIEGL VMX-450 laser scanners. The study predominantly focuses on the intricate dynamics of two-lane-two-way roads. The analysis was conducted in rural settings as their general layout is less complicated which facilitates the isolation of the different physical road features so that their influence on the performance of the AVs can be investigated.

34 km of highways were considered, encompassing diverse layouts, including an 11 km stretch of road featuring varying vertical grades, to delve into the profound impact of vertical curves. Additionally, 15 km of roads were scrutinized for their varying widths. Furthermore, 8 km of segments with critical horizontal curves were meticulously examined.

Furthermore, in the applied simulations, the autonomous vehicles' visual perception was replicated using different Velodyne sensors owing to their wide adoption in AVs [90], ensuring an accurate representation of the environment. To simplify the analysis, the research excludes the consideration of dynamic objects such as pedestrians, cyclists, and vehicles. This decision was made to reduce the intricacy of tracking moving objects and anticipating their behaviour, particularly in busy driving environments [91] [92]. Moreover, none of the currently implemented algorithms used to detect pedestrians are advanced and reliable enough to operate in real time and with high accuracy [93]. Therefore, this study focuses on rural environments and limits the analysis to two-way-two-lane roads for consistency.

1.5 THESIS STRUCTURE

The remainder of this thesis is divided into five chapters. Chapter two reviews the literature on several topics, including Autonomous Vehicles, Simulations on Autonomous Vehicles, Applications on Digitized infrastructure using LiDAR point clouds, and Quantification of the AVs' environment complexity. Chapter three outlines the adopted methodology in this research along with the introduction to the tested road segments. Chapter four discusses the results of different simulations with a summary of the main findings. Chapter five offers a comprehensive overview discussion on the results of its previous chapter. Finally, Chapter six highlights the study's main findings and the research contributions and defines the research limitations along with future research prospects.

2 LITERATURE REVIEW

2.1 AUTONOMOUS VEHICLES

The notable advancement of technology has helped to alter the goal of vehicles from a mere means of transportation to the inclusion of comfort, safety, and convenience. The extensive research in the field of vehicles, where the incorporation of breakthrough technologies and improvements has been thoroughly discussed, concluded that introducing the idea of self-driving vehicles is the next groundbreaking innovation that would serve as the pinnacle of exhaustive research in this field.

Attempts to develop a self-driving vehicle started with the advent of the 20th century. In New York City, the Linriccan Wonder was a model equipped with an antenna controlled by a following vehicle using radio impulses. These radio signals dictated the model's movements and directions, making it one of the earliest attempts and rudimentary forms of AVs [94]. This model was later improved and renamed Phantom Auto and was displayed in December of 1926 in Milwaukee by Achen Motors.

Others attempted to develop embedded-circuit-powered electric vehicles. Typically, circuits were embedded in the test roadways or labs and were controlled by radio waves. This strategy was demonstrated by the efforts of GM (General Motors) when they promoted Norman Bel Geddes's Futurama at the World's Fair in 1939 [94]. RCA Labs developed something similar where embedded wires controlled their developed small-scale model on a laboratory floor. A larger-scale deployment of the RCA model was soon followed by Leland Hancock, where the experiment was carried out on a 121.92m section of an actual highway [94].

The United Kingdom's Transport and Road Research Laboratory demonstrated the success of driverless vehicles with a Citroen DS model on a test track in 1960. The vehicle relied on magnetic cables embedded into the road for its navigation. It was reported that the vehicle managed to maintain a speed of 130 kph without any significant deviation of either speed or direction through different weather conditions, surpassing the human control capabilities at that time [94]. Comparably, The Phileas electronic guidance system in Eindhoven, Netherlands, contributed to the success of one of the earliest autonomous public transport systems in the early 1990s. The magnetic system allowed the vehicles to operate at speeds that reached 70 kph and provided two

modes of operation, either fully autonomous or semi-autonomous, by exploiting the embedded magnetic markers inside the road surface to guide the vehicles through the streets [94].

Video monitors and video tape recorders were used for navigation in 1991 when the University of Bundeswehr Munich tested out the twin robot vehicles VaMP and Vita-2 on a three-lane highway in Paris for more than 1000 km under standard traffic conditions and with speeds that exceeded 120 kph. During this experiment and with occasional human interventions, they showcased the ability of their models to change lanes, follow other vehicles closely, and even pass other vehicles [95].

The introduction of computer vision technology was first adopted by the Autonomous Land Vehicle Project, funded by the U.S. Department of Defence. The developed models relied on LiDAR and autonomous control to navigate at speeds that reached 31 kph [96]. Later on, HRL Laboratories tested the AVL model along 610m of an off-road setting. During the test, the vehicle maintained a speed of 3.1 kph while navigating through natural obstacles, vegetation, and steep slopes. This experiment emphasized the potential of LiDAR technologies in autonomous navigation [97]. In another test in 1995, Dickmann's autonomous S-class Mercedes-Benz embarked on a roundtrip between Germany and Copenhagen, where the vehicles drove on the Autobahn with speeds exceeding 175 kph. The vehicles incorporated radar-based computer vision technologies and microprocessors with memories, allowing real time reactions. The models were reported to achieve a solid 95% autonomous driving whilst successfully carrying out a wide range of driving manoeuvres [98].

Machine learning algorithms were first introduced in 1995 to assess their potential to improve autonomous driving performance. Navlab of Carnegie Mellon University developed a semi-autonomous car that utilized neural networks to control the steering manoeuvres. The 5 km cross-country trip known as "No Hands Across America" or NHOA demonstrated the ability of the vehicle to attain 98.2% autonomous driving, albeit the braking and throttle tasks were still the onus of humans [94].

The Defense Advanced Research Project Agency (DARPA) launched a competition in 2003 that required the participating AVs to complete an off-road desert course without the aid of road markings [99]. Although no vehicles completed the first Challenge held in 2004, five AVs succeeded in the 2005 grand DARPA challenge, marking a huge milestone for the research

community [100]. The DARPA *urban* challenge took place in 2007 to address the more complex tasks, such as driving in urban areas with heavy traffic and frequent intersections [99]. Four vehicles successfully met the challenge, including teams from Carnegie Mellon University, Stanford University, Virginia Polytechnic Institute and State University, and the Massachusetts Institute of Technology. The competing teams used different sensing, localizing, and perception techniques [101].

The VisLab Intercontinental Autonomous Challenge (VIAC) initiated a more advanced test challenge in 2010 to expose AVs to a more realistic real-world environment, including common obstacles such as pedestrians and cyclists. Two AVs (a leader and a follower) were involved in this challenge, driving across multiple European and Asian countries. The vehicle encountered different urban, highway, and offroad scenarios throughout this challenge. In addition, the following vehicle detected the leading vehicle and employed platooning techniques. They were successful in completing this challenge with the absence of any prior knowledge of the road information. [102].

The heavy reliance on priori information like HD maps, which carry static information about the surrounding environment [103], has been observed by several AV developers such as Google [104]. Although this approach helps tremendously in decreasing the computational demands on the AV, it also restricts the adaptability of the AVs to novel situations like construction zones and potholes [103]. One potential solution to this problem is connecting the AVs to their surroundings by allowing vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication. The vehicle can acquire invaluable information about road geometry, speed limits, and traffic light information from the infrastructure [105]. Likewise, through contact with other vehicles, the positioning, lane-changing intentions, and vehicle states can be exchanged [105].

In 2011, a challenge was held in the Netherlands under the name of the "Grand Cooperative Driving Challenge" (GCDC) in an attempt to accelerate and advocate for cooperative driving through both V2I and V2V communications [103]. A similar challenge called the European Truck Platooning Challenge took place in 2016, where the V2V technology was tested in a larger-scale, real-world environment. In this challenge, the AVs successfully platooned behind human-

driven cars in a course that covered different European cities like Sweden, Germany, Belgium, and the Netherlands [103].

The ongoing developments in autonomous driving prompted the Society of Automotive Engineers (SAE) to define six levels of driving automation [106]. No automation is provided for level 0. Rudimentary automation functions such as adaptive cruise control and anti-lock braking system (ABS) are offered in level 1. The integration of crash avoidance systems and emergency braking can be found in level 2 [107]. Partial automation up to level 2 has been successfully realized and is readily available in the market. However, the real challenge starts with level 3 automation [42]. Conditional automation begins with level 3, where the driver's attention is only called to respond to emergencies, whereas, during normal driving conditions, the human is not required to control the vehicle. Automated driving systems (ADS), in general, have their restrictions, defined as Operational Design Domain (ODD) [93]. For instance, level 3 automation only applies under specific weather conditions and road types, as it also considers the vehicle's hardware and software capabilities. They are expected to work in normal weather conditions on highways [42]. Human supervision is not required for levels 4 and 5 of automation. The only difference is that level 4 is operational in some ODDs since it needs the support of detailed maps and the existence of certain types of infrastructures. If these conditions go unmet, the vehicles automatically park themselves to stop the trip as a part of a fail-safe system [108]. On the other hand, level 5 is designed to have full automation without any human intervention.

By the early 2000s, level 2 driving automation became more prominent and commercially available. Audi's autonomous TTS research vehicle introduced driver assistance functions like adaptive cruise control and side monitoring for safer lane-changing. Equipped with multiple sensors that range from radars, laser scanners, and ultrasonic sensors, this $40 million system, funded by the European Union's HAVEit (Highly Automated Vehicles for Intelligent Transport) project, demonstrated its ability to navigate bends on the road safely by adopting slower speeds [109]. This milestone model ushered a new era towards collision-free autonomous driving.

Recently, Toyota utilized laser and radar technologies to engineer its model so that system failures do not contribute to collisions [110]. On the other hand, Nissan developed the Infiniti Q50 model using cameras and radars for its navigation. Actuators and sensors were employed to assess collision avoidance and cruise control systems [111]. That said, Nissan had plans to produce their

fully autonomous vehicle in 2020 [112]. Audi's Q7 vehicle, with the help of multiple radar and camera sensors, provides traffic jam assistance, utilizing both adaptive cruise control and lane-keeping, and functions at speeds up to 60kph. It also has predictive control features that use standard GPS navigation systems along with information extracted about the road's curvature, roundabouts, and crossings. [113].

In 2018, Audi provided the public with the first level 3 autonomous vehicle [114]. As for Volvo, they initiated their pursuit of autonomous driving in 2006. And in 2017, a significant milestone was achieved with the release of their autonomous vehicle (AV) test model [109]. They outlined their ambitious goal of bringing a fully autonomous model to the commercial market by 2021 [110]. Similarly, Google started its pursuit of full automation in 2009 and managed to complete about three million miles in four states in the U.S. using the WAYMO model [109]. Meanwhile, Tesla planned to equip all their models with advanced self-driving technologies in 2014; presently, all available models in the market possess the Autopilot feature [112]. Nevertheless, Tesla envisions that fully autonomous models will be available by 2025 [112].

Forecasts often disagree on those anticipated launch dates of fully autonomous vehicles disclosed by different automotive companies, seen as somewhat optimistic. Realistically, experts suggest it will be at least a decade before level 5 autonomous vehicles are ready for deployment [115]. Table 1 summarizes the projected dates for achieving AV models capable of level four and five autonomous driving and the infrastructure's preparedness for Connected Autonomous Vehicle (CAV) technology. It is readily evident from the anticipated timelines that level 5 autonomous vehicles are more likely to be available within the next 10 to 20 years. Comprehensively transforming the infrastructure to support CAV technologies, on the other hand, is thought to be realistically available within 20 to 60 years.

Moreover, the expected market penetration suggests that half of the newly produced vehicles will be autonomous by 2045, expanding to half of the entire vehicle fleet by 2060, according to Litman [13]. As for the anticipated benefits projections, reduced driver stress and independent mobility can be attained for wealthy customers when the AVs are limited, and their market price is relatively high. On the other hand, achieving the same benefits for moderate-income customers and having affordable autonomous taxis and micro-transit services can be

fulfilled if AVs are common and affordable. Finally, improving traffic mobility and reducing congestion will only materialize if dedicated lanes are provided to permit platooning.

Table 1. Summary of Forecasts for Realizing Fully Autonomous Driving Technology.

Source	Level 4	Level 5	CAV Environment
Zmud [116]	2021	2025-2030	-
Litman [13]	2020-2030	2020-2040	2060-2080
Bloomberg [114]	2018-2020	2028-2030	2040-2060
Kuhnert [117]	2020-2030	2025-2030	-
Gehrke [114]	2018-2021	2018-2021	2040-2050
Shaheen [118]	2018-2021	2023-2040	2045-2070

Currently, some AV models, indeed, operate on level 4 autonomous driving [115]. However, due to the lack of adequate infrastructure and supporting legislation, their deployment has been restricted to a few small regions with urban environments with a speed limit of only 50 kph [115]. Such regulations prompted these models to be used primarily for ridesharing purposes. WAYMO, NAVYA, and Magna are among the presently available level 4 AVs [115].

Numerous companies, including Audi, Lyft, Uber, WAYMO, Tesla, Apple, Renault, and Ford, have openly acknowledged their ongoing efforts to test level 5 autonomous vehicles for future public use. However, no level 5 autonomous vehicle has yet been released for commercial use [115]. Despite this, these companies have successfully identified the technology and technical specifications required to attain level 5 autonomous driving [115].

Several challenges impede the full realization and deployment of fully autonomous vehicles, encompassing technological, safety, ethical, and political aspects [112]. A significant technological hurdle arises from the immense data influx into the AV's onboard computer, primarily from sensors like LiDAR, which complicates real time data processing and subsequently impacts the vehicle's efficiency and safety. The concept of vehicle-to-anything communication (V2X) holds promise as a solution to this challenge. However, the current infrastructure cannot support this technology yet, and existing networks lack the robustness needed to support the anticipated high volume of data exchange [112]. Additionally, safeguarding the privacy of the extensive data collected by AVs is a vital concern [119]. The need to ensure the security of

transmitted information, with cyberattacks posing risks to sensing, communication, data, and control security, is vital [120].

Another aspect that poses a hurdle to commercializing autonomous vehicles is their environmental perception. Avs' development level is determined primarily by the robustness and reliability of their sensing technologies [121]. And since the adopted sensors are still not fully developed to maintain the required reliability levels, any trivial mistakes can be catastrophic [112]. An example of the immaturity of equipped systems is the Tesla-related fatality, where the Autopilot system failed to identify a white truck, confusing it with the ambient clouds, thus killing the driver [122]. Furthermore, the correct and timely response to unforeseen loss-of-control incidents like skidding is not on par with human reactions [123].

Unmanned autonomous vehicles face a significant challenge in matching or improving upon human factors, such as ethical decision-making on the road. While these vehicles excel at following traffic rules and safe navigation, they lack the "Human Touch" needed to make moral decisions in complex situations that involve human emotions, morals, and judgment [112]. This lack raises concerns about potential biases in the algorithms or AI used in AVs [112]. It is a concern because it is believed that the public will only switch from manned to unmanned once the ethical principles guiding AVs are clear [54].

A different facet that impedes the progress of AVs is the current policies and regulations. Liability, in particular, emerges as a pivotal concern for the widespread adoption of AVs. Currently, drivers are typically liable for any car-involved collisions [124]. However, determining the primary responsible party becomes rather unclear in the context of accidents involving unmanned vehicles. In such cases, multiple entities come into play, including the driver, car manufacturers, software developers, and AV developers, further complicating the liability landscape [125].

2.1.1 Summary

The advent of AVs carries many promises, ranging from significantly enhanced road safety and improved traffic efficiency to fostering inclusivity for the elderly and individuals with disabilities. Additionally, their positive impact on environmental conservation cannot be overstated. In essence, the evolution of AVs not only represents a technological milestone but also holds the

potential to revolutionize societies, making transportation safer, more efficient, and accessible while contributing positively to the environment and economy. Understanding the challenges and solutions in this transformative journey is crucial to harnessing the full benefits of this groundbreaking technology.

This section offers a comprehensive exploration of the evolution of AVs, tracing their development from inception to their cutting-edge present state. Delving into the historical context, it highlights the key milestones and innovations that have shaped the trajectory of AV technology. Additionally, the text addresses the contemporary challenges that have hindered AVs' progress and widespread deployment. It delves into these obstacles, shedding light on complex issues ranging from technological hurdles to ethical considerations and legal frameworks. Crucially, this section highlights the challenges and examines the innovative solutions proposed in existing literature. Investigating these proposed solutions provides valuable insights into the strategies and technologies that hold the potential to overcome the obstacles, thereby paving the way for the full realization of AVs.

2.2 AUTONOMOUS VEHICLE SIMULATIONS ON REALISTIC LIDAR DATA

Achieving a meticulous replication of the intricate physical road infrastructure and an exacting simulation of the physics involved in sensing processes is critical to bridging the substantial divide between theoretical simulations and real-world applications [85]. Closing this gap not only refines the development of AV technologies but also augments their reliability, safety, and efficiency.

Before LiDARs, laser finders and camera images were used to simulate a virtual environment where autonomous land vehicles controlled by a sophisticated 3-layer back-propagation neural network were tested for their ability to execute road-following tasks [126]. The results demonstrated the vehicle’s ability to effectively navigate real roads under specific field conditions.

Neuhaus et al. [127] utilized the 3D point cloud data captured by Velodyne HDL-64E LiDAR sensor in assessing autonomous navigations in unstructured environments. Drivable areas are analyzed using an innovative algorithm that evaluates local terrain roughness. This data is valuable for path planning algorithms, enabling them to perform non-binary decisions such as

choosing between different terrains, providing a more detailed and context-aware understanding of the terrain, and enhancing the precision of autonomous path planning.

In attempting to close the gap between AV simulations and the real world, Fang et al. [90] managed to develop an innovative framework for generating highly realistic 3D point data annotated for training Deep Neural Network (DNN) models by utilizing mobile LIDAR scanners, typically employed in land surveys, to directly capture detailed 3D scans of road scenes, eliminating the need for constructing artificial environmental models. The team developed data-driven methods to accurately determine obstacle poses and shapes to enhance the simulation's authenticity and reduce costs. Actual traffic scene data was utilized to synthesize obstacle placements within the virtual environment. Additionally, a novel LIDAR renderer, incorporating physical models and real hardware statistics, was developed, resulting in simulated LIDAR point clouds virtually indistinguishable from real-world data. This approach significantly narrowed the performance gap between simulated and real LIDAR data to a remarkable 1~6% across diverse applications, achieved without fine-tuning or blending real and simulated data. Furthermore, incorporating a mere 10%~20% additional real data for model fine-tuning led to substantial improvements in the model's performance.

A similar work that realized the importance of leveraging real-world data in the simulations was conducted by Manivasagam et al. [85]. Through their work, the researchers advocate for leveraging real data collected by their self-driving fleet in diverse cities to enhance simulation realism. They curated an extensive catalogue of 3D static maps and dynamic objects from real-world situations. Using this catalogue, they devised an innovative simulator that integrates physics-based and learning-based approaches. This simulator employs ray casting across 3D scenes and a deep neural network to introduce deviations from physics-based simulations, resulting in genuine LiDAR point clouds. The resulting tool, LiDARsim, proved invaluable for testing perception algorithms in rare scenarios and evaluating safety-critical situations through comprehensive end-to-end closed-loop assessments. Significantly, this framework differs from that conducted in [85] in adopting a standard LiDAR sensor for their data acquisition, which allowed their mapping to be more cost-effective and highly scalable.

Li et al. [128] proposed augmented autonomous driving simulation (AADS) by combining LiDAR and cameras to scan real street scenes. Then, they introduced foreground vehicles and

pedestrians to their 3D virtual environment using CG models. However, the overall simulated traffic flows were based on real trajectories collected from the scanned data. The annotation of their data was performed by a third party to facilitate their scan-to-simulation approach. In contrast to traditional approaches, their method offers more scalability and realism. Likewise, Fang et al. utilized MLS data to create a virtual environment that can directly reflect the complexity and diversity of real-world geometry [129]. Then, by applying CAD models to capture the obstacles' poses, such information was incorporated into the virtual environment to enrich and enhance the AV simulations. This method demonstrated that a combination of real and simulated data can attain over 95% accuracy in the simulations.

In contrast, several other works have used synthetic LiDAR data in their AV simulations since they do not require the same heavy manual annotation work as the scanned LiDAR data. Hence, they promise to streamline and increase the efficiency of AV simulations [87]. For instance, Yue et al. extracted their synthetic LiDAR annotated datasets from the famous computer game (GTA-V), simulating a virtual scanning vehicle within the game's environment to capture realistic driving scenes [87]. Their proposed method improved the validation accuracy for point cloud segmentation using a neural network model by 9% on the real-world benchmark.

Wang et al. [130], motivated by the recent success of deep learning in 3D data analysis that relies on large volumes of annotated data, developed a framework that simulated LiDAR sensors in the CARLA [77] autonomous urban driving simulator to generate synthetic LiDAR data. Their approach was inspired by the notable achievements of deep learning in 3D data analysis, which often relies on extensive volumes of annotated data. This method demonstrated that incorporating synthetic data significantly improves the performance and accuracy of AVs in simulation environments.

2.2.1 Summary

In summary, the literature shows the benefits of utilizing real-world LiDAR point clouds in the simulations because they can bridge the gap between sim-to-real. Open-source simulators, such as those presented in [77] and [131], rely on handmade 3D features and implement inadequate physics assumptions, resulting in ill-conceived representations of real-world sensory data. Real time graphics engines, such as Unreal [132] and Unity [133], also exhibit a sizeable sim-to-real gap

mainly because the generated environment lacks diversity and does not accommodate important sensor simulation features. The frameworks discussed were actively used to enhance the perception modules of the AVs in an attempt to improve their overall safety. Similarly, the framework proposed here seeks to achieve the same goal by quantifying environment complexity for AVs based on real-world data to bolster the accuracy and validity of the results. It is also worth noting that the use of synthetic LiDAR data in simulations is yet to be on par with using realistic data since it was shown to have diminished accuracy [90].

2.3 Applications of LiDAR Infrastructure Digitization in Transportation Engineering

LiDAR is an acronym for light detection and ranging. As an optical remote sensing technique, it relies on emitting thousands of laser beams from its laser scanners to sense its surroundings by collecting the information from the reflected beams, carrying valuable information about the position of the hit targets. The continuous scanning operations result in highly detailed 3D point clouds replicating the surrounding environment digitally. LiDAR is also equipped with navigation sensors like GNSS receivers and IMUs to help with the system localization and correctly calculate its geo-referenced position [134]. GNSS provides position and timing information for LiDAR for it to locate itself [135], whereas IMU sensors can measure both linear accelerations and spin rates, which can be used to identify the orientation and 3D position of objects [136].

The accuracy and density of LiDAR data are highly dependent on the collection method. The data can be collected aerially using drones or helicopters or terrestrially, either statically by mounting the LiDAR on a tripod or making it mobile by mounting the system on a moving platform like a car or truck. Mobile laser scanning (MLS) has been widely recognized to be the more expedient form of LiDAR data as it is known for the high level of detail for its outputs, cheap cost of operation, high point density, and the ability to scan below bridges and inside tunnels [83]. Other factors can also influence the density of the point cloud data, such as the sensor's range and the incidence angle, known as density anisotropy [137]. The weather and environmental conditions can also impact the point density and the speed by which the data was collected [138].

The appeal of using LiDAR technologies in the transportation field has increased over the past decade owing to the meticulousness of its outputs and the high degree of convenience of their

collection methods [139]. LiDAR is an appealing alternative to traditional surveying tools, which are slow, often inaccurate, and disrupt traffic operations. Since LiDAR performs its scanning while complying with highway speed limits, outputs are faster or more time-efficient and do not cause any tangible disruption to traffic operations. Further benefits include their 360-degree field of vision, ensuring minimal loss of detail regardless of the device's orientation and active sensing, giving them the flexibility to operate under different lighting conditions [139].

The prospective utilization of LiDAR data in different road inventory applications has long been discussed in the literature [140]. Williams et al. [83] and Landa et al. [141] have demonstrated the potential of LiDAR data in different road inventory applications, which can be classified into the extraction of road pavement, lane marking, road edge, alignment information, and vertical and lateral clearances [139]. Even though extracting semantic information from LiDAR data can be cumbersome, researchers have recognized the value of the extracted information, leading to significant dedication to extensive research in this area.

The isolation of road surface points from the rest of the point cloud data is recognized as an essential step for two reasons. One is the provision of a reference point for extracting other features and minimizing the data size for further processing operations [139]. Lam et al. [142] used Kalman filtering to capture the segmented planes of the road, and then mathematical models were fitted to the extracted planes through Random Sample consensus (RANSAC). It was reported that the results were satisfactory. More recently, the RANSAC algorithm was utilized by Hu et al. to extract the road surface plane from ground points [143]. After that, a point density threshold was set to facilitate the extraction of road surface points. On the other hand, Wang et al. [144] first defined the profile of the road and then captured its edges based on the height difference of specific road segments. The results were not clearly demonstrated for this study.

More recently, Yadav et al. [145] tested a method of splitting the point cloud into ground and non-ground points in India with a correctness of 98.3%. Even more so, they have managed to identify the road surface and non-road surface points from the extracted ground points. This was done in three stages. First, they applied data structuring and ground filtering techniques where a 2D grid representation of points was implemented to get a primary set of ground points. Then, the road surface extraction was carried out using the intensity and density of points to filter out ground

points, followed by region growing and road point recovery. Finally, they concluded with road boundary refinement.

The LiDAR scanlines were analyzed by Zhou et al. [148] to differentiate between the ground and non-ground points by tracing the scanlines of different points. On a similar note, an overlapping cylinders representation of points was introduced by Yadav et al. [146], where the ground points were identified at the bottom base of the developed cylinders. The identified points were then subjected to a refinement process by employing a height roughness-based approach. Other researchers pursued voxel-based approaches [147] [148] [149]. For example, Yu et al. [150] proposed a method that segmented the point cloud into a 2D XY grid, and then an upward-growing voxel approach was applied to connect voxels with their upward neighbours. The ground elevation was compared to the constructed voxels to identify the ground and non-ground points based on a defined threshold.

The extraction of lane markings and road edges has been a point of interest for researchers, particularly those invested in developing AVs. Typically, AVs rely on detecting the road's edge and lane marking for guidance and navigation. Generally, in urban settings, curbs facilitate the extraction of road edges, unlike in rural settings where curbs are often absent. Additionally, such information is in high demand in other geospatial applications such as road safety evaluation, traffic accident analysis, and intelligent transportation systems [151].

One of the proposed methods by researchers for curb extraction was presented by Zhou and Deng [152], and it involved the identification of sudden height differences where the maximum difference was calculated to locate the curbs of the road. This method was tested on point clouds obtained from aerial laser scanning (ALS) and MLS, where the completeness figures varied between 53% and 92% for ALS. As for the MLS data, the figures ranged between 54% to 83%.

A real time road edge and surface extraction procedure was presented by Zhang [153], where elevation filters were applied along with pattern recognition techniques. A false positive rate of 0.83% was reported in detecting most road-curb and road points. Nevertheless, the details of the testing environment were not clearly demonstrated, whereas Serna and Marcotegui [154] presented a framework that utilized a range of images for curb extraction. The λ-flat zones algorithm was employed for ground and non-ground segmentation. They relied on the height and

elongation attributes in the identification of curbs. The data sets used in this study were collected from France and the Netherlands where the completeness results ranged from 54% to 65%, while a 91% to 95% range was seen in terms of correctness.

An average precision of 84.49% and recall rate of 82.87% for curb detection was reported by [155]. The authors opted to apply a plane-based method for on-road and off-road point segmentation. The algorithm was tested on LiDAR data collected by an AV, where the curbs were identified based on their spatial properties. Jaakkola [156], on the other hand, proposed a framework that used image processing tools and heuristic thresholds on a TIN surface generated from LiDAR point clouds to extract pavement markings and curbstones. Different pavement marking types were extracted using this method, such as pedestrian crossings, and the authors reported around an 80% detection rate. However, Kumar et al. [162] sought to extract the road's edges by employing point cloud rasterization techniques, Gradient Vector Flow (GVF), and Ballon Parametric Active Contour Models. The authors tested this method on 50m road sections, where they reported that the failure in road edge detection was attributed to low point density near the edges.

Range-dependent thresholding was adopted by [157]-[158]. Intensity-based thresholding coupled with binary morphological operations was used to extract lane marking information [157]. The authors reported that 86% of markings were identified accurately. However, Guan et al. [158] carried out the road edge identification by the height difference analysis of the extracted blocks of data from the point clouds as a final step. Segmentation filters with multiple thresholds were applied for lane marking extractions. The authors tested their framework in two road segments where 96% completeness and 83% correctness values were reported.

More recently, deep learning and spatial statistics techniques were implemented by [159] in detecting road edges. The method used the pre-trained RandLA-net for the semantic segmentation of the point cloud data and then applied voxelization and spatial statistical analysis to detect the road edges. Additionally, the authors reported the capability of their proposed method in identifying the locations of objects that belong to the car class to overcome the occlusion problem of parked vehicles. Meanwhile, [160] utilized both the spatial and geometric information of the input point cloud data to extract curbs in complex scenes to aid city managers in monitoring

the status of the roads and urban street reconstruction. The authors reported values north of 98% in both completeness and correctness measures.

Mi et al. [151] proposed a framework that relied on generating super voxels to extract potential curbs with high efficiency. Supervoxels are a series of spatially adjacent voxels that share common properties [161]. The methodology entailed grouping the prospective super voxel curbs and producing a continuous boundary for the road. After that, the extracted road boundaries were fitted, tracked, and completed to obtain geometric parameters such as boundary location, roadway widths, turning radius, and slopes. The authors reported the high success of this method through precision and recall scores, which amounted to 95% and 91%, respectively.

Traffic sign extraction is another useful application for the digitization of infrastructure. Owing to the valuable information that the signs carry for drivers, governments have been actively seeking ways to streamline the monitoring and maintenance process for these signs [162]. Hence, the literature reveals several research efforts to optimize the extraction of traffic signs using LiDAR data. For instance, Vu et al. [139] proposed a method to identify and classify traffic signs in real time using intensity filtration techniques and Principle Component Analysis (PCA). The authors applied their method in a test track and reported high success rates that ranged from 84% to 96%.

Weng et al. [139] proposed a framework using intensity filtering techniques, hit count, and elevation information. By leveraging these methods, they obtained the geometric shapes and 3D shape context and then used them for traffic sign identification and classification. Nevertheless, not much information is known about the success rate of this method. Earlier, Ai and Tsai [163] explored the success of implementing a framework that again looked to filter their point cloud by applying intensity filtering techniques and hit counts but coupled them with MUTCD offset values in detecting and classifying traffic signs. The authors reported a 94% detection rate in urban highways compared to 91.4% in rural areas. The false positives were linked to several factors such as low retro-reflectivity, insufficient height, and occlusions.

In Alberta, Canada, Gargoum et al. [164] attempted to both detect the traffic signs and map them using an assortment of different techniques, including intensity filtering, density-based clustering (DBSCAN), geometric filtering, used to clear clusters outside a defined threshold and buffer zone filtering, implemented to remove any clusters located in unreasonable locations for

road edges. The authors applied their method on three highways and reported staggering detection rates ranging from 94% to 100%. On a different note, the premise of Guan's work [165] in detecting traffic signs stood on the previous knowledge of sign pole heights and roadway widths from LiDAR data. First, the point cloud was voxelated, and then the geometric properties of adjacent voxels were obtained to identify the potential sign candidates. Finally, different intensity, size, and position filtering were applied to the elected voxel clusters to extract the signs. The noise was addressed by applying Euclidean distance clustering algorithms. An 86.8% detection rate was observed in this study.

A different approach was adopted by Javanmardi [166], where traffic signs and light poles were detected through the implementation of surface reconstruction algorithms coupled with k-means clustering. In the proposed method, objects with high elevations were captured by introducing a sliding cuboid, and a RANSAC algorithm was employed to filter out false candidates. Success rates of 94.5% were reported by the authors in detecting traffic signs. More recently, Gouda [162] explored the efficiency of adopting machine learning techniques to extract traffic signs. The proposed method was to improve the performance of the PointNet++ neural network in outdoor scenes. Using data collected from seven highway segments in Alberta, Canada, the author reported high success rates with a 99.2% recall and a 98% F1 score, demonstrating the adopted framework's efficiency and accuracy.

In similar works, a fully automated algorithm was developed by Gouda et al. [167] to detect the positions of the light poles on the road owing to their increased safety concerns if not placed correctly in the designated roadside clear zones. Using MLS data of 28 km of highways in Alberta, Canada, the proposed algorithm scored north of 98% in precision, recall, and F1 metrics, outlining its effectiveness.

One more application for the MLS data in the field of transportation is the extraction of geometric features of the road, such as slopes, horizontal and vertical curves, sight distances, and clearances. There are ample guidelines that regulate the cross-sectional design of the road. For instance, the superelevation value on horizontal curves and the curve's radius are governing design factors that ensure the stability of vehicles. The minimum length of the vertical curves is also essential in highway design and for its safe operations. LiDAR data present a valuable and viable

alternative to conventional data collection methods for their time efficiency, minimal disruption of traffic, and high accuracy [139].

The cross-sectional information has been captured by [168] through the modification of the orientation of the laser scanner. A region of interest (ROI) was identified perpendicular to the road, and the points within this region were exposed to linear regressions to estimate the cross slopes. The study was conducted in a controlled environment and demonstrated results within 0.28% of the digital readings. In a different study, Holgado-Barco proposed a semi-automatic algorithm to extract slope information from LiDAR data [169]. First, the segmentation of the ground surface was achieved using height thresholds. They applied intensity-based filters to extract lane markings, which were used to calculate the slope information by considering the lane width and the difference in elevation between their ends. The authors reported minor variations in their estimations when the algorithm was tested on two highway segments.

More recently, A fully automated algorithm to estimate the cross and side slopes was presented by Gargoum [170]. First, points that juxtaposed the road's axis were extracted. Then, Multivariate Adaptive Regression Splines and linear regressions were applied to identify the points that had a change in slope. The authors applied their method on two highway segments in Alberta, Canada, and a difference of 0.08% to 0.22% was reported in the cross-slope calculations.

One of the earliest studies that attempted to extract vertical alignment information from LiDAR data was carried out by Souleyrette [171]. ArcGIS was used to localize the road's centerline, and then multiple linear regressions were applied to estimate the grade information. The findings of this study showed a significant deviation in their calculations when compared with field measurements. Adopting a similar method, Zhang [172] attempted to refine the detection of the road edges and centerlines. That, in turn, yielded a 5% accuracy in the grade estimation.

A later study by Higuera de Frutos and Castro [173] involved GNSS data collected along a road's centre line to reconstruct its vertical profiles. The authors attempted to automate acquiring different elements of a road's vertical profile information using the collected centre line points by obtaining grade points, parabolic curve points, and border points from classifying the points along the profile. A mean error of 8cm was reported in the study when the algorithm was tested on a rural highway in Spain. Earlier studies have also attempted to use the same principle of exploiting

the GNSS data in extracting vertical profile information [174]. It is important to note that GNSS data are one of the outputs of MLS systems, which is the main focus of this literature review.

Similar efforts have been directed toward the extraction of horizontal alignment information. Holgado et al. [175] proposed a semi-automatic method for curve extractions. The point cloud data was subjected to segmenting, parametrizing, and filtering. The lane markings were distinguished in the segmentation step, and then the data was classified into straight segments, circular arcs, and clothoids (which exhibit a variation in curvature). After that, the radii and transition lengths were estimated. On data collected from a Spanish highway, the algorithm demonstrated a 2% error in the length of circular arcs and 0.4% in the circular radius compared to topography calculations.

The attributes of the horizontal curves had an average difference of less than 3% when tested on different LiDAR segments in a study by Gargoum [176]. The authors captured the endpoints on the curve by analyzing the change in azimuth along the trajectory of the road. The point of intersection and deflection angles of the curve were identified using Linear regression models. Then, the point of curvature (PC) and point of tangency (PT) were identified, enabling the authors to calculate the radius of the curves. A recent study by Shalkamy [177] proposed a framework for detecting horizontal curve elements on a network level using LiDAR data. This large-scale, fully automated algorithm followed the principles outlined in [176]. It achieved high accuracy attainment levels ranging from 96% to 100% when tested on 242 km of highways in Alberta, Canada.

Sight distance assessment studies using LiDAR have seen a growing interest from researchers owing to the time and effort hurdles when traditional methods are used. In a study by Castro et al. [178], the sight distance requirements were assessed using ArcGIS, where digital terrain models (DTM) with viewsheds for observers were developed. The visible parts of the viewshed were converted to polygons, and the intersection between these polygons and the trajectory of vehicles was identified. Then, the distance between the intersections and the observers was measured to quantify the available sight distance. The results were compared to the Trivium software, and it was found that there is no statistical significance in the differences between the two results. Later on, the same authors [179] attempted to improve the automation level of the method by overlaying the observer and target points onto the DTM models directly from GPS

surveys. The findings were compared to their earlier results. However, differences were found to be not statistically significant.

Finally, assessments of vertical clearances for bridges and tunnels are also possible using LiDAR data. Typically, this work is done manually from site inspections using tools like theodolites and total stations. However, this process can benefit greatly from the advantages of utilizing LiDAR data, which promises to streamline this process, cause no disruption to traffic operations, and eliminate human errors [139].

Puente et al. [184] implemented a semi-automated framework to assess the vertical clearance in tunnels in Spain. The process begins by capturing cross-sections of the tunnel at regular intervals. Subsequently, an algorithm was employed to detect lane markings, pinpointing the boundaries of the travel lanes within the tunnel. Convex hulls are applied to assess clearance, creating a semicircular representation of the tunnel's cross-section and achieved by connecting points along the road edges to corresponding points on the tunnel roof. The effectiveness of these algorithms was rigorously evaluated across multiple cross-sections within the tunnel. Remarkably, the results demonstrated a negligible error rate, with the variance between the clearance estimated through this innovative method and the actual measurements taken in the field not surpassing 1% for most cross sections.

With emphasis on vertical overhead objects' vertical clearance assessment in Alberta, Canada, Gargoum et al. [180] proposed a framework that involved delineating the road's path and eliminating non-ground points from the highway data. A Nearest Neighbour search method was employed to identify objects above the road surface. Subsequently, these overhead objects were organized into distinct groups using a density-based clustering algorithm. A statistical kurtosis assessment was applied to categorize these objects, distinguishing between bridges and non-bridges. Additionally, the algorithm calculated the minimum clearance for each identified object. The effectiveness of this methodology was verified through testing on three diverse highways, one of which was a lengthy 242 km highway corridor.

2.3.1 Summary

In recent years, LiDAR point clouds have become a focal point in transportation, evident in their widespread applications across various tasks within the field. Infrastructure digitization through

LiDAR technology has ushered in a revolutionary era, offering automated and semi-automated methods that surpass traditional techniques. These advanced methods ensure safety and significantly enhance accuracy, all without disrupting regular operations.

This section is a valuable resource, shedding light on cutting-edge techniques employed in current research to maximize the potential of LiDAR data. The literature resoundingly attests to the effectiveness of LiDAR-based approaches in diverse applications. They excel in on-road tasks such as extracting road surfaces, lane markings, and road edges. Similarly, they prove adept at capturing crucial roadside details, including traffic signs, vegetation, and pole-like objects. Moreover, LiDAR technology has demonstrated exceptional capabilities in assessing vital road geometric features, such as vertical and horizontal alignments, sight distances, and vertical and lateral clearances.

Finally, the advent of LiDAR-based algorithms marks a significant leap forward in transportation. These advancements are poised to fundamentally transform transportation infrastructure, creating safer, more reliable road networks and facilitating smoother traffic mobility.

2.4 QUANTIFYING THE COMPLEXITY OF THE AV ENVIRONMENT

The onboard computer equipped by the AVs, which is required to operate in real time, has the onus of perceiving the environment surrounding the AV, processing the incoming information, and making the most apt decisions to ensure the safe operations of the ego vehicle. The traffic environment poses a big challenge for autonomous systems as they are typically open, non-deterministic, hard to predict, and dynamic [181]. Identifying the complex situations is essential in advancing AVs' safety, which would expedite their mass adoption.

Wang et al. [182] proposed a modelling and assessment framework that can quantify the complexity of the AV's environment. Their approach involved establishing fundamental and additional environment complexity models systematically evaluating four key environmental aspects: road conditions, traffic features, weather conditions, and potential interferences. The overall environment complexity can be calculated using a preset scoring system for the different environment features based on experts' judgment. The Analytic Hierarchy Method (AHM) determined the relation between different attributes. A weighting scheme based on subjective and

objective considerations was implemented to calculate the overall complexity of the environment. Taking both subjective and objective considerations in the analysis, two tests involving lane changing and crossing a simple intersection were conducted to validate this method and show its effectiveness in estimating the complexity of the environment.

A similar, automated framework that bases its measured complexity on the road type (RT), scene type (ST), challenging conditions (CC), and traffic elements was developed by Wang et al. [183]. Traffic elements focus exclusively on vehicles, considering a maximum of the closest eight neighbouring vehicles. The calculation criteria centred on performance for the teams participating in the Intelligent Vehicle Future Challenge (IVFC). The main tasks considered were lane keeping, detection of vehicles, pedestrians, and signals. Using both LiDAR point clouds and image data, a Support Vector Regression (SVR) was implemented to train a machine-learning algorithm to predict the complexity of the scenes. With an accuracy of 93.23% and 68.7% for the training and test data, respectively, the proposed framework was validated using three experiments that modelled different road and traffic conditions.

Gravity models were proposed by Zhang [184] to assess the complexity of the surrounding environment, where the level of driving complexity was measured as the extra cognitive burdens exerted by the traffic environment on the drivers. That said, the proposed method could not directly obtain the complexity values, and many relevant parameters were miscalibrated in the calculations. Following the same concept, Yang et al. [185] divided the environment into static and dynamic elements to develop their environment complexity model. The static features were studied using the grey relation analysis. At the same time, the complexity of the dynamic elements was quantified based on the improved gravitation model, adding an extra explanatory variable into the function to explain the degree of contribution of the driving strategy. The proposed method was validated on a car-following driving scenario based on the 2018 World Intelligent Driving Challenge. The proposed static and dynamic models were compared to the existing models and were more convenient and efficient.

Focusing on the dynamic traffic elements only, [186] proposed a framework that captured the objective human drivers' judgment on the complexity of the driving environment. In this method, the complexity of the environment was defined based on the interactions of the ego vehicle with the vehicles surrounding it. The proposed model can be established by first designating the

influence area around the ego vehicle where the lateral distance is taken as a specific number of lanes, while the longitudinal distance considers the event of rear-end collisions by adopting the theory of Responsibility-Sensitive Safety (RSS). Second, using information entropy theory, the vehicle pair complexity is studied by considering the encounter angle, relative velocity, and relative distance between vehicles. Lastly, the vehicle pair complexities are aggregated. Data collected from multiple sensors, including radar and video cameras, were used to validate the efficacy of this method. On applying this framework to three case studies that involved lane changing, car following, cutting in, and cutting out, the results showed that the produced complexity curves managed to quantify and time the changes in environmental complexity. One drawback to this method is its inability to describe the static environment complexity.

Multiple researchers utilized the potential field theory in their models. A highway potential field function was proposed by Wolf et al. [187] to aid the AV's obstacle avoidance system. This function constructed the potential field as a wedge while considering the velocity potential of other vehicles. A novel concept of driving risk-field was presented by Wang et al. [188], where three different fields were established for the moving objects, road environment, and driver characteristics. The characteristics of traffic agents and traffic rules were considered in developing a potential field model proposed by Tu et al. [189] to aid in the lateral planning module of AVs in urban environments. Likewise, a model predictive path planning controller, based on the potential field concept, was developed by [190] for obstacle avoidance. Based on Lennard-Jones potential, the safety risk of CAV platooning and vehicle following were considered in the unified vehicle function developed by Jia et al. [199].

Similarly, Cheng et al. [191] based their environmental complexity evaluation model on the potential field theory. The environment elements are represented by a positive point charge or uniformly charged wires that create a potential field in their vicinity. The total potential field of a particular environment can be calculated by superimposing individual fields. The quantified complexity is determined by the impact of the defined static and dynamic environmental elements on the AV based on the viewpoint scene. The virtual electric quantity of the different environment elements is calibrated using the AHM, where non-motorized vehicles have the highest values and static traffic elements like lane markings have the lowest values. This method was verified on virtual traffic environments using the PreScan software and on real traffic scenarios using driving

videos of the Huawei AV. The quantified complexity was found to be consistent with the complexity of the given scenes. Finally, when comparing this method to expert scoring, the error percentage was within $\pm$ 5%, again demonstrating the efficiency of the proposed model.

2.4.1 Summary

In summary, multiple attempts have been dedicated to quantifying the complexity of the environment surrounding the AV. Generally, there are objective and subjective quantification methods. Some methods relied on implementing experts' subjective evaluations and AHP's [182]; others sought to base their quantifications on road semantic and traffic element complexities to model static and dynamic elements [183]. Gravity models were proposed to develop quantitative models[184] [185]. Other researchers based their calculations on the spatial-temporal interactions between different vehicles and implemented the concept of information entropy theory to explain the non-linear relationship between them [186]. Finally, the most commonly implemented models were based on the potential field theory and were adopted by researchers in [191] [192].

Nevertheless, the existing literature's evaluation of the surrounding environment's complexity has limitations. Using predefined subjective ratings for various road and weather conditions, relying on expert opinions is simplistic and lacks comprehensiveness, failing to capture an accurate perception of AVs. Moreover, the literature predominantly focuses on camera and video data, overlooking the potential of LiDAR point clouds in assessing complexity. Additionally, existing models calculate complexity in regions of interest around the AV without accounting for the specific sensor specifications. Consequently, the proposed framework aims to offer a more insightful understanding of the static environment's complexity. It achieves this by accurately simulating the actual perceived environment using commercially available LiDAR sensors through digital twin environments, addressing the aforementioned limitations. Furthermore, the complexity of the environment is captured through the estimation of the required processing data rates, which also anticipates the AVs' level of performance in different driving environments and weather conditions.

3 METHODOLOGY

This research examines the interaction between the AV and the surrounding static environment. It aims to study and quantify the impact of relevant road features on the computational requirements of AVs. Specifically, the study calculates the expected data rate that the AV needs to process in real time to ensure safe operations under ideal weather and driving conditions.

The analysis will be conducted in two ways: first, it will encompass the entire scanned environment, including all road and roadside features. Second, it will focus solely on the right-of-way section of the road, containing pertinent road features. This dual approach aims to effectively gauge the influence of the road's geometry (including vertical and horizontal curves and roadway width) on the AV's performance. The contrasting data rate requirements between the comprehensive scene and the relevant road features will also be highlighted.

The implemented methodology adopts an end-to-end system approach, wherein the operational modules are not treated in isolation but as interconnected. The calculated required data rates are directly linked to the data collected by the AV sensors, meaning that the data rate values encompass all the necessary operations, starting from perception and extending through the decision-making and planning stages. The more points collected at an individual frame, the higher the required processing power for safe AV operations, or in other words, this location would be more challenging to navigate by the ego vehicle. Inversely, the frames with low data rate requirements would mean less processing power would be required and, by extension, easier to navigate. That said, locations that exhibit low data rate values can also indicate a loss of information for the AV, which might be critical for its operations. Hence, more emphasis will be directed towards the frames that exhibit the highest and lowest values at any given road section.

An overview of the implemented methodology will be presented in this section. The general high-level framework is illustrated in Figure 1.

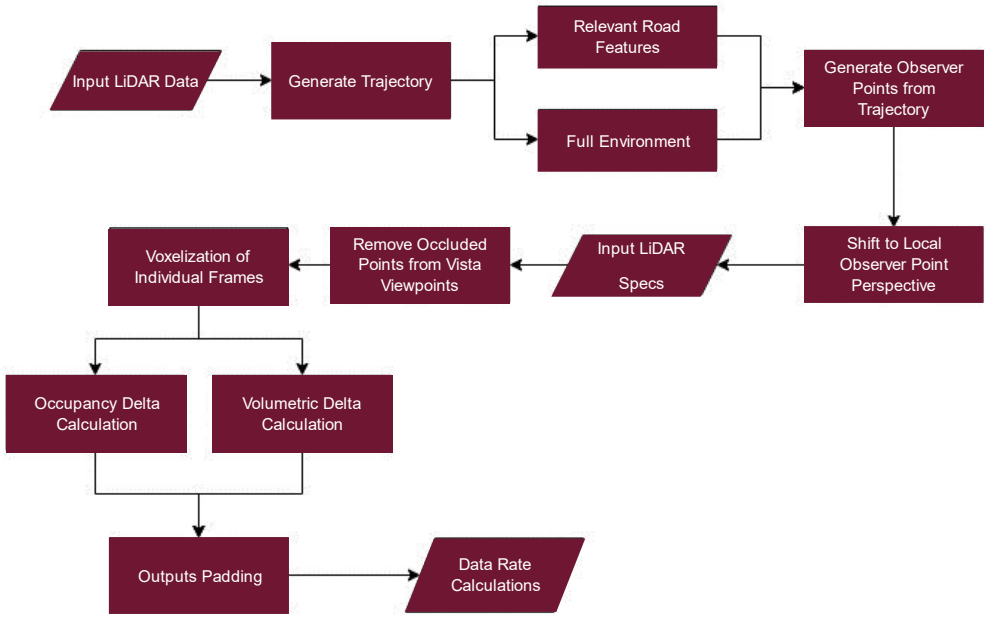

Figure 1. High-Level Framework for Methodology

3.1 Point Cloud Data

The primary input of this method is the 3D LiDAR point cloud data generated from mobile laser scanners (MLS). This data is known for its points' density and accuracy, making them the most suitable data for this type of simulation. Aerial laser scanner (ALS) data can also be used. However, their low point densities, especially at underpasses, below bridges, and underneath canopies, make them less convenient.

The data used in this research was collected by Alberta Transportation using Tetra Tech PSP-7000, a proprietary multifunction pavement surface profiling vehicle. This vehicle is equipped with a REIGL VMX-450 system [193] that can collect 360° LiDAR data using 2 RIEGL VMX-450 laser scanners, producing up to 1.1 million measurements/sec and has up to 400 lines/sec scanning rate. This system has two GNSS antennae and an IMU system to provide accurate georeferencing information. Additionally, the system can be supplemented with a camera system that provides images that can be integrated with the LiDAR data to get RGB information about the surroundings. This system is mounted on top of the vehicle's roof rack to maximize the field of view for the sensors.

The density of the point cloud data can be influenced by multiple factors like LiDAR sensor specifications, where the precision and rotational speeds can impact the density of the points. Furthermore, the speed of the moving vehicle where the LiDAR sensor was mounted during scanning also dictates the point cloud densities, where lower speeds would result in higher point densities. For that reason, the scanning was conducted at normal traffic flow speeds of up to 100 kph.

The National Cooperative Highway Research Program (NCHRP) [194] has put forth some guidelines that define the level of granularity for the point cloud density and the accuracy needed for different LiDAR point cloud applications. It has indicated that for purposes that involve autonomous navigations, the minimum recommended point cloud density is between 30-100 pts/m^2. It was also suggested that any value between 0.05 to 0.2m would be accepted for the accuracy of the scanning sensor,

The pavement was sliced into tiles of known areas to investigate the input point cloud density. An area of 20x10 m^2 was used in the calculations, and then the number of points was recorded to estimate the number of points per square meter for each sliced area. The used data averaged values north of 150 pts/m^2, which exceeds the recommended values by a margin. The RIEGL sensor operates with an accuracy of 8 mm [193], marginally satisfying the accuracy requirements and validating the data quality used in this research.

3.2 PREPROCESSING

The point cloud data was manually checked and verified to ascertain their integrity by comparing them to the real environment and ensuring no discernable loss of road information. Any excess road points are filtered to prepare the point clouds for later calculations. Figure 2 illustrates the validation of a road section with its point cloud counterpart to ensure that the sudden gap in the right-side vegetation was natural and that the virtual environment accurately reflects the actual scene.

Figure 2. Validation of Surrounding Environment

3.2.1 Point Cloud Filtration

Point cloud filtration addresses any issues the sensors might have encountered during their scanning operations, like malfunctions or complete occlusion due to moving vehicles. Scanned points not associated with the road or the environment, like rain, snow, or dust particles, were manually removed from the point clouds. Additionally, since this framework focuses on the interaction between the ego vehicle and the statical environment only, the influence of moving targets like vehicles, pedestrians, and cyclists was ignored. Hence, they were manually removed from the point clouds as they can influence the calculations by occluding some road points beneath and behind them.

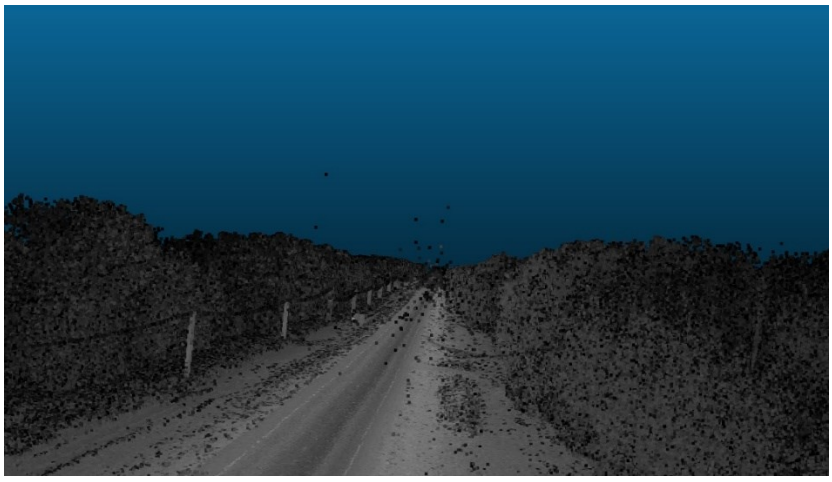

Figure 3. Excess Points in The Point Cloud

Figure 4. Vehicle in The Point Cloud

3.2.2 Relevant Road Features Trimming

After cleaning the road sections, the relevant road features region (RRFs) is then trimmed out and extracted from the original point cloud. A Python code was developed to generate bounding boxes $\mathcal{B}_i$ that are centred and rotated around the trajectory points $\boldsymbol{r_i}$, which are uniformly spaced across the point cloud $\mathcal{P}$ (road points extraction procedure is explained in detail in section 3.4). Each bounding box $\mathcal{B}_i$ is defined with a set length and width; no constraints are imposed for the z direction. By aggregating the bounding boxes at each road point along the trajectory, a smooth region that simulates the right-of-way region can be extracted from the original point cloud $\mathcal{P}$. The width of the bounding box typically depends on the road itself. Nevertheless, it should encompass the width of the driving lanes, road shoulders, and an extra space beyond the shoulder that can extend to a couple of meters to accommodate roadside clear zones [195], such that all road and traffic signs are present in the analysis. Algorithm 1 outlines the procedures implemented for right-of-way trimming.

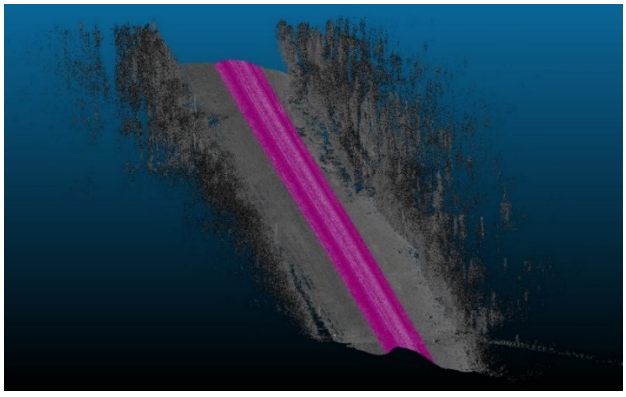

Figure 5. Trimmed Road Section (Highlighted in Purple)

Algorithm 1 Performing right of way trimming

Input: $\mathcal{P}$, $\mathbf{r}[\,] = \{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N\}$, $\mathbf{R}[\,] = \{\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_N\}$, h, R
Output: $\mathcal{P}_{\text{trimmed}}$

function TRIMPOINTCLOUD($\mathcal{P}$, $\mathbf{r}[\,]$, $\mathbf{R}[\,]$)
  for $i \leftarrow 0$ to N **do**
    $\mathcal{B}_i \leftarrow GetBoundingBox(i)$
    $\mathcal{B}_i \leftarrow \mathcal{B}[i] \times \mathbf{R} + \mathbf{r}_i$ ▷ Align bounding box along each road point
    $\mathcal{P}_{\text{sliced}}[i] \leftarrow SliceFromBound(\mathcal{P}, \mathcal{B}_i)$ ▷ Obtain points from $\mathcal{P}$ inside $\mathcal{B}$
  end for
  $\mathcal{P}_{\text{trimmed}} \leftarrow Unique(\mathcal{P}_{\text{sliced}}[\,])$ ▷ Reconstruct the path from all slices
  return $\mathcal{P}_{\text{trimmed}}$
end function

Symbols in Algorithm 1

Symbol	Description
$\mathcal{P}$	Original point cloud.
$\mathcal{P}_{\text{sliced}}[\,]$	Array containing slices of $\mathcal{P}$. Each slice is an array that contains XYZ points within bounding box $\mathcal{B}_i$ without constraints imposed on the z-coordinate.
$\mathcal{B}$	(4, 3) array containing four XYZ points defining a bounding box at $z = 0$. Bounding box $\mathcal{B}_i$ is a rectangle determined by a set width (perpendicular to the direction of travel) and length (along the direction of travel) at road point i.
$\mathbf{r}$	Road points from the trajectory. Road point $\mathbf{r}_i$ is an XYZ point given in UTM coordinates.
$\mathbf{R}$	Rotation matrices. Rotation matrix $\mathbf{R}_i$ is calculated from the forward vector $\mathbf{f}_i$ and the leftward vector $\mathbf{l}_i$ at point i.

Functions in Algorithm 1

Function	Description
$GetBoundingBox$	Obtain a bounding box $\mathcal{B}_i$ at a certain road point.
$SliceFromBound$	Obtains points from $\mathcal{P}$ that are within $\mathcal{B}_i$, on the XY plane.
$Unique$	Obtains unique points from all slices $\mathcal{P}_{\text{sliced}}[\,]$.

Figure 6. Algorithm (1) - Right of Way Trimming

3.3 TEST SEGMENTS

Throughout this analysis, multiple test segments with varying terrains and road geometries were used for various simulations to better understand the AV performance in different driving environments. Table 2 summarises the different road sections used in this analysis. Specifically, an 8 km stretch from Highway AB-816, situated southeast of Red Deer in a rural setting, was

chosen to analyze rolling terrain and vertical curve effects. Similarly, a 3 km section from AB-32, located in the rural west-central area of Alberta, was utilized for the same vertical curve assessments.

Moving on to roadway width variations, a 15 km portion from AB-3, also known as Crowsnest Highway in southern Alberta, was analyzed. Lastly, a 4 km segment from Highway AB-11 (David Thompson Highway) in central Alberta, amidst mountainous terrains near Banff National Park, was investigated to study horizontal curves. A 4 km stretch of Highway AB-1A, connecting Canmore to Calgary, was also selected for the horizontal curve analysis due to the environment's mountainous nature.

Table 2. Summary of Test Segments.

ID	Highway ID	Length (km)	Terrain	Characteristics
1	AB-816	4	Rolling	Vertical Curves
2	AB-32	3	Rolling	Vertical Curves
3	AB-816	4	Rolling	Vertical Curves
4	AB-3	4	Level	Roadway Width
5	AB-3	4	Rolling	Roadway Width
6	AB-3	3	Level	Roadway Width
7	AB-3	4	Level	Roadway Width
8	AB-11	2	Mountainous	Horizontal Curves
9	AB-11	2	Mountainous	Horizontal Curves
10	AB-1A	4	Mountainous	Horizontal Curves

3.4 TRAJECTORY GENERATION

The sensor has to be placed in its natural position in the virtual environment, typically on top of the AV, to generate accurate simulations. So, first, the position of the AV with respect to the road itself in the virtual environment has to be identified. Once this is done, the sensor's location can be specified to be above ground by a certain distance.

In this analysis, it was assumed that the AV would follow the trajectory of the data collection vehicle, done by capturing the road points that had a scan angle rank of zero, or in other words, the road points that lie directly beneath the vehicle that collected that point cloud data. A parametric curve then smoothed the trajectory $\vec{P}$(t). By setting a uniform spacing of 1m between the trajectory points that lie on $\vec{P}$(t), a set of road points $\boldsymbol{r} = \{\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{r_n}\}$ was created, achieved by reparametrizing $\vec{P}$(t) using the arc length, resulting in a function $\vec{P}$(s) where s has a value of 0, 1, …, N.

Three direction vectors are defined for each road point $\boldsymbol{r_i}$ on the parametric curve $\vec{P}$(t): forwards $\overrightarrow{f_i}$, leftwards $\overrightarrow{l_i}$, upwards $\overrightarrow{u_i}$. These vectors are essential in identifying the orientation of the road points. The upward vectors U = $\{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_N}\}$ can be derived by calculating the normal of the positive z component for a plane that contains nearby points and is centred around any given road point $\boldsymbol{r_i}$. The best least square fit for the defined plane is used to get accurate estimates of the upward vectors. As for the search range for nearby points, it is defined by a cube that has an edge length of 2 m. The forward vectors F = $\{\overrightarrow{f_1}, \overrightarrow{f_2}, \ldots, \overrightarrow{f_N}\}$ were obtained using the upward component and the parametric curve $\vec{P}$(t) using this equation: $\overrightarrow{f_i} = \vec{P}(t) - \overrightarrow{u_i} \cdot \frac{\vec{P}(t) \cdot \overrightarrow{u_i}}{\overrightarrow{u_i} \cdot \overrightarrow{u_i}}$. Last but not least, the leftwards vectors L = $\{\overrightarrow{l_1}, \overrightarrow{l_2}, \ldots, \overrightarrow{l_N}\}$ were obtained by having the cross-product of the upward and forward vectors $\overrightarrow{\boldsymbol{l_i}} = \overrightarrow{u_i} \times \overrightarrow{f_i}$. All the directional vectors were normalized to have a magnitude of 1.

After calculating the direction vectors for each point, observer points can be obtained by translating the road points upwards by a fixed distance of 1.8m, which takes into account the height of the ego vehicle and an additional clearance to accurately capture the position of the LiDAR sensors on top of vehicle's roof. AV models typically mount LiDARs at elevations close to this

figure [196] [197]. The translated observer points will inherently have the directionality of the leftward, forward, and upward vectors of the road points. $\overrightarrow{o_i} = r_i + 1.8 \cdot \overrightarrow{u_i}$.

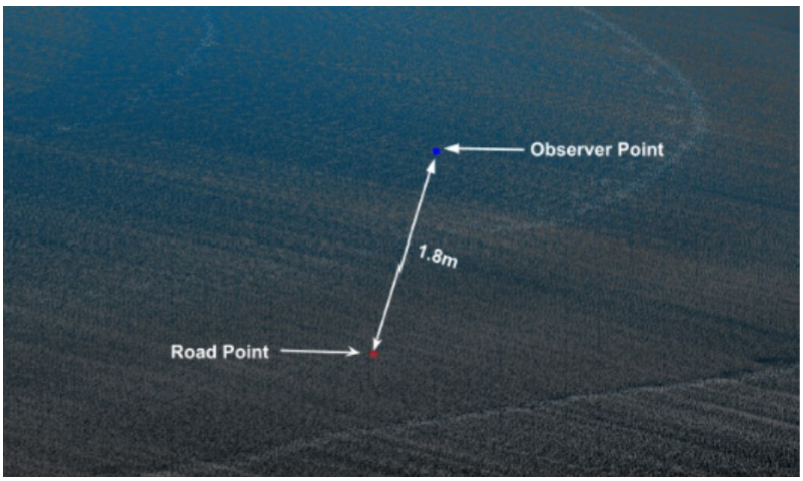

Figure 7. Road Point Vs Observer Point

Figure 8. Trajectory Points Demonstration

3.5 VISTA VIEWPOINT SYNTHESIS

The primary objective of this research is to investigate the dynamic nature of the surrounding environment and how AVs will interact with it. To achieve this, the FOV of the AV has to be simulated at different observer points along the road to capture what the AV would perceive from separate locations. By doing so, the data rate values required to process the scenes at each position or observer point can be determined effectively.

The VISTA simulator is one of many open-source, data-driven simulators available to researchers capable of implementing simulation-based techniques to develop sound algorithms for mobile agents [198]. VISTA is mainly sought-after for its ability to rapidly generate novel viewpoints and simulate sensors like cameras, LiDAR, and event-based cameras [198]. The generated perspective-based viewpoints exhibit excellent fidelity and precision based on predefined sensor configurations, allowing for a meticulous visualization of the surroundings around the vehicle at any instant along its trajectory.

First, the perspective must shift to a local road point to produce accurate synthesized viewpoints by applying a rigid transformation to the point cloud. The point cloud has to be rotated and translated using the trajectory information to implement such a transformation. For a road point r_i , the shifted point cloud from this point's perspective will be $\mathcal{P}_i = \mathcal{P} \times R_i + t_i$, where $\mathcal{P}$ is the original point cloud, R is the rotation matrix $R \in \mathbb{R}^{3x3}$ and t stands for the translation of the point cloud using the coordinates of the road point. Once done, the Euclidean distance between the newly found origin and the road points is calculated to cull all points not within the predefined sensor range. Finally, the elevation of the culled road points is subtracted by 1.8m to shift the perspective from a road point into an observer point. The culled range can be seen in Figure 9. An overview of this process is outlined in Algorithm 1.

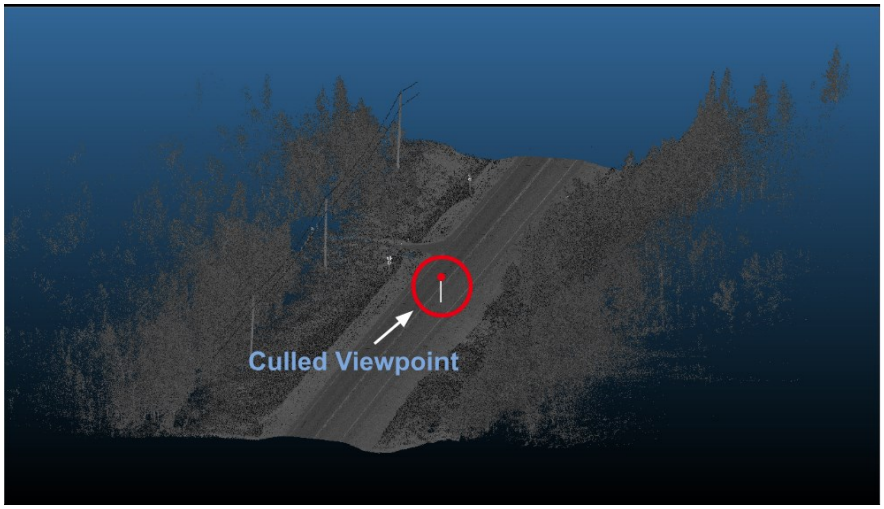

Figure 9. Culled Points Within the Sensor Range.

Once the global 3D point cloud is shifted into a local view from the perspective of the synthesized sensor viewpoint located at the previously defined observer point through rigid transformation, some of the road points that were once visible in the original 3D point cloud data

will likely be occluded due to the difference in perspective. Hence, an algorithm is needed to prevent blending points in the foreground with those in the background.

Algorithm 2 Preparing the point cloud for Vista

Input: $\mathcal{P}$, $\mathbf{r}[\,] = \{\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N\}$, $\mathbf{R}[\,] = \{\mathbf{R}_0, \mathbf{R}_1, \ldots, \mathbf{R}_N\}$, h, R
Output: $\mathcal{P}[\,]$

function CONVERTTOLOCAL($\mathcal{P}$, $\mathbf{r}[\,]$, $\mathbf{R}[\,]$, h, R)
  for $i \leftarrow 0$ to N **do**
    $\mathcal{P}_i \leftarrow \mathcal{P} - \mathbf{r}_i$ ▷ Translate point cloud
    $\mathcal{P}_{i,\,\text{culled}} \leftarrow CullByRange(\mathcal{P}_i, R)$
    $\mathcal{P}_{i,\,\text{rotated}} \leftarrow \mathcal{P}_{i,\,\text{culled}} \times \mathbf{R}_i$ ▷ Rotate
    $\mathcal{P}_{i,\,\text{observer}} \leftarrow \mathcal{P}_{i,\,\text{rotated}} - h$ ▷ Center point cloud on observer
    $\mathcal{P}_{\text{observer}}[i] \leftarrow \mathcal{P}_{i,\,\text{observer}}$
  end for
  return $\mathcal{P}_{\text{observer}}[\,]$
end function

Symbols in Algorithm 2

Symbol	Description
$\mathcal{P}$	Original point cloud.
$\mathbf{r}$	Road points from the trajectory. Road point $\mathbf{r}_i$ is an XYZ point given in UTM coordinates.
$\mathbf{R}$	Rotation matrices. Rotation matrix $\mathbf{R}_i$ is calculated from the forward vector $\mathbf{f}_i$ and the leftward vector $\mathbf{l}_i$ at point i.
h	Observer height given in meters.
R	Maximum range of the sensor in meters.

Functions in Algorithm 2

Function	Description
$CullByRange$	Cull out points from $\mathcal{P}$ of Euclidean distance greater than or equal to R from the road point $\mathbf{r}_i$.

Figure 10. Algorithm (2) - Point Cloud Preparation for VISTA

VISTA converts the 3D space view into a 2D image with polar coordinates representation where each pixel is represented by yaw (α) and pitch (β) angles. In addition to that, the ray distance (d) that connects between the observer point location and each pixel is calculated. The size of each

pixel will be equivalent to the horizontal and vertical accuracies of the used sensor. Additionally, the total number of pixels in each direction will be the range of yaw or pitch angles divided by the accuracy of each respective direction. If the resolution is high, more pixels and fewer points will be found within each pixel.

$$\alpha = \arctan\left(\frac{P_{i,y}}{P_{i,x}}\right) \qquad \beta = \arcsin\left(\frac{P_{i,z}}{d}\right) \qquad d = ||P_i|| \qquad (1)$$

For pixels that contain multiple points with variable depth values inside them, VISTA prioritizes the point that has the least depth and assigns its value to the whole pixel. This process, in principle, is similar to traditional raycasting techniques [130], which handle occlusions by electing points closest to the sensor and filtering out the points in the background.

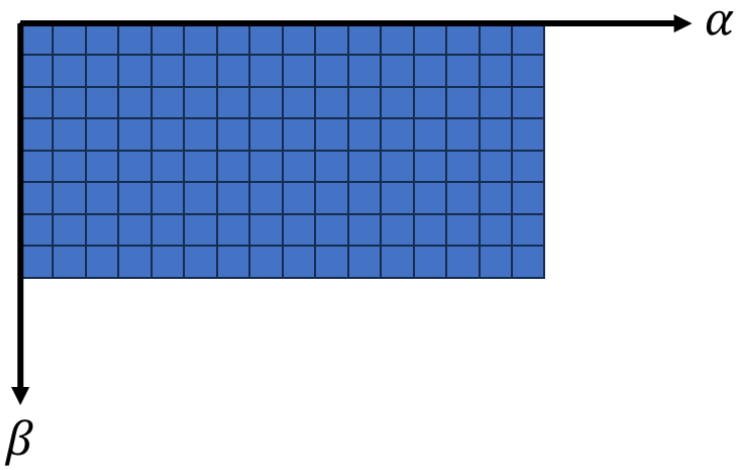

Figure 11. Image Representation Polar Coordinates with Yaw (α) And Pitch (β) Angles

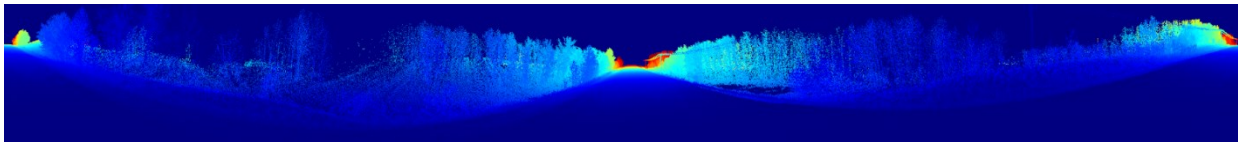

Figure 12. Image Representation from Local Observer Point Perspective with Depth Values

Unlike traditional ray casting algorithms, the occlusion handling from VISTA is more advanced as it adds an extra layer of robustness in identifying the occluded pixels. To determine whether a pixel is visible or occluded using VISTA, its distance (d) is compared to the average distance of its neighbouring rays. If the average distance of neighbouring pixels is less than the depth of the pixel in question, this means that there are points in the foreground of this pixel that occlude it from the sensor's perspective, and hence, this pixel is removed from the image.

Compared to traditional raycasting, the qualitative enhancement in VISTA's synthesized outputs is evident in Figure 13. Despite the sensor being positioned within the central circle of the road, the highlighted tree at the bottom of the frames should be occluded due to a line of trees obstructing the sensor's line of sight. Strikingly, it remains visible in the raycasting output. Moreover, the raycasting output displays an unreasonably high vegetation density at the frames' top, where self-occlusion should occur naturally. In contrast, VISTA shows a much lower density in this area. This discrepancy arises because raycasting sometimes allows rays to pass between neighbouring points that should be considered obstacles. Consequently, VISTA emerges as a more reliable tool for handling occlusion calculations, producing outputs that are considerably more accurate.

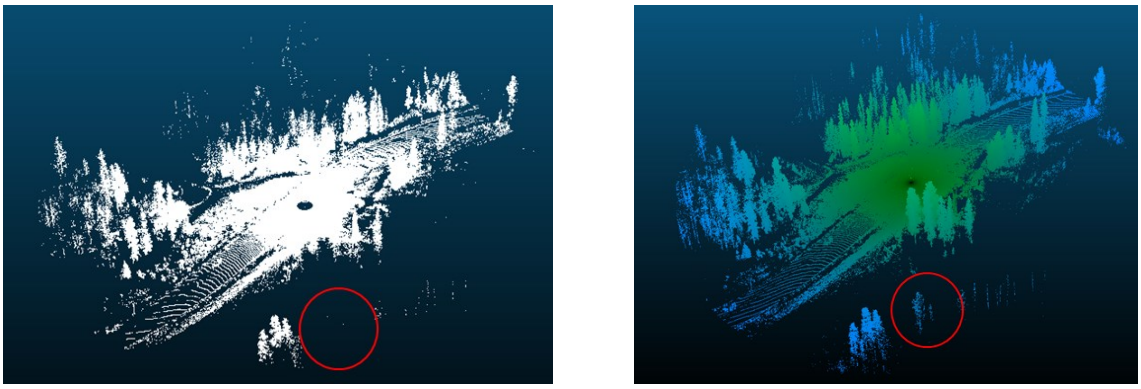

Figure 13. Vista Outputs (White) vs. Raycasting Output (Colored)

VISTA utilizes a variable called Culling Radius (λ) to define the region of neighbouring pixels for processing. The culling radius determines the number of neighbouring pixels to include in the calculations, thereby providing flexibility for users to choose the most suitable culling radius value that matches their data and the type of sensor used in simulations.

A culling radius of 0 will result in traditional raycasting outputs, as the algorithm will not use information from the surrounding pixels to check the visibility of the pixel in question. A value of one would mean that the algorithm will create a search radius one pixel wide around the queried pixel, as seen in Figure 14. VISTA then calculates the average depths of the surrounding pixels and compares it to the depth of the pixel in question to make a more informed decision about its visibility. That said, if the culling radius value is too high, it leads to inaccurate point cloud representations. When the search radius widens, the algorithm loses its locality and incorporates more distant points in its average neighbouring depth calculations. Naturally, this invites more

noise into the calculations, resulting in unstable outputs. For example, points far from the sensor are often sparse and low in numbers. By increasing (λ), far points are more susceptible to occlusions as the algorithm would include more neighbouring pixels, which will probably have shorter depths. Conversely, points located near the sensor are more densely packed, and if the search radius is increased, there is a bigger chance they would remain visible to the sensor.

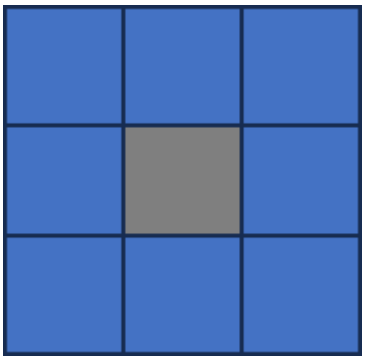

Figure 14. Culling Radius = 1 for the Middle Pixel (Neighboring Pixels in Blue)

Although the differences might be subtle through the visual examination of the outputs, Figure 15 to Figure 20 show the discrepancies in the VISTA outputs using different culling radius values. The output with a culling radius of 0 is very dense, as shown in the number of captured points (401,655). Such a number is unrealistic from this perspective. Using a culling radius of 8 returned a total of (238,167) points for the same frame. However, it was observed in the outputs that the level of detail is heavily obscured in the mid to far regions, which again does not accurately capture the outputs from this perspective. Finally, using a value of 2 for the culling radius produced more consistent and reasonable outputs along the sensor's range where the density of points was most fitting for the analysis. Hence, it was adopted for this study.

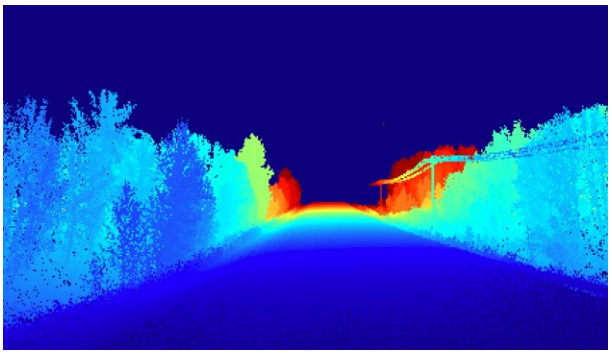

Figure 15. VISTA Image Representation with λ = 0

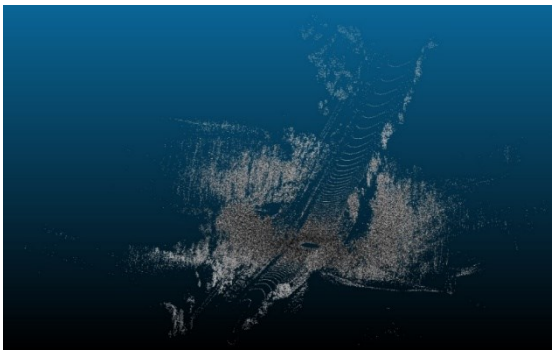

Figure 16. VISTA Outputs with λ = 0

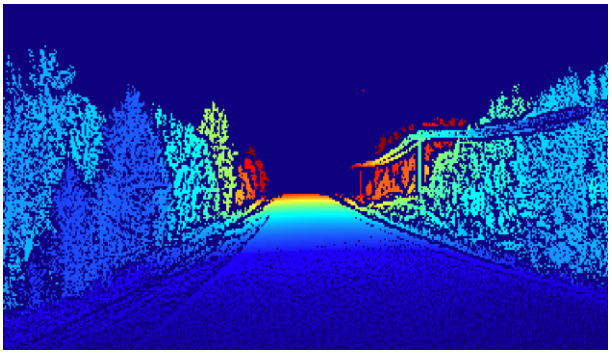

Figure 17. VISTA Image Representation with λ = 2

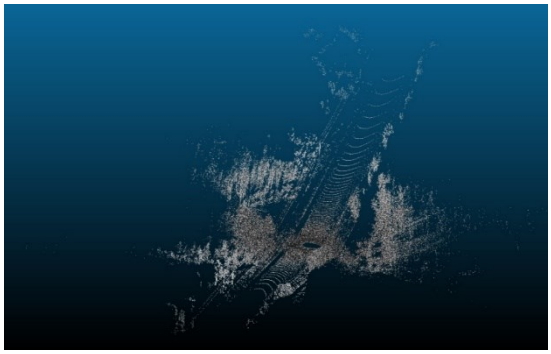

Figure 18. VISTA Outputs with λ = 2

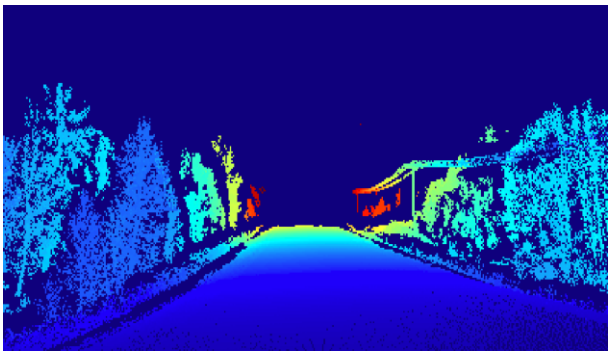

Figure 19. VISTA Image Representation with λ = 8

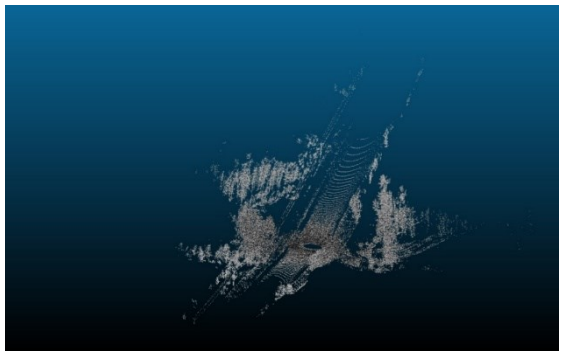

Figure 20. VISTA Outputs with λ = 8

Once the VISTA simulator is finished removing the occluded pixels, it proceeds with converting the points from the image representation back to the cartesian coordinate system by transforming the calculated (α_i), (β_i), and (d_i) of each pixel back to (x, y, z) by applying the inverse

of the transformation used in equation 1. Algorithm 3 presents an overview of the VISTA viewpoint synthesis.

Algorithm 3 Simulating the vehicle's viewpoint using VISTA

Input: $\mathcal{P}[\,]$,

Output: $\boldsymbol{\mathcal{P}}[\,]$

function SIMLIDAR($\mathcal{P}[\,]$, $\boldsymbol{\theta} \leftarrow \{\theta_l, \theta_h\}$, $\boldsymbol{\phi} \leftarrow \{\phi_l, \phi_h\}$, R, $\boldsymbol{\delta} \leftarrow \{\delta_\theta, \delta_\phi\}$, λ)
  for $i \leftarrow 0$ to N **do**
    $\mathcal{P}'_i \leftarrow ProjectToImage(\mathcal{P}_i, \boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\delta})$
    $\mathcal{P}'_{i,\,\text{culled}} \leftarrow CullOcclusions(\mathcal{P}'_i, \lambda)$
    $\boldsymbol{\mathcal{P}}[i] \leftarrow SynthesizePointcloud(\mathcal{P}'_{i,\,\text{culled}})$
  end for
  return $\boldsymbol{\mathcal{P}}[\,]$
end function

Symbols in Algorithm 3

Symbol	Description
$\mathcal{P}[\,]$	Array containing N of our previously prepared point clouds in local coordinates. Point cloud $\mathcal{P}_i$ is an (n, 3) array of XYZ points.
$\mathcal{P}'$	Point clouds projected to an image format.
$\boldsymbol{\mathcal{P}}[\,]$	Point clouds of the synthesized LiDAR viewpoint. $\boldsymbol{\mathcal{P}}_i$ is an $(m, 3)$ array of XYZ coordinates making up viewpoint i.

Functions in Algorithm 3

Function	Description
$ProjectToImage$	Project the XYZ point cloud into the VISTA simulator's image format.
$CullOcclusions$	Cull occluded points from the image representation based on the given culling radius λ.
$SynthesizePointcloud$	Reproject our point cloud from the image representation back to XYZ points.

Figure 21. Algorithm (3) - Synthesized Viewpoints from VISTA

3.6 DATA RATE CALCULATIONS

After developing the synthesized viewpoints for each observer point along the vehicle's trajectory and removing the occluded points based on each unique perspective, the following step in the framework is to quantify the data rate requirements for the AV for each frame or scene based on the distribution of points within each frame.

A recently published report estimated the data rate requirements that would prevent AVs from causing roadway fatalities. The report established that to achieve such a goal, the AV had to possess the ability to see sooner, clearer, and farther. In order to provide the AV with such capabilities, the onboard computer had to process a tremendous amount of sensor data in real time to reconstruct the entire scene around the ego vehicle with sufficient accuracy and frequency that would enable the AV to evade any possible obstacles in its path. The report developed an equation that captured all the necessary elements for estimating the required data rates [84].

Since there has not been much literature regarding the development of equations that would enable the rough estimation of the AV's expected performance under different weather and driving conditions, the authors of this paper opted to adopt the equation developed by the aforementioned report and use it not to report on definite, absolute data rate values. Instead, the goal is to use the equation to explain AVs' performance when encountering different driving environments and how their equipped LiDAR sensors are expected to interact with such environments. Additionally, the research's attention is directed toward the static environment and does not consider the other dynamic objects on the road, like vehicles, cyclists, or pedestrians.

The adopted equation had some assumptions which were also considered in this research. The first assumption is that calculations would be performed to simulate an end-to-end model for the onboard computer, meaning that for the onboard computer to complete the four core tasks of environment perception, localization, decision-making, and vehicle control, it would have to process the estimated data rate requirements in real time. The second assumption is that the sensors' point cloud data must be voxelated to calculate delta (Δ), which is the percentage of occupied voxels. Additionally, the equation suggests that the onboard computer will implement atomic norm compressed sensing, reducing the required calculations by the fraction of occupied voxels.

Compressed sensing algorithms like the Atomic Norm are known for their ability to handle signal processing problems by exploiting the signals' sparsity, allowing them to reduce the required number of measurements while maintaining a certain level of performance [199] [200]. Since AVs are naturally equipped with sensors that operate by transmitting and receiving signals like LiDAR and radar [201], this algorithm was incorporated to address the big data problem caused by LiDAR sensors by minimizing the needed computations.

For each frame (i) in the point cloud, the corresponding data rate value can be calculated using Equation 2.

$$\boldsymbol{d}[i] = \frac{R(\theta_h - \theta_l)(\phi_h - \phi_l)}{\delta_R \delta_\theta \delta_\phi} \cdot \frac{32Fb\Delta_i \log\left(\frac{1}{2\Delta_i}\right)}{3\mathrm{SNR}_{\max\,\mathrm{range}}} \tag{2}$$

The adopted equation considers multiple factors instrumental in estimating the required data rates for the AV to process its surroundings. And since the ego vehicle's primary source of sensory vision is the LiDAR, the considered factors are mainly related to the LiDAR equipped with the AV. In the proposed equation, R is the maximum sensor's range in meters, θ is the azimuth or horizontal angle of the sensor, ϕ is the elevation or vertical angle of the sensor, and they are both in degrees, δ_R, δ_θ, δ_ϕ are the range, azimuth, and elevation precisions, respectively, which are also in meters and degrees. F is the refresh rate for the used sensor in hz.

Analog-to-digital (ADC) LiDAR digitizes the oncoming signals to obtain range information [208], and b is the accuracy in bits for each sampled signal, which measures the required data rates in bits/s. For this study, a value of 12 is used. Turning to delta (Δ), it is the percentage of occupied voxels, and finally, SNR_{max} is the signal-to-noise ratio at the maximum range in dB.

All the terms of the equation are sensor-specific except for the delta (Δ), meaning that they can be treated as constants unless the sensor is changed because each sensor has its operational field of view that is dictated by its range (R), its horizontal (θ) and vertical angles (Ø) along with their precisions. The frame rate is also sensor-related as it denotes the frequency by which the sensor captures a scene.

The signal-to-noise ratio (SNR) is a term that enables the simulation of weather conditions in the analysis. It is established that LiDAR operating on higher SNR levels will enhance their

detection range and precision [202]. Furthermore, it is understood that in inclement weather conditions, the performance of LiDAR is affected significantly as the medium through which the signals of the LiDAR travel tend to carry a lot of noise. Especially at longer ranges, the signal's power reflected off a target is heavily attenuated, which might cause mis-detections if the signal strength falls below the noise strength [203], thus increasing the number of computations to process the surroundings accurately and, by extension, increase the data rate requirements. For instance, as the rain intensity increases, the point cloud perceived by the LiDAR decreases, directly affecting the AV's ability to detect and track objects [204]. To maintain optimal performance levels, the strength of the signal has to be higher than the strength of the ambient noise, resulting in a minimum SNR of more than one. A value of 12 dB, which simulates a signal-to-noise power ratio of 16, has been used in this analysis to simulate normal weather conditions [205].

As outlined in [206], heavy rain weather causes attenuation to 905nm LiDAR's signal strength with a rate of $35\ dB/km$. It was reported that LiDAR lose 10-15 dB from the strength of their signals within 50 m of fog [207]. Meanwhile, in a controlled lab experiment, the received signal power for lasers operating on 905nm wave length through thick fog and smoke was observed to be approximately -30 and -40 dB, respectively [208].

$$SNR(dB) = 10 \cdot \log(\frac{Signal\ Power}{Noise\ Power}) \quad (3)$$

Delta is a variable that captures the density and complexity of the surrounding environment from the perspective of the equipped LiDAR sensor. It can be calculated by first voxelating the point cloud space and then finding the portion of the voxels occupied by points. In this framework, a voxel is either occupied or empty, so if a single voxel is occupied by more than one point, only a single point is kept, and the rest of the points are removed from the calculations.

Voxelization of point clouds is a process that is typically used to reduce the volume of the point cloud data by their discretization into a 3D grid in space where the point cloud is represented by a set of voxels [146]. Typically, voxels are laid out in a cartesian space where their geometry, length, width, and height are usually defined by a single value that dictates the voxel size, which also defines the resolution of the point cloud as a whole [209]. However, in this framework, voxelization was carried out in spherical coordinates to simulate the ring pattern of LiDAR and the inherent sparse nature of points that are distant from the sensor [210].

In this spherical system, the voxel size will be smallest near the sensor and largest at the end of the sensor's range. By doing so, there will be more voxels in the near range to accommodate the higher number of points and fewer voxels in the far range to match the lesser number of points at the periphery, thus minimizing the loss of information or points during the analysis while maintaining efficient and accurate results [211]. The voxel size is contingent on the precision of the used LiDAR sensor in the simulations, and a demonstration of voxel sizes along the sensor range using a spherical coordinate system is shown in Figure 22.

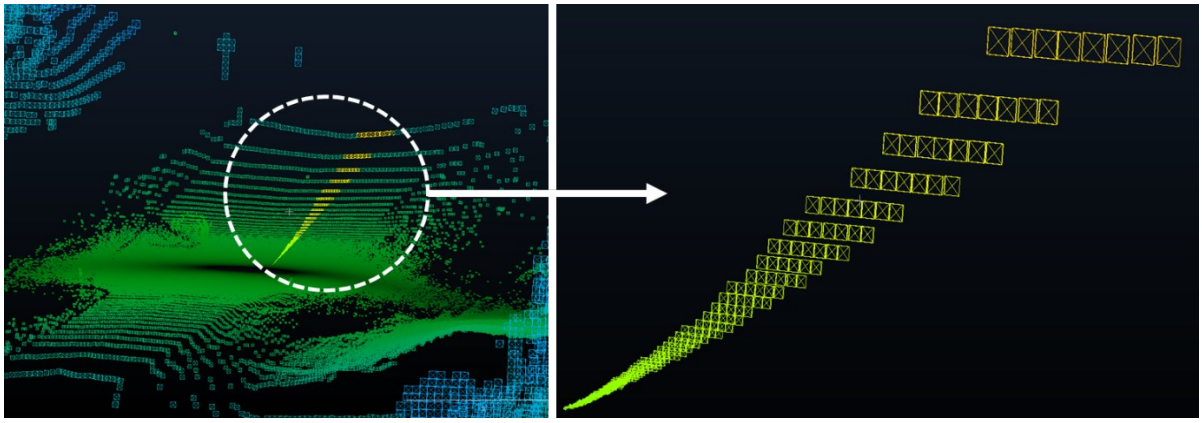

Figure 22. Varying Voxel Sizes Along the Sensor's Range

In the spherical coordinate system, each point in space is defined by azimuth (θ) and elevation (∅) angles, along with the range (ρ), which is the Euclidean distance from the observer point. Converting the cartesian system to the spherical system can be done through the following equations, where θ is measured counterclockwise from the positive x-axis, and its values range from 0° to 360°, while ϕ is calculated from the negative z-axis and can have values between 0° and 180°. Figure 23 demonstrates a voxel in the spherical grid space. Note that the angles calculated in this equation are in degrees.

$$\boldsymbol{\mathcal{P}}(x,y,z) \Rightarrow \boldsymbol{\mathcal{P}}(\rho,\theta,\phi)\begin{cases} \theta = \tan^{-1}\left(\frac{y}{x}\right) \\ \phi = \cos^{-1}\left(\frac{z}{\sqrt{x^2+y^2}}\right) \\ \rho = \sqrt{x^2+y^2+z^2} \end{cases} \tag{4}$$

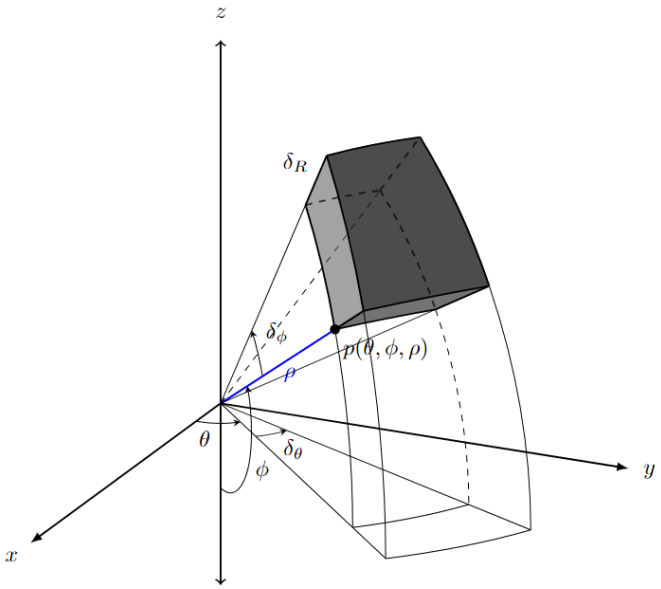

Figure 23. Voxels in Spherical Coordinates

Identifying the locations of all occupied voxels in the space involves calculating the indices of each occupied voxel. Then, the floor of the indices is taken since they correspond to specific locations on the spherical grid; hence, they have to be integers. In doing so, the points are grouped into their respective voxel. The voxel index corresponds to the location of the occupied voxel in the spherical grid and can be obtained by dividing the spherical coordinates of each point by their respective precisions ($\delta_R, \delta_\theta, \delta_\phi$) to get their projection inside the spherical grid system defined by the sensor configuration. Finally, after flooring the voxel indices, they are reverted to their spherical coordinates by multiplying them with the defined precisions.

The density of the points at any region is dictated by the angular precision, which, in principle, controls the minimum gap between two consecutive distinguishable points in the cloud, as seen in Figure 24 and Figure 25. Naturally, the spacing between two neighbouring laser beams increases away from the sensor since they are all fired from the location. Hence, this gap reaches its biggest value at the maximum sensor range, leading to a notable reduction in point density. The gap is minimal when close to the sensor, resulting in significantly higher point cloud densities [212].

$$\boldsymbol{Gap = distance\ to\ the\ target\ x}\ \tan\left(\boldsymbol{vertical\ resolution}\right) \quad (5)$$

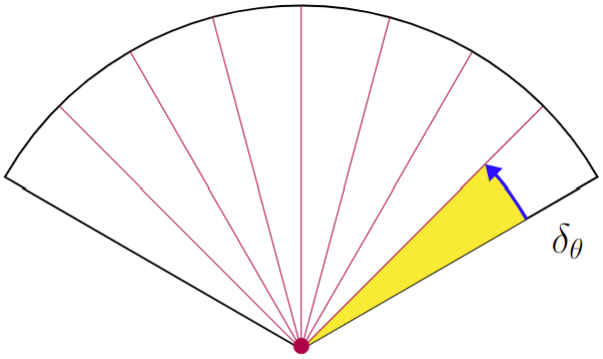

Figure 24. Top View for Sensor's Angular Resolutions (Horizontal Precision)

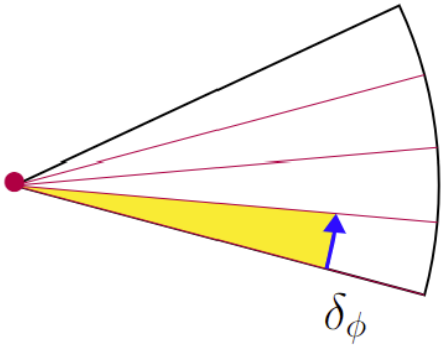

Figure 25. Side View for Sensor's Angular Resolutions (Vertical Precision)

Figure 26 showcases a planar view of the voxelated space in polar coordinates, where the points captured by the sensor are in blue, the occupied voxels are in red, and the green points correspond to the unique base point of the voxel.

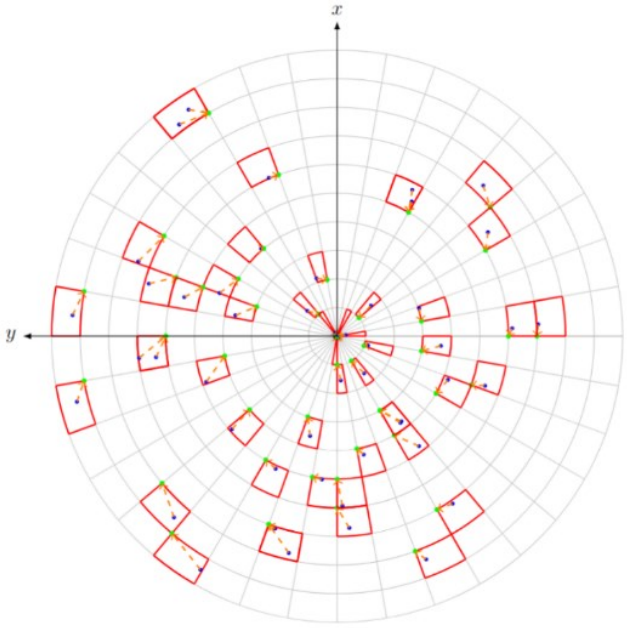

Figure 26. Planar View for Occupied Voxels and Their Base Points

Upon voxelating the point cloud, there are two approaches to calculating the percentage of occupied voxels. In this research, the analysis will be mainly carried out using the occupancy approach, and the second approach (volumetric approach) will highlight the difference between the two methods.

3.6.1 Occupancy Method

The primary methodology employed for this analysis centres around the occupancy method. This approach effectively complements using a spherical coordinate system for voxel representation. It considers the characteristic ring pattern effect produced by LiDAR sensors, where point density is notably greater near the sensor and decreases with distance. In this method, the calculation of the delta involves determining the ratio of the total number of occupied voxels (k) to the total number of voxels within the space. Points are assigned equal contribution weights regardless of their position from the sensor in this approach. However, the proximity of points to the sensor results in a higher contribution due to their dense presence. Conversely, the contribution of more distant points is lower, attributable to their sparser distribution. The delta value for each frame (i), following the occupancy approach, can be calculated using Equation 6.

$$\Delta_i = \frac{k}{\left\lfloor \frac{\theta_h - \theta_l}{\delta_\theta} \right\rfloor \times \left\lfloor \frac{\phi_h - \phi_l}{\delta_\phi} \right\rfloor \times \left\lfloor \frac{R}{\delta_R} \right\rfloor} \tag{6}$$

3.6.2 Volumetric Method

The volumetric method adopts a different approach in calculating delta where more weight is assigned to the distant points rather than the closer ones to account for the scarcity in their numbers. This calculation is done by utilizing the voxel size, which is maximum at the sensor's maximum range and minimum in the vicinity of the sensor. Hence, the contribution of the distant points will be high owing to their bigger size. This method attempts to simulate the complexity of processing distant points as they are typically scarce in numbers and will require higher sampling rates to be accurately probed by the sensor. Such points pose a significant challenge for AVs because recognizing and responding to them accurately is inherently difficult [213].

The corresponding delta can be calculated by dividing the aggregate volume of all occupied voxels $\mathcal{V}_i$ by the total volume (V) defined within the bounds of the sensor's field of view (FOV). To obtain $\mathcal{V}_i$, the voxel indices must revert to spherical coordinates by multiplying the voxel indices by their respective precisions, resulting in the base points p_i of all occupied voxels where $p_i = \{p_1, p_2, \dots, p_k\}(\theta, \phi, \rho)$. From the voxel representation $\boldsymbol{v_i}$, for each unique voxel's base point (j), the volume of the occupied voxel can be calculated through integration by knowing the

precisions of the sensor. Upon calculating the volume of all occupied voxels up the k^{th} base point, $\mathcal{V}_i$ can then be obtained by summing the volume of the occupied voxels as seen in the following equation:

$$\mathcal{V}_i = \sum_{j=1}^{k} \int_{\rho_j}^{\rho_j+\delta_\rho} \int_{\phi_j}^{\phi_j+\delta_\phi} \int_{\theta_j}^{\theta_j+\delta_\theta} \rho_j^2 \sin\phi_j \; d\theta \, d\phi \, d\rho \tag{7}$$

Finally, delta can be calculated using the volumetric approach by dividing the volume of occupied voxels by the total volume covered by the sensor FOV using Equation 5.

$$\Delta_i = \frac{\mathcal{V}_i}{\int_{R_l}^{R_h} \int_{\theta_l}^{\theta_h} \int_{\phi_l}^{\phi_h} R^2 \sin\phi \; d\phi \, d\theta \, dR} \tag{8}$$

3.7 PADDING REGIONS

Owing to the 360° vision of the LiDAR sensor, the surroundings can be effectively scanned both ahead and behind the sensor, with a distance equivalent to its operational range. Hence, to overcome the problem of having a start and an end to the test segments, where at the beginning of the road section, the region behind the sensor is typically empty, and conversely, at the end of the road section, the region ahead of the senor is also vacant, a solution is devised through the introduction of padding regions at these extremities of the test segments. These regions are introduced at both the start and the end of the road section and have a range equivalent to the sensor's range. In doing so, the consistency of the calculations will not be compromised, as this will ensure that the calculation region for each frame along the road section is constant and equal to twice the sensor's range. Figure 27 shows an example of road section padding.

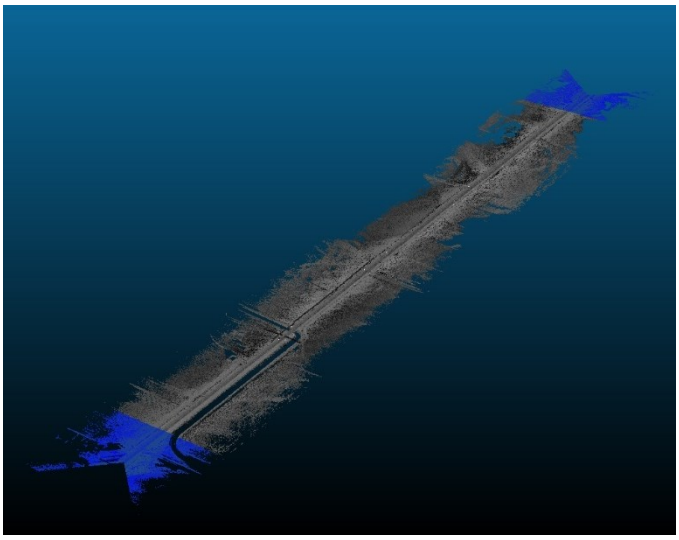

Figure 27. Padded Regions in Blue at The Ends of a Road Section

3.8 SENSOR SPECIFICATIONS

The various parameters of the sensor specifications used in the simulations were modelled on real-world sensors. In particular, The state-of-the-art Velodyne Alpha Prime LiDAR [214] sensor, also known as (VLS-128), is used as the primary sensor for this framework. Table 3 illustrates the adopted values for the range of the sensor (R), the ranges for both the horizontal (θ) and vertical angles (ϕ), along with the sensor's precisions (δ) and operating rotational rate (F). Figure 29 and Figure 30 demonstrate a graphical representation of the sensor's FOV.

Figure 28. Velodyne Alpha Prime (VLS-128) LiDAR

Table 3. VLS-128 Sensor Specifications.

Parameter	R	θ_h	θ_l	ϕ_h	ϕ_l	δ_R	δ_ϕ	δ_θ	F
Value	245m	180°	-180°	15°	-25°	0.03m	0.11°	0.11°	20hz

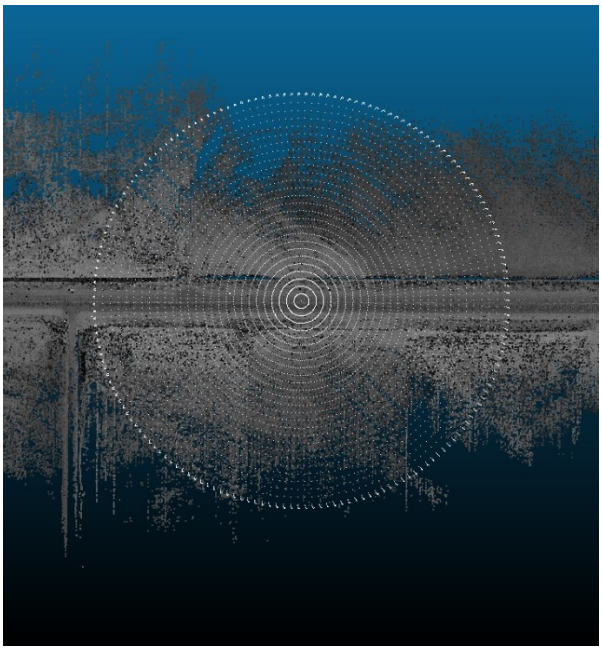

Figure 29. Sensor Field of Vision (Top View)

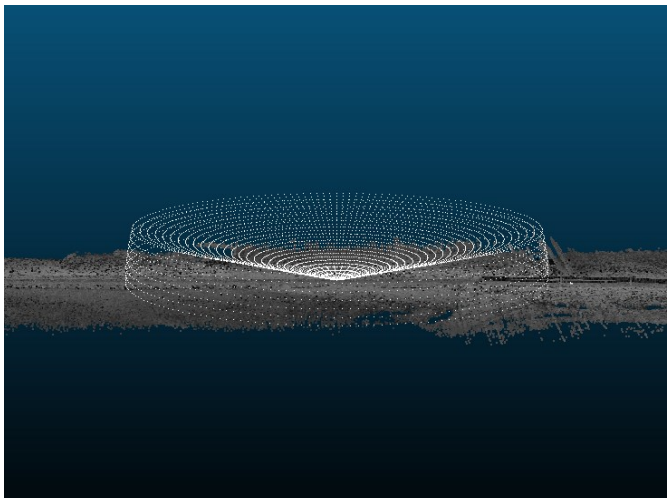

Figure 30. Sensor Field of Vision (Side View)

4 ANALYSIS AND RESULTS

The simulations implemented in this analysis are divided into two folds. The first fold comprises the full environment (FE) analysis, and the second fold focuses on the road's main relevant features (RRFs). This approach enables a comparative evaluation of results, elucidating the extent of additional computational demands arising from extraneous roadside elements like vegetation.

Also, this analysis is dedicated to attempting to understand the relationship between the anticipated performance of the AV in different road environments in terms of the required processing power that the AV has to possess to traverse the road safely. The dynamic objects such as pedestrians, vehicles, and cyclists were neglected to reduce the analysis's complexity and isolate the different elements of the road geometry to better explain their influence on the ego vehicle's performance. Moreover, the impact of occlusion will also be demonstrated and discussed.

The foundation of this analysis rests on the premise that the ego vehicle predominantly relies on a LiDAR sensor as its primary sensory input, situated at a height of 1.8 meters above road level. Furthermore, ideal weather conditions are presumed throughout the analysis. Lastly, the specifications for the LiDAR sensor values are drawn from Table 3. The outputs of this analysis measure the expected values for the data rate requirements at each frame of the road section. The analyzed frames are spaced 1m apart across the road section. Furthermore, images from Google Maps Street View [221] will be provided along with the graphical data rate requirements outputs to validate that the analysis findings are reflected in the real-world environment.

4.1 VALIDATION OF IMPLEMENTED METHOD

Before engaging in the principal analysis, some validations were required to ascertain the quality of the proposed framework. Some control sections were extracted from the LiDAR point cloud data to this end. Such control sections were selected to match specific criteria, mainly the sections' flatness and the absence of any potential noise in the data that might cause random output fluctuations. Knowing that the vertical grade is rarely flat for highways, the selected road sections mostly lay on a steady vertical grade. Additionally, the analysis ignored the roadside elements due to the inherent randomness in their densities. Through this analysis, the variation in the collected LiDAR point cloud densities can be demonstrated and quantified, and their impact on the data rate

requirements can be assessed. In this analysis only, the range of the sensor was assumed to be 100m as the road sections that were fit to use as control sections were often short in length.

Using Cloud Compare software [215], the point cloud density (d) can be measured for every point within the road section. This was achieved by establishing a search radius ($\mathcal{R}$) and then dividing the number of points within this region ($\mathcal{N}$) by the corresponding volume of the defined neighbourhood. This tool was used to check for the consistency of road point densities for all the road sections used in this analysis; an arbitrary value of 0.5m for the search radius was used throughout the calculations.

$$d = \frac{\mathcal{N}}{\frac{4}{3} \cdot \pi \cdot \mathcal{R}^3} \tag{8}$$

Four control sections were used for this analysis, with lengths between 600 and 800 m. The change in the vertical gradient was minimal across the analyzed sections, as shown in Figure 31 and Figure 35. The chosen road sections had no significant crest or sag vertical curves that might cause major disturbance on the outputs. Additionally, there were no horizontal curves in the control sections to maintain the integrity of the analysis.

From the point density calculations, it was clear that the region surrounding the road trajectory had the highest point density, whereas the regions farthest from the road trajectory had lower values, as seen in Figure 32 and Figure 36 because the path chosen for the simulated vehicle and sensor was modelled to mirror the course of the scanning vehicle. The regions closest to the LiDAR sensor naturally have higher point densities, also reflected in the simulations. Moreover, it was also observed that the different point density regions had variations in their values across the road sections, reflected in the data rate calculations, and influenced the change in the data rate values. In addition, the extracted trajectory used in the analysis is not perfectly straight and can sometimes deviate slightly, adding to the fluctuations of the data rate values.

On comparing the different calculation approaches, it was noticed that the data rate values recorded by the occupancy method were consistently higher than those of the volumetric method because of the difference in methodology between the two approaches, as seen in the results. This is because the ring effect of LiDAR sensors and the implementation of a spherical coordinate system for the voxels prompts the density of the point cloud in the vicinity of the sensor to be

much higher. Therefore, the number of occupied voxels will also be much higher than those on the periphery of the sensor. Consequently, in the volumetric method, the voxels with higher contribution are few, resulting in lower (Δ) delta values and overall lower data rate values, unlike the occupancy method where all voxels are given the same weight, which means that the high number of occupied voxels in the near end will have more influence on the data rate calculations.

Moreover, it can be inferred from this analysis that the volumetric method is much more sensitive to fluctuations in values as it relies on the point density around the sensor along their distribution in space compared to the occupancy method, seen in the range of changes in values between the two approaches in the different road sections where higher variations were reported using the volumetric method.

Table 4 demonstrates the recorded maximum deviations from the average data rate values across the four control sections using the occupancy and volumetric approaches. Following the occupancy method, the recorded change in values for the control sections ranged from 0.6% to 1.1%, with an average of approximately 0.85% change in the data rate values. On the other hand, the volumetric method, being more sensitive to fluctuations, produced slightly higher variations in the data rate values that ranged from 1.6% to 3%, which amounted to an average of 2.3%.

From this analysis, it was understood that the input point cloud density for the road points, along with the generated trajectory, influenced the fluctuations in the output data rate values but only slightly. These variations were seen to be insignificant using the occupancy approach, which is the primary calculation method adopted in this analysis. In contrast, the volumetric method has a more observable impact on road point density. Nonetheless, even this impact can be regarded as minimal overall.

Table 4. Data Rate Calculations for Different Control Sections.

Control Section	Length	% Maximum Deviation Using Occupancy Method	% Maximum Deviation Using Volumetric Method
1	800m	0.6%	2.8%
2	600m	1.16%	2.6%
3	850m	0.86%	1.6%
4	700m	0.84%	3%

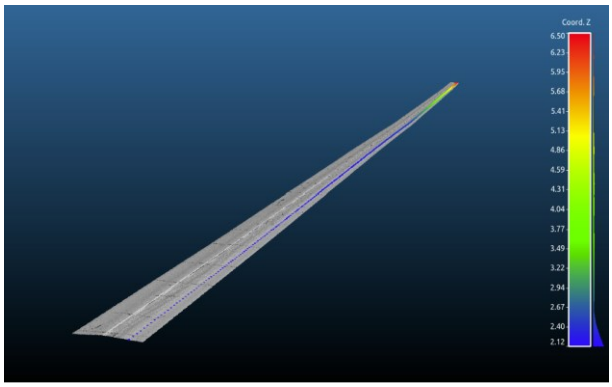

Figure 31. Vertical Profile for Control Section (1)

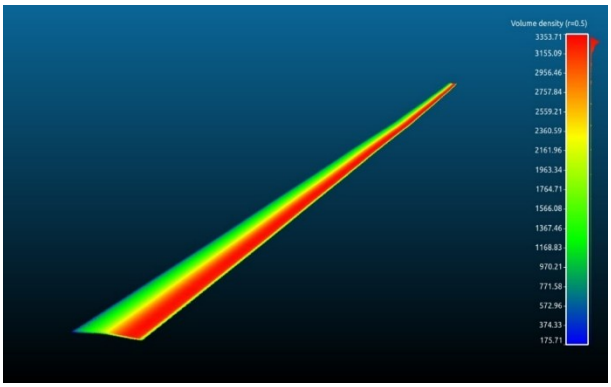

Figure 32. Point Density for Control Section (1)

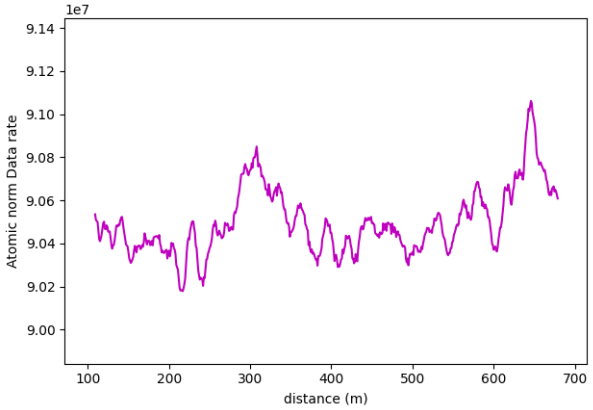

Figure 33. Occupancy Approach Data Rate Values for Control Section (1)

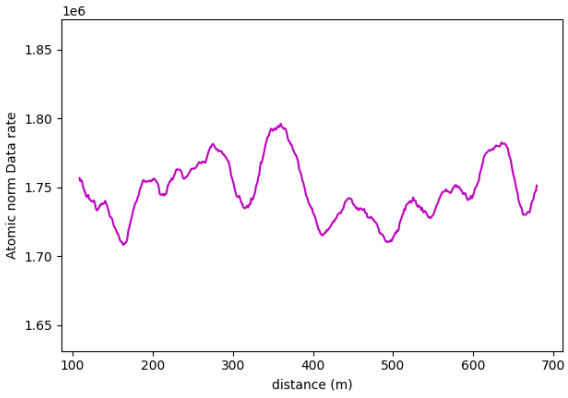

Figure 34. Volumetric Approach Data Rate Values for Control Section (1)

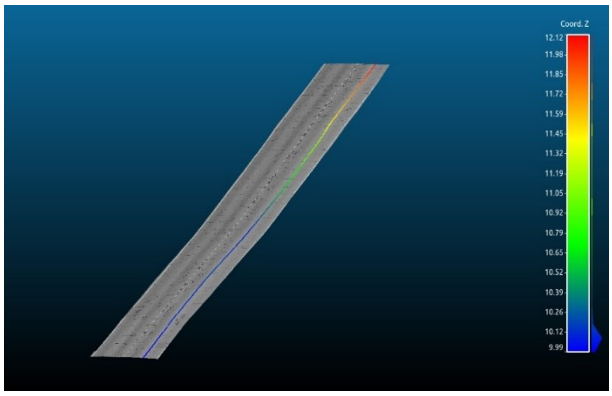

Figure 35. Vertical Profile for Control Section (2)

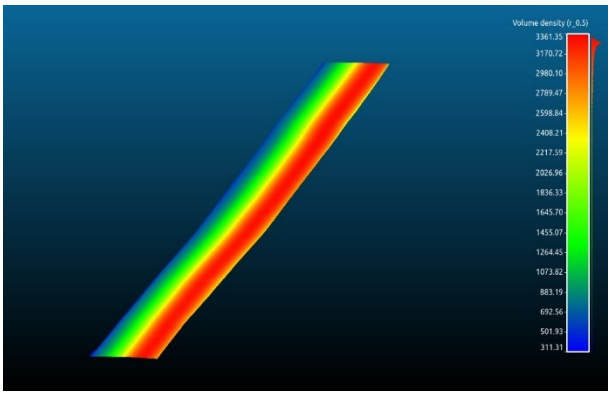

Figure 36. Point Density for Control Section (2)

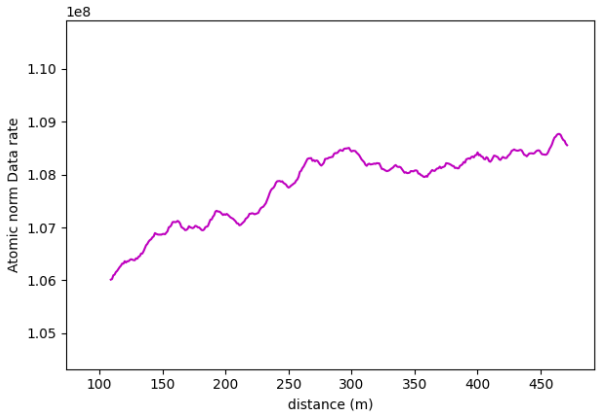

Figure 37. Occupancy Approach Data Rate Values for Control Section (2)

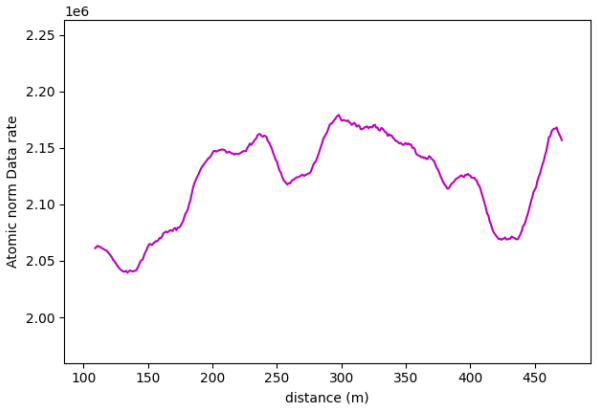

Figure 38. Volumetric Approach Data Rate Values for Control Section (2)

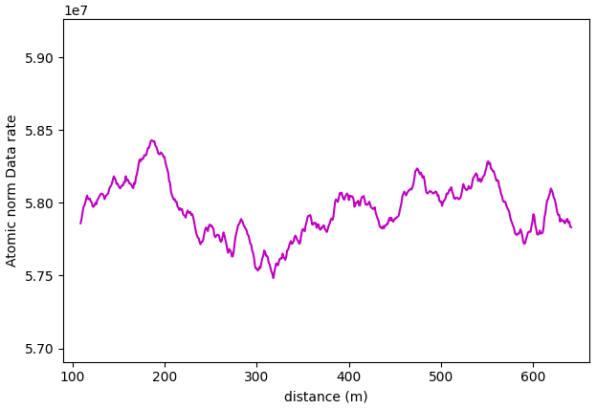

Figure 39. Occupancy Approach Data Rate Values for Control Section (3)

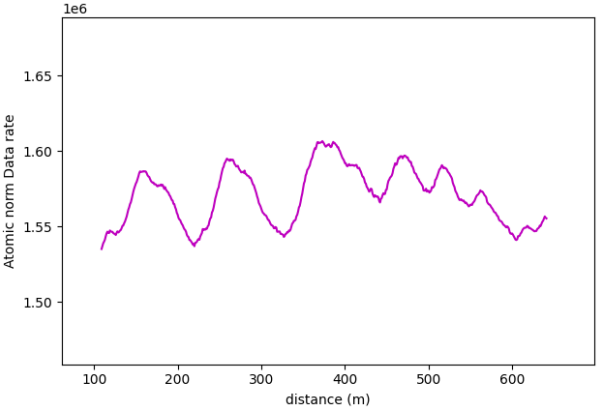

Figure 40. Volumetric Approach Data Rate Values for Control Section (3)

4.2 VERTICAL CURVES

The primary objective of this analysis is to examine the impact of vertical curves on data rate requirements in road sections. Vertical curves, which are common features found in most roadways, present a challenge for AVs because of their potential to compromise the line of sight or reduce visibility due to occlusions. Understanding the implications of these curves is crucial for ensuring AVs' safe and efficient operation.

Figure 41. Road Section (1) FE Layout

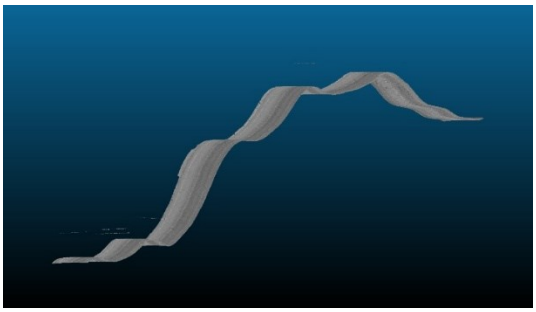

Figure 42. Road Section (1) RRFs Layout

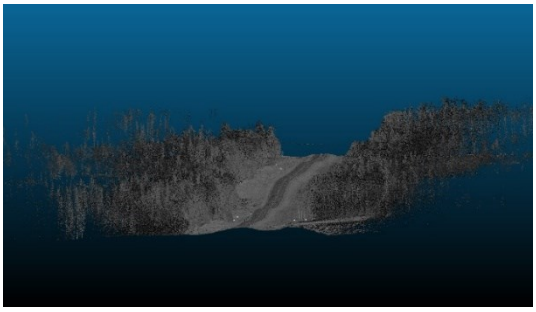

Figure 43. Road Section (2) FE Layout

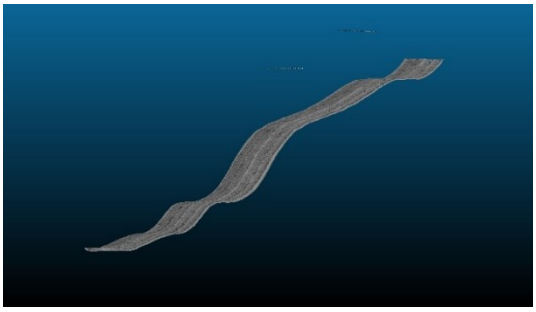

Figure 44. Road Section (2) RRFs Layout

Figure 45. Road Section (3) FE Layout

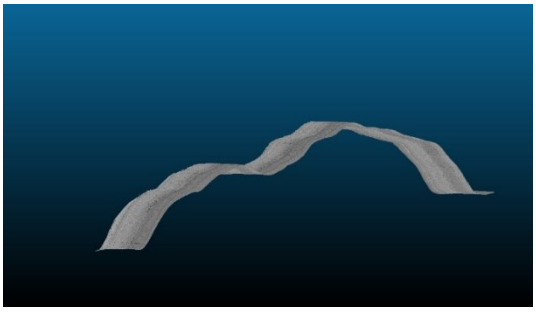

Figure 46. Road Section (3) RRFs Layout

To conduct a comprehensive investigation, the analysis compared the vertical curve impact once on the relevant road features (RRFs) and once on the full environment (FE). The analysis employed LiDAR vision simulations on three road sections (1-3) with varying vertical grades to simulate real-world scenarios; Figure 41 to Figure 46 showcase the general layout of the analyzed road sections. This analysis aims to investigate the influence of the vertical curves on the data rate requirements and overall performance of AVs in a controlled environment.

4.2.1 Vertical Curves (RRFs) Analysis

The average values of the data rate requirements at the flat road sections will be compared to their counterparts at vertical curve locations, including sag and crest curves, to isolate the vertical curves' influence on the AVs' performance.

On studying the data rate requirements graph for the road sections, a pattern was observed where the data rate values were, on average, lower at the crest vertical curves when compared to the sag vertical curves due to occlusion from the road geometry. For instance, the lowest values in the first road sections were traced to different crest locations, such as frames 665 and 1525. Conversely, the highest values on the graph, for example, at frames 480 and 881, were located on sag curves. These frames are demonstrated from Figure 50 to Figure 53. The lowest data rate value at crest locations was approximately $1.01x10^8\ bits/s$ while the highest value at the sag locations was $1.104x10^8\ bits/s$. Knowing that the average values at flat sections of the road were $1.05 \times 10^8\ bits/s$, a 5.1% increase in values was observed in sag curves, and a 3% decrease was seen at crest locations for the first road section. The sudden drop in values around the 2500m mark was due to an on-road vehicle occluding some road points.

Similar trends were evident in the second road section. Once again, on average, the data rate values at sag curve locations outperformed those at crest curves. In the second road section, the lowest recorded value at crest curves, exemplified by frames 824 and 1621, stood at $0.98x10^8\ bits/s$, whereas the highest at sag curves, represented by frames 600 and 1135, reached $1.04x10^8\ bits/s$. When these values are compared to the average value at flat sections, which averaged $1.01x10^8\ bits/s$, an increase of 8.3% at sag curves and a 4.4% decrease in values at crest curves was observed.

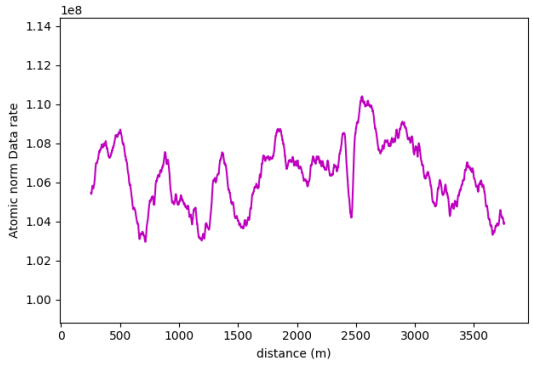

Figure 47. Data Rate Requirements for Road Section (1)

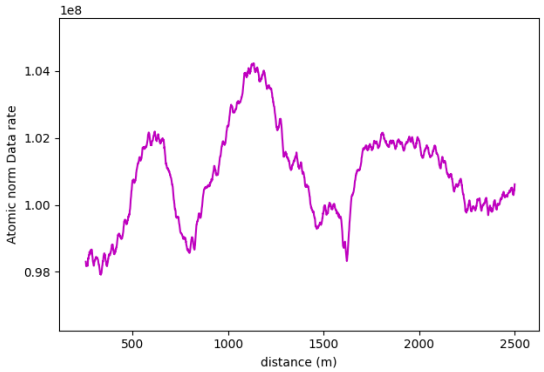

Figure 48. Data Rate Requirements for Road Section (2)

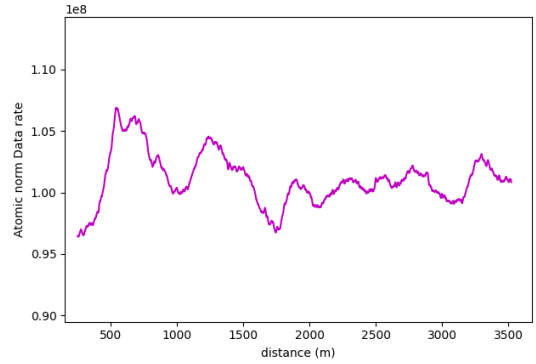

Figure 49. Data Rate Requirements for Road Section (3)

This consistent pattern persisted in the third road section, where frames such as 257 and 1700 exhibited the lowest values, primarily due to their location on crest curves, averaging 4.4% lower than values observed on flat road sections. Conversely, frames like 543 and 1240 recorded the highest values, with a maximum increase of 7% compared to flat sections.

Naturally, the visible range of the sensor is controlled by the geometry of the road, where the grade percentage is a crucial feature. For instance, the first road section's crest curve at frame 1525 was characterized by +3.6% and -0.8% tangents, respectively. As illustrated in the VISTA outputs in Figure 62, the range of vision on the -0.8% slope extended to approximately 200 meters, while the range on the +3.6% slope was restricted to 72 meters. Similarly, the crest curve at frame 1743 in the third road section was formed by +3.24% and -0.5% tangents. Upon examination of the VISTA outputs, shown in Figure 63, it was noted that the range of vision ranged between 115 meters at the +3.24% grade and 190 meters at the -0.5% grade, demonstrating that the road's

geometry indeed occludes parts of the road, typically those lying beyond the sensor's short range (beyond 50-60 meters).

By investigating the number of points captured within the sensor's short range and comparing them to the total number of points within an individual frame, it was established that, as a general trend, when only the RRFs are considered, approximately 95-97% of the captured points are located within the initial 50-60 meters around the sensor.

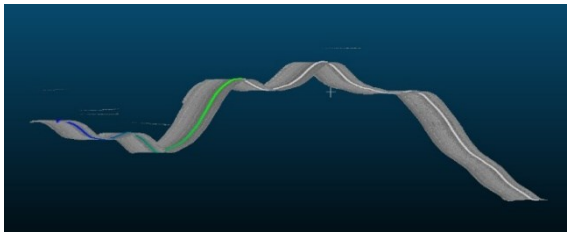

Figure 50. Road Section (1) - Frame 1525 (Crest) Location on The Road Section

Figure 51. Road Section (1) - Frame 1525 Location on Google Maps

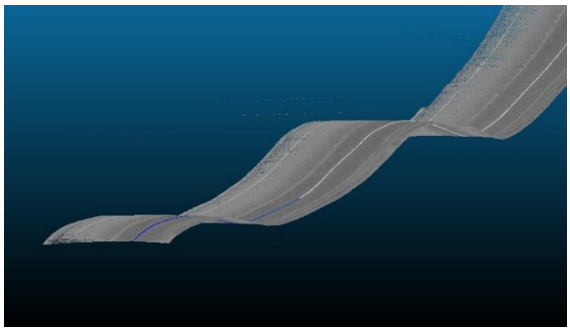

Figure 52. Road Section (1) - Frame 480 (Sag) Location on The Road Section

Figure 53. Road Section (1) - Frame 480 Location on Google Maps

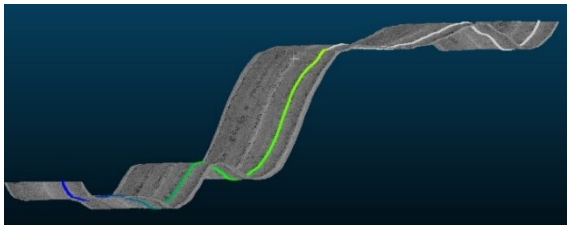

Figure 54. Road Section (2) - Frame 1621 (Crest) Location on The Road Section

Figure 55. Road Section (2) - Frame 1621 Location on Google Maps

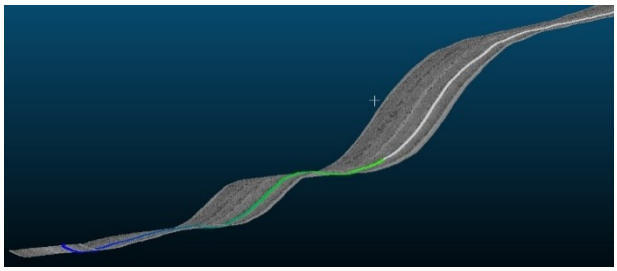

Figure 56. Road Section (2) - Frame 1135 (Sag) Location on The Road Section

Figure 57. Road Section (2) - Frame 1135 Location on Google Maps

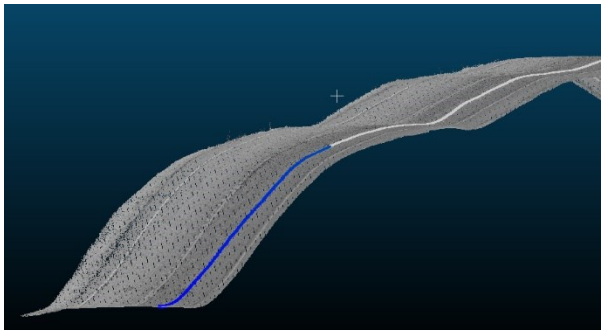

Figure 58. Road Section (3) - Frame 257 (Crest) Location on The Road Section

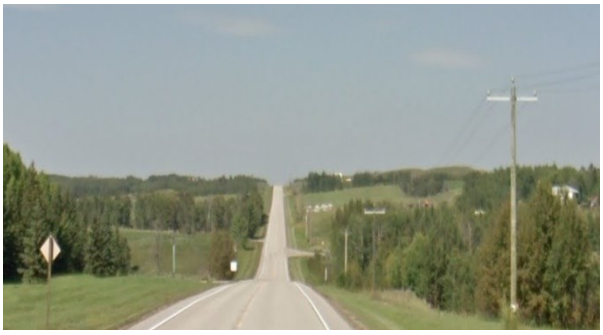

Figure 59. Road Section (3) - Frame 257 Location on Google Maps

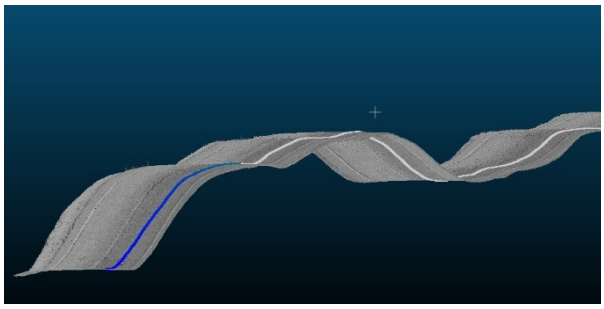

Figure 60. Road Section (3) - Frame 543 (Sag) Location on The Road Section

Figure 61. Road Section (3) - Frame 543 Location on Google Maps

This observation is exemplified by comparing frames 2552 and 665 in the first road section, situated on sag and crest curves, respectively. As shown in Figure 64, the VISTA outputs of frame 2552 (depicted in white) revealed a range of vision extending to 200 meters on each side, whereas it was constrained to 93 meters in frame 665 (shown in red). Frame 2552 contained 44,500 points, while frame 665 held 41,300 points. The number of points beyond the 90m range of vision in frame 665 amounted to only 756 points. This quantity constitutes a mere 1.6% of the total points in frame 2552. Such observations align with the typical behaviour of LiDAR sensors, wherein the

number of points in the far range is significantly lower than those close to the sensor, thus explaining why there are small margins among the data rate values in the vertical curve analysis, especially when considering the RRFs only.

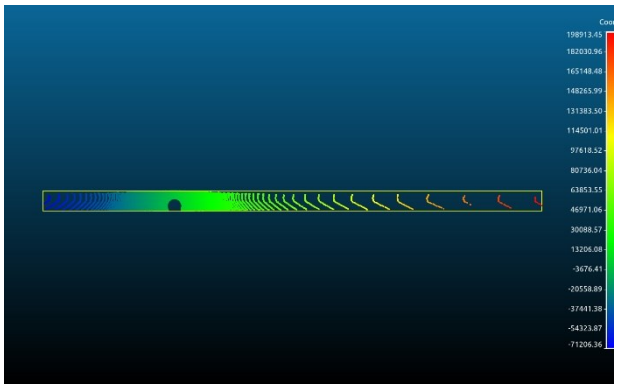

Figure 62. Road Section (1) - Frame 1525 VISTA Outputs (Vision Range = 72 - 200m)

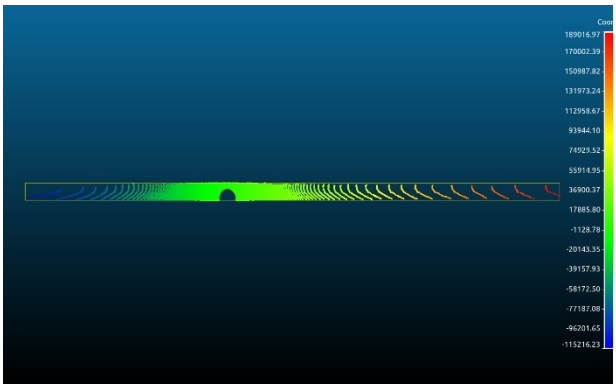

Figure 63. Road Section (2) - Frame 1743 VISTA Outputs (Vision Range = 115 - 190)

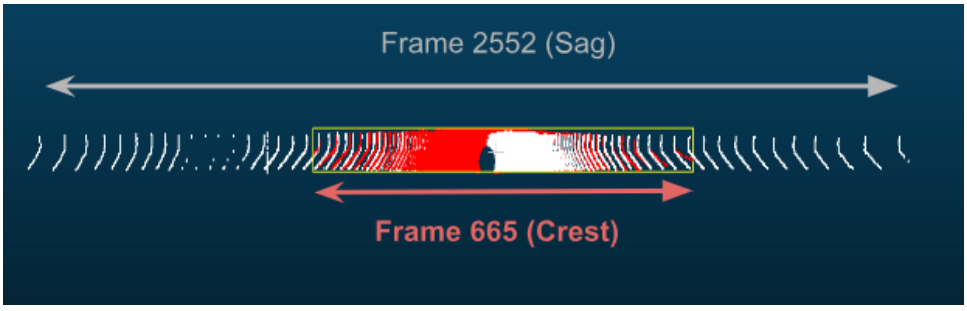

Figure 64. Road Section (1) - Comparing the Extent of Vision Between Frame 2552 Located on a Sag Curve (Depicted in White) and Frame 665 Located on a Crest Curve (Depicted in Red)

4.2.2 Vertical Curves (FE) Analysis

This analysis will compare the data rate requirements of the full environment (FE) and the relevant road features analysis (RRFs). Two distinct observations became apparent in this comparison. Firstly, the influence of vertical curves, whether sag or crest curves, was generally masked, and the previously established patterns were hardly visible. Secondly, data rate requirements increase substantially if the analysis considers the full scene.

Within the FE analysis, roadside features, particularly vegetation, will mainly exercise control over data rate values. The highest data rate value in the FE analysis is expected to be primarily found at locations with high vegetation density near the AV. Road geometry will now

serve as an auxiliary feature that can either boost or reduce data rate requirements at various vertical curve locations.

For instance, in road section (1), it was observed that the data rate requirements have experienced a fivefold increase in the FE analysis compared to the RRFs analysis. As for the data rate values, frame 880 was situated on a sag curve, resulting in a high value in the RRFs analysis. However, in the FE analysis, the density of surrounding vegetation was low, and the right side of the road was occupied by a waterbody, as demonstrated in Figure 66 and Figure 67, leading to a reduction in the total number of points captured by the sensor and, consequently, a decrease in overall data rate requirements. On the contrary, frame 480 yielded high values in both the RRFs and FE analyzes because it was located not only on a sag curve but also in an area with high vegetation density. These factors, in turn, contributed to it marginally having the highest data rate requirements in this road section.

On analyzing the second road section, the vegetation density was generally higher than the first and third sections. Consequently, a notable increase in data rate values was observed across the entire road section, particularly within the first 1.5 km, where vegetation density on both sides of the road remained consistently high. As a result, minor fluctuations were encountered in data rate values within this region. Figure 69 and Figure 70 illustrate the differences in the environment layout and vegetation density. Frame 1100 exhibited the highest data rate value, primarily due to its location's relatively high vegetation density. Meanwhile, frame 1600 recorded one of the lowest data rate values, owing to that area's comparatively lower vegetation density. Moreover, the positioning of frame 1100 on a sag curve likely contributed to its high value. Similarly, a crest curve at frame 1600 likely contributed to its low value.

However, this is not always the case. When examining the data rate values of frames 600 and 824, which were situated on a sag curve and a crest curve, respectively, frame 600 naturally had a higher value than frame 824 in the RRFs analysis. Nevertheless, in the FE analysis, it was observed that data rate requirements at frame 824 were higher than at frame 600, indicating that the primary determinant of data rate values is vegetation density, which was slightly higher at frame 824.

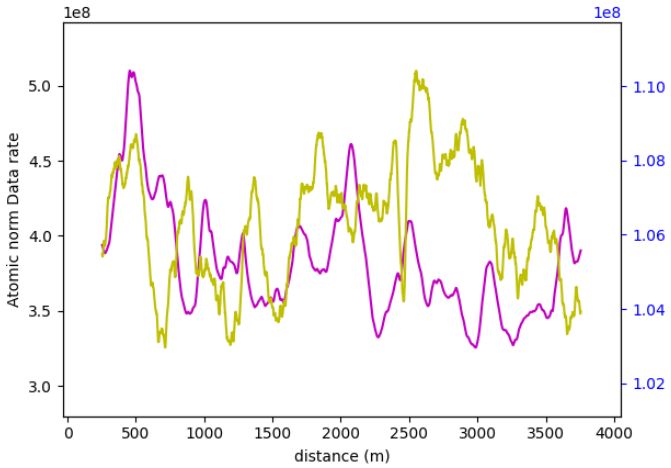

Figure 65. Road Section (1) FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

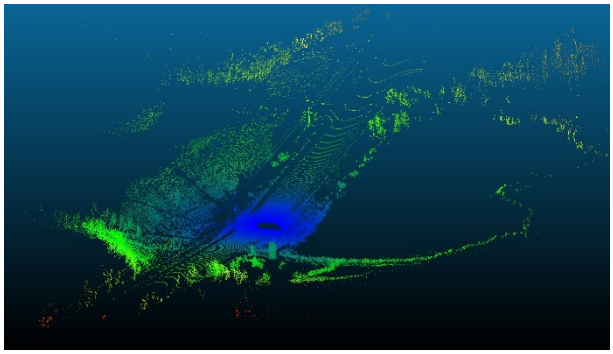

Figure 66. Road Section (1) - Frame 880 VISTA Output

Figure 67. Road Section (1) - Frame 880 Location on Google Maps

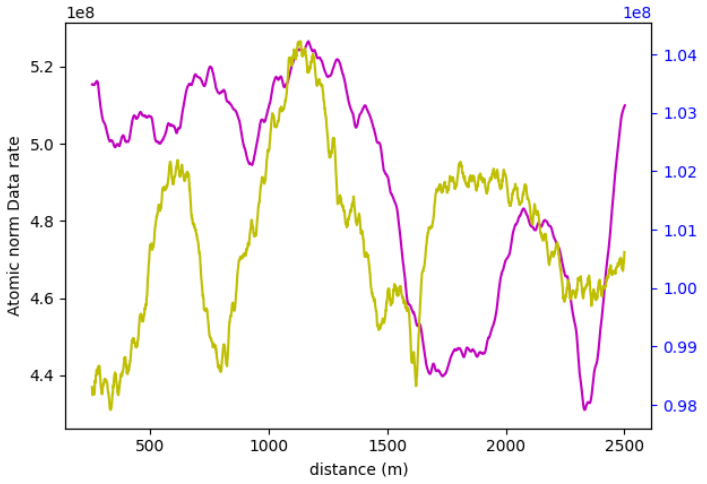

Figure 68. Road Section (2) FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

Figure 69. Road Section (2) – High Vegetation Density on Both Sides of Frame 1100

Figure 70. Road Section (2) – Low Vegetation Density on One Side of Frame 1600

Finally, in the third road section, similar observations were made as the first two, with data rate values being approximately five times higher. Regarding the influence of the vertical curve on the full environment (FE), frames 257 and 543 serve as prime examples to illustrate the limited impact of road geometry in the presence of roadside features. Frame 257, positioned on a crest curve, yielded lower data rate values than frame 543, situated on a sag curve in the RRFs analysis. Nevertheless, this pattern was reversed in the FE analysis since the vegetation density at frame 543 was the lowest across the entire road section, which resulted in recording the lowest data rate requirements. Figure 72 and Figure 73 show the difference in the vegetation density in frames 257 and 543.

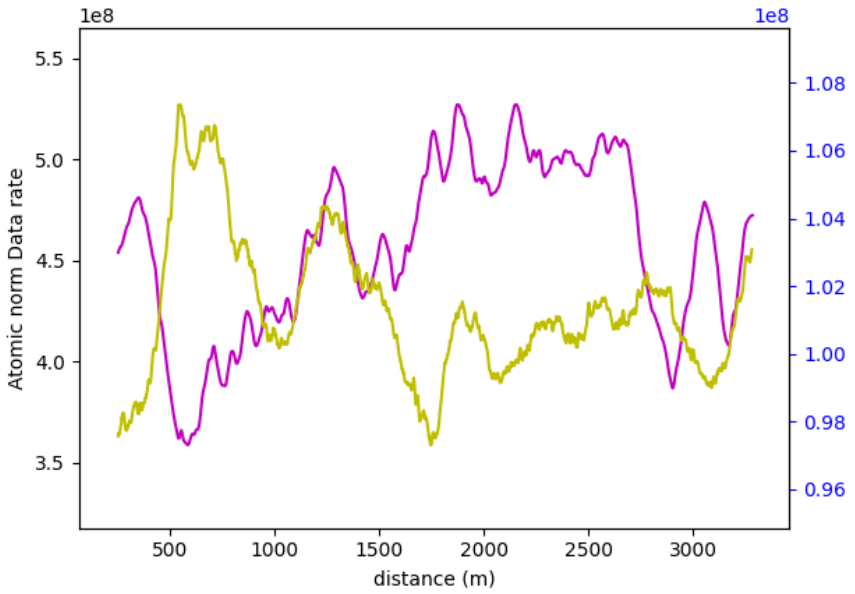

Figure 71. Road Section (3) FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

Figure 72. Road Section (3) – High Vegetation Density of Frame 257

Figure 73. Road Section (3) – Low Vegetation Density of Frame 543

4.2.3 Summary

LiDAR has limitations in its scanning operation regarding the vertical FOV. As mentioned previously, the distance between two consecutive LiDAR beams increases the further they travel away from the sensor owing to the resolutions of the azimuth and elevation angles, meaning that the number of points collected at the far end of the range will be significantly lower than those near the sensor. Additionally, owing to the 360° vision of the LiDAR sensor, at positive or negative grades, the forward vision of the sensor will have a different perspective than the backward vision. For example, suppose the vehicle is going up on a positive grade. In that case, the forward portion of the FOV will generally have a limited range because the points ahead are at higher elevations and will be more likely to occlude each other, unlike the backward portion, which will have more range due to the elevated position of the sensor relative to the road points below minimizing the occlusion problem. Typically, this balances out the total number of points captured by the sensor, whether the sensor is at an upgrade or a downgrade, and by extension, the data rate requirements would be similar.

A clear pattern of peaks and drops throughout the tested road sections was observed throughout the analysis of vertical curves, which closely resembles the inverse of the vertical geometry of the roads themselves. Frames found at the crests of the vertical curves have, on average, lower values than frames found at straight sections. On the other hand, frames at the centre of sag curves will have higher values than straight sections on average because, at sag locations, the LiDAR sensor can hit more points as the vertical angle of the sensor at this position

covers a higher portion of the road. On top of the crest curve, in addition to the self-occlusion of the road, it is expected that an increased number of the beams fired from the sensor would land beyond the sensor range, and hence, fewer points will be hit. Therefore, the sensor will have less point density in total.

However, using the occupancy method in the calculations, where all points are assigned equal weights regardless of their position, has resulted in minor differences between the values at the crest curves and sag curves because even though there might be occlusions that occlude some parts of the road, the number of occluded points from the LiDAR vision have low impact compared to the bulk of points within the visible short range of the sensor. Hence, their influence on changing the data rate values is weak. Table 5 compares the values at the different vertical curve types to those at the flat section of the road, where a maximum of 7% increase in values is recorded at sag curves. As for crest curves, a maximum of 4.2% decrease in values was recorded.

Table 5. Summary of Change in Data Rate Requirements for Vertical Curves RRFs Analysis.

Section	Min Value (Crest)	Average Value (Flat)	Max Value (Sag)	% Change Relative to Flat Values
1	$1.01x10^8$	$1.05x10^8$	$1.104x10^8$	-3.9% → +5.1%
2	$0.98x10^8$	$1.01x10^8$	$1.04x10^8$	-3% → +3%
3	$0.96x10^8$	$1.0x10^8$	$1.07x10^8$	-4.2% → +7%

Regarding the FE analysis, it was observed that the density of the roadside features dictates the data rate values. Areas with high vegetation density scored the highest data rate requirements and vice versa, masking the effects of the road geometry. From the perspective of AVs, processing environments with heavy vegetation to eliminate some hazards, such as wildlife collisions, would require them to process, on average, almost 4.8 times more data than they should if they only consider the RRFs for their navigation, as outlined in Table 6.

Table 6. Vertical Curves Calculations for FE and RRFs Analysis.

Section	RRFs Maximum Value	FE Maximum Value	% Increase
1	$1.1x10^8$	$5.1x10^8$	363%
2	$1.05x10^8$	$5.25x10^8$	400%
3	$1.08x10^8$	$5.25x10^8$	386%

4.3 ROADWAY WIDTH

The volume of data to be processed by the AV is subject to change, influenced by the surrounding environment and the road's geometry. One specific aspect of road geometry that warrants investigation is roadway width. Therefore, the primary research question for this study centres on how alterations in roadway width affect the data rate requirements, especially in areas where the number of lanes increases or decreases. Two separate analyzes addressed this issue: one encompassing the FE and another focusing solely on the RRFs. The influence of the variation in the roadway width will be assessed with and without the roadside elements, and expected changes in the data rate requirements will be quantified.

This study examined four distinct road sections, each with a length of 4 km, situated on the same highway, all mainly configured as two-way, two-lane roads with passing lanes. The average data rate requirements in the 2-lane zones will be compared to those in the 3-lane zones to quantify the extent of change in the data rate requirements between the two distinct zones.

4.3.1 Roadway Width (RRFs) Analysis

Road section (4) began with a standard two-lane configuration, and approximately halfway through this section, an extra lane was introduced in one direction. This change is vividly depicted in the data rate value graph, presented in Figure 80, which clearly illustrates an increase in data rate values as the AV enters the three-lane zone. In the two-lane zone, data rate values ranged from 1.26×10^8 to 1.35×10^8, with an average of approximately 1.30×10^8. In contrast, within the three-lane zone, data rate values ranged from 1.43×10^8 to 1.49×10^8, with an average value of roughly 1.46×10^8, indicating an average increase in data rate values of approximately 12.3%.

It's important to note that the graph displayed several abrupt drops in data rate values near the 500m and 2000m marks, associated with temporary loss of road information due to obstructing vehicles during data collection, as shown in Figure 84. These localized drops were not factored into any of the calculations.

Figure 74. Illustration of 3-Lane Zones from Google Maps

Figure 75. Illustration of 2-Lane Zones from Google Maps

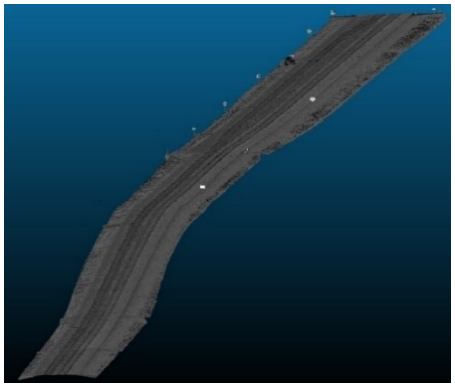

Figure 76. Road Section (4) Layout

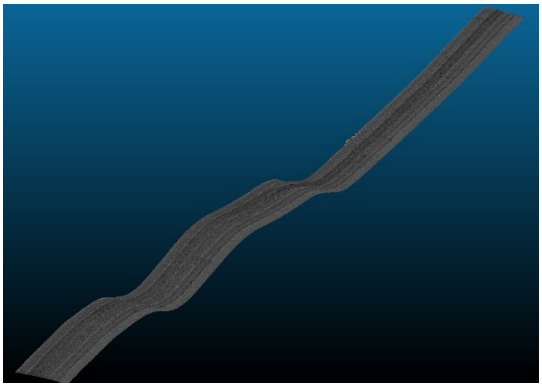

Figure 77. Road Section (5)

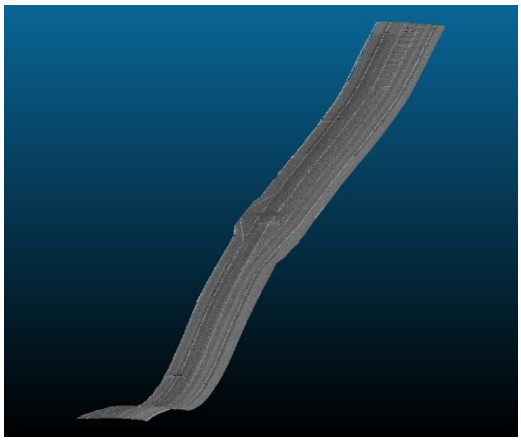

Figure 78. Road Section (6)

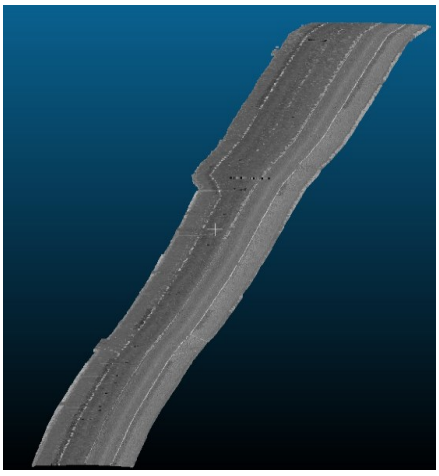

Figure 79. Road Section (7)

In road section (5), a notable drop in data rate values became evident around the 2.3 km mark, as seen in Figure 81. The segment preceding this point featured three lanes, while the region beyond this mark had just two lanes. The influence of a vertical curve in the three-lane zone was observed around the 1 km mark, where there was a decrease in the data rate values at the crest curve, followed by an increase at the sag of the vertical curve as the sensor's field of vision is augmented in this region and hence the data rate requirements increase.

Again, the data rate values within the three-lane zone ranged from 1.29×10^8 to 1.39×10^8, with an average of approximately 1.34×10^8. In contrast, data rate values within the two-lane zone were in the range of 1.11×10^8 to 1.20×10^8, averaging around 1.15×10^8. Consequently, this marked an approximate 16.5% value increase between the two zones.

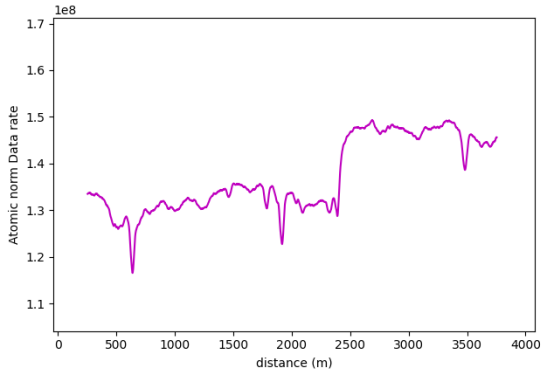

Figure 80. Data Rate Requirements for Road Section (4)

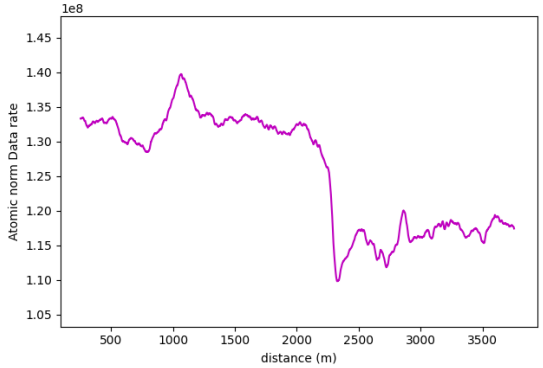

Figure 81. Data Rate Requirements for Road Section (5)

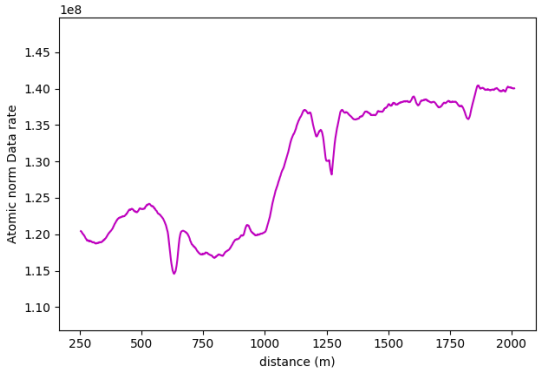

Figure 82. Data Rate Requirements for Road Section (6)

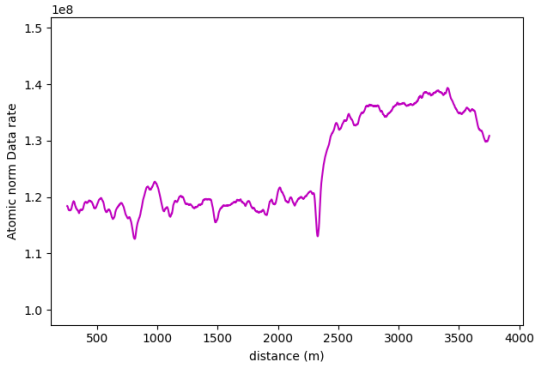

Figure 83. Data Rate Requirements for Road Section (7)

Similar patterns were observed in road section (6), which started with two lanes and ended with three. That is why the graphs showed that the second half of the data rate graphs was higher than the first. The average value at the 3-lane zone was equal to 1.38×10^8 while that at the 2-lane zone was equal to 1.20×10^8 resulting in an average 17.4% increase in the data rate values between the regions. It is worth noting that the drop observed around the 1.2 km mark, shown in Figure 85, was attributed to a big truck on the road occluding a part of the road pavement, resulting in an abrupt drop in the data rate values. Nevertheless, as mentioned previously, this region was not regarded in the calculations.

Finally, road section (7) exhibited the same pattern regarding the change in data rate values at the two separate zones. The average value at the 2-lane zone was 1.20×10^8 and that at the 3-lane zone was 1.37×10^8 meaning that there is an average 13.2% increase in data rate values.

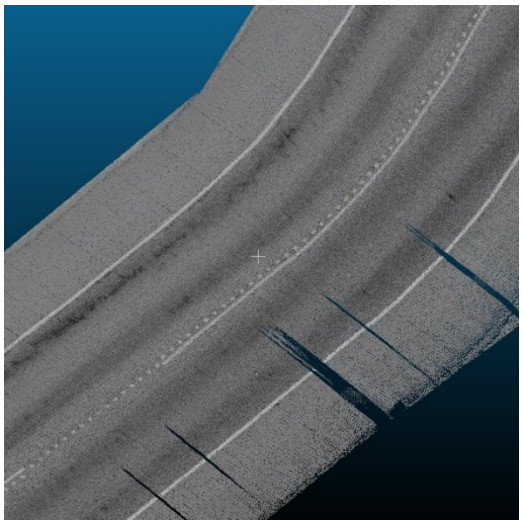

Figure 84. Gaps in Road Section (4)

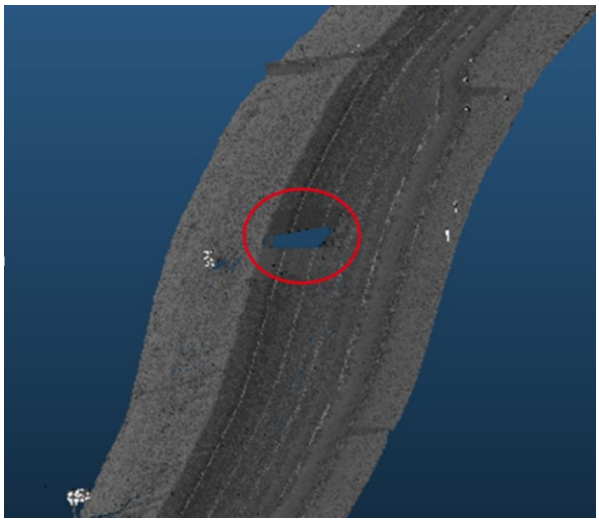

Figure 85. Void in The Pavement of Road Section (6)

4.3.2 Roadway Width (FE) Analysis

On analyzing the full road environment for the same four road sections, the first observation that was clear from incorporating the roadside features in the calculations is that the data rate requirements more than doubled. The scene becomes more complex for the AV to process in real time, albeit most of the points collected by the AV's sensor in its point cloud are irrelevant to its navigation operations.

Furthermore, the previously observed patterns when the roadway width increases had been masked by the addition of the roadside elements in the analysis, as witnessed in Figure 86. It was also observed in Figure 91, which compares the data rate requirements with and without the irrelevant roadside elements. Also, the roadside features have heavily influenced the data rate values. For example, in road section (4), the density of the point cloud was highest around the 1.7 km mark; therefore, the data rate requirements were maximum. On the other hand, around the 0.5 km mark, the low density of the point cloud prompted the data rate values to plummet. Figure 87 and Figure 89 demonstrate the difference in the environment at these designated regions, where it was observed that the surrounding environment has more elements like grass or buildings near the 1.7 km mark. Another observation was that the region that recorded the highest data rate values was composed of only two lanes, which again highlights the fact that with the addition of the roadside elements, the patterns seen in changing the width of the road become insignificant. Generally, an approximate increase of 128% was recorded when the FE was included in the analysis.

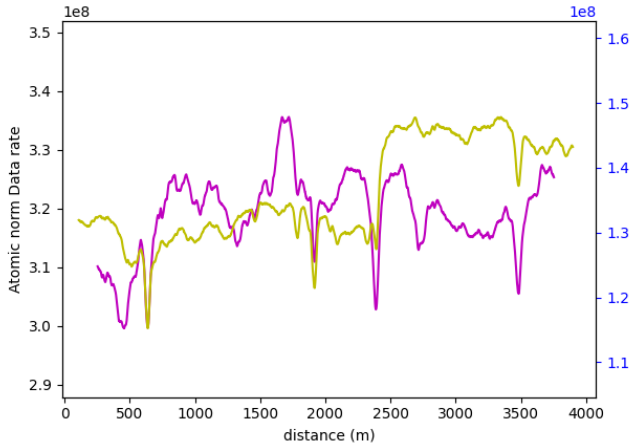

Figure 86. Section (4) FE Analysis (Purple + Left Axis) vs RRFs Analysis (Yellow + Right Axis)

The same inferences can be drawn from the analysis of the other road sections. Again, by investigating the impact of the roadside features on the second road section, it was readily apparent that the established patterns in the variation of the roadway width are nonexistent. Even more so, it was observed that the two-lane region, on average, had higher values than the rest of the road section.

Figure 87. Environment for Road Section (4) at the 0.5 km Mark

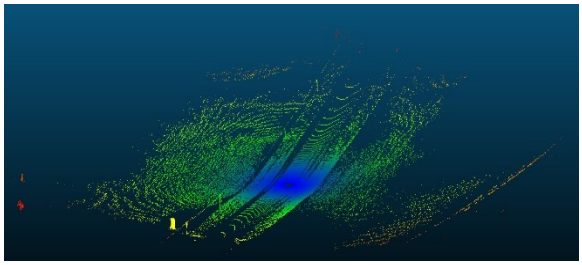

Figure 88. VISTA Output for Frame 500

Figure 89. Environment for Road Section (4) at the 1.7km Mark

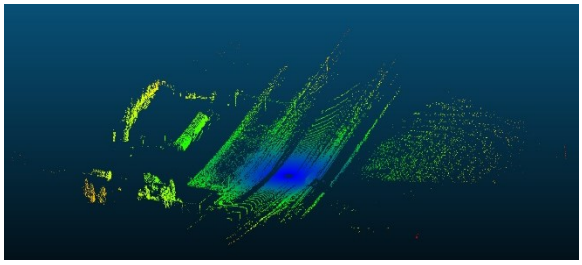

Figure 90. VISTA Output for Frame 1700

The magnification in the data rate requirements has been quantified to be around 145%. Moreover, the VISTA outputs show that the roadside elements chiefly control the calculated values using the whole environment. The hike in the data rate values around the 1 km mark, previously reported in the analysis of the relevant features, was exacerbated with the inclusion of roadside features, which coincidentally were high in density in this region. Additionally, the position of this region on a sag curve contributed to this jump in values. On the other hand, the VISTA output at frame 500 demonstrated the relatively lower point density of the surrounding environment, lowering the data rate requirements.

The impact of the roadside features in the analysis is summarised in Table 8. The expected increase in data rate requirements for four road sections included in this analysis ranged from 128% to 148%, demonstrating that for the AV to safely navigate this highway under favourable weather conditions and evade any possible WVC, it would need to process an extra amount of data that is more than twice the amount it would need for its primary navigation operations that rely on the relevant road features.

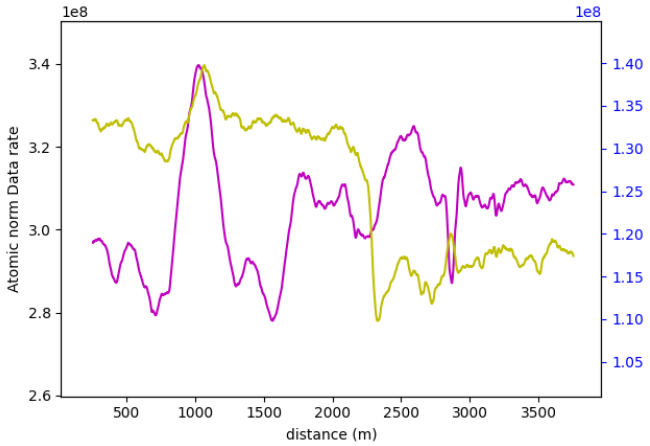

Figure 91. Road Section (5) FE Analysis (Purple + Left Axis) vs RRFs Analysis (Yellow + Right Axis)

Figure 92. Environment for Road Section (5) at the 0.5 km Mark

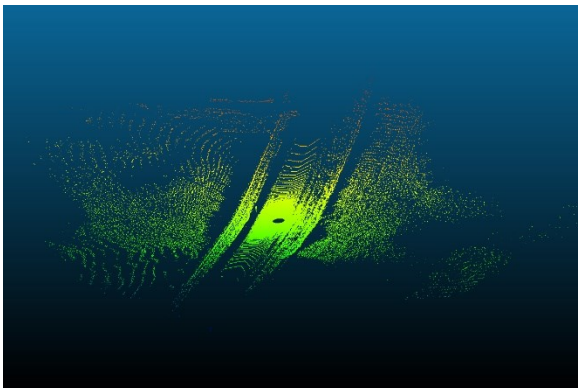

Figure 93. VISTA Output for Frame 500

Figure 94. Environment for Road Section (5) at the 1 km Mark

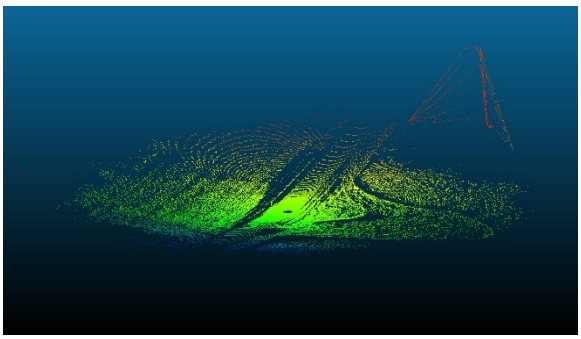

Figure 95. VISTA Output for Frame 1000

Summary

This analysis highlights the impact that the road's width has on the AVs' performance. More often than not, the general layout of the road geometry changes along any trip. In the RRFs analysis, it has been established that the addition of an extra lane would require the AV to process 12.3-16.5% additional data to maintain its safe operations.

Additionally, the FE analysis has revealed that including roadside features increases the complexity of the scene surrounding the AV due to increasing the total number of points that must be processed in real time by the onboard computer. For the four analyzed road sections, generally, the surrounding environment had low vegetation density, and as a result, the data rate requirements were seen to increase around 2.3 times only. Therefore, it would be easier for the AV to process its surroundings and less susceptible to wildlife collisions. Finally, the patterns established in the RRFs analysis on adding an extra lane had been masked in the FE analysis owing to the randomness and inconsistency of the vegetation throughout the analyzed road sections.

Table 7. Summary of RRFs Data Requirements for Roadway Width Analysis.

Section	2-Lane Zone			3-Lane Zone			Increase Percentage
	From	To	Average	From	To	Average	
4	$1.26x10^8$	$1.35x10^8$	$1.3x10^8$	$1.43x10^8$	$1.49x10^8$	$1.46x10^8$	12.3%
5	$1.10x10^8$	$1.2x10^8$	$1.15x10^8$	$1.29x10^8$	$1.39x10^8$	$1.34x10^8$	16.5%
6	$1.15x10^8$	$1.25x10^8$	$1.20x10^8$	$1.35x10^8$	$1.40x10^8$	$1.38x10^8$	15%
7	$1.15x10^8$	$1.25x10^7$	$1.20x10^8$	$1.34x10^8$	$1.39x10^8$	$1.37x10^8$	14.1%

Table 8. Roadway width Calculations for FE and RRFs analysis.

Road Section	RRFs Max Data Rate Values	FE Max Data Rate Values	% Increase
4	$1.49x10^8$	$3.4x10^8$	128%
5	$1.39x10^8$	$3.4x10^8$	145%
6	$1.42x10^8$	$3.3x10^8$	132%
7	$1.39x10^8$	$3.45x10^8$	148%

4.4 HORIZONTAL CURVES

The following analysis delves into the crucial aspect of horizontal curves and their influence on the data rate requirements of AVs while considering their surroundings with a focus on the full environment around AVs, including challenging off-road conditions characterized by buildings, vegetation, and mountains. The presence of sight obstructions can lead to critical situations at these curves, affecting sight distance and causing occlusion, which, in turn, creates an area where there is a high loss of information for AVs and ultimately leads to a decline in data rate requirements. A modified analysis is conducted in the attached Appendix where the roadside features were kept only on the side of the road that caused the occlusions to attenuate the influence of roadside features.

Horizontal curves play an essential role in AVs' navigation and safety. They require careful deliberation and planning, including speed, steering, and data processing adaptation to ensure a smooth and secure driving experience. Understanding how horizontal curves affect data rate requirements is paramount for enhancing the performance and reliability of autonomous driving systems. The analysis encompasses the complete surrounding environment around AVs. Unlike controlled environments, real-world scenarios often present challenges like buildings, dense vegetation, and towering mountains. These features obstruct sightlines, rendering horizontal curves even more critical for human drivers and AVs. Therefore, it is necessary to account for such obstructions to derive accurate conclusions regarding the impact of horizontal curves on data rate requirements because AVs rely heavily on sensor data and real time processing to navigate these curves safely. When sight distance is compromised, AVs experience occlusion, leading to a loss

of crucial information, such as road information, potential obstacles, and the presence of other vehicles.

Three road sections comprising critical horizontal curves and situated on a mountainous terrain were considered for this analysis. Although the presence of the vegetation and other offroad elements in the analysis are expected to have random variations on the data rate values as their density is inconsistent and unevenly distributed along the road, the effect of the critical horizontal curves was clearly evident, primarily due to the anticipated occlusions at these locations.

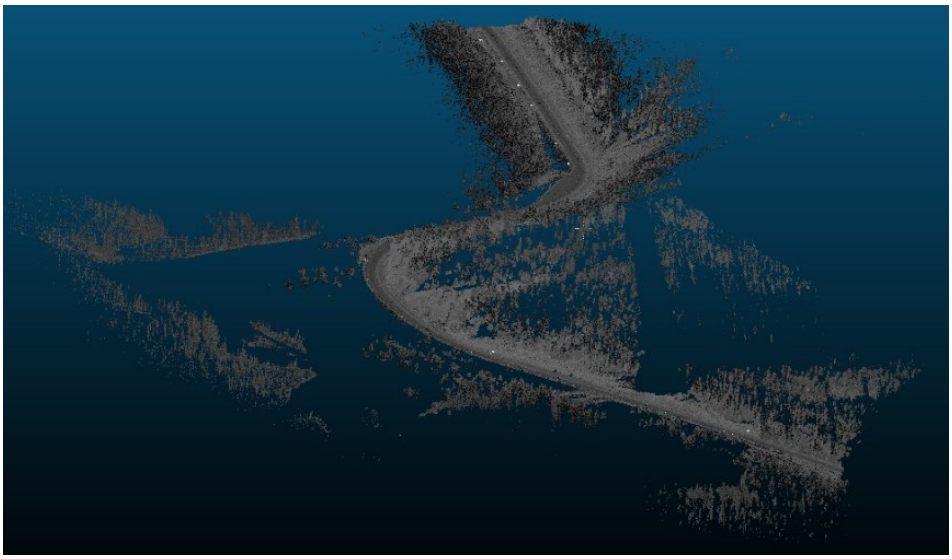

Figure 96. Road Section (8) Layout

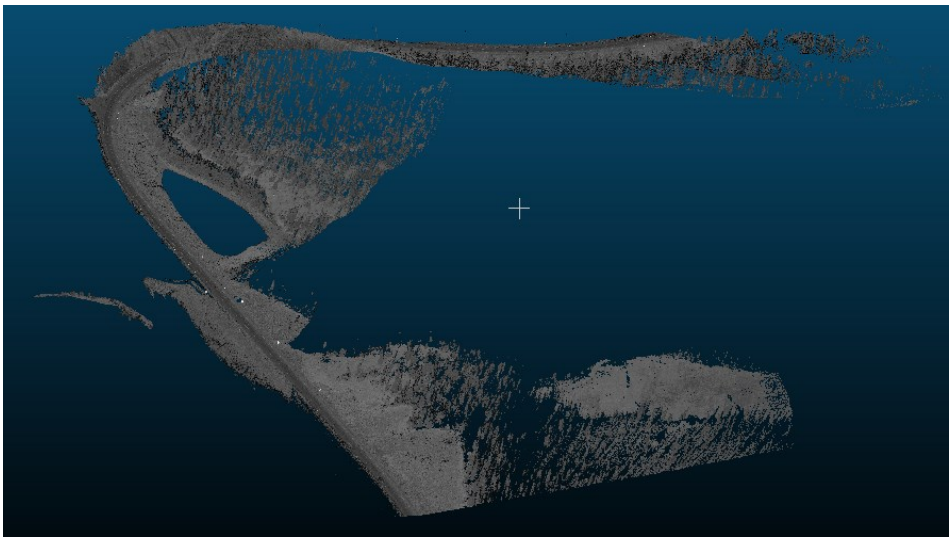

Figure 97. Road Section (9) Layout

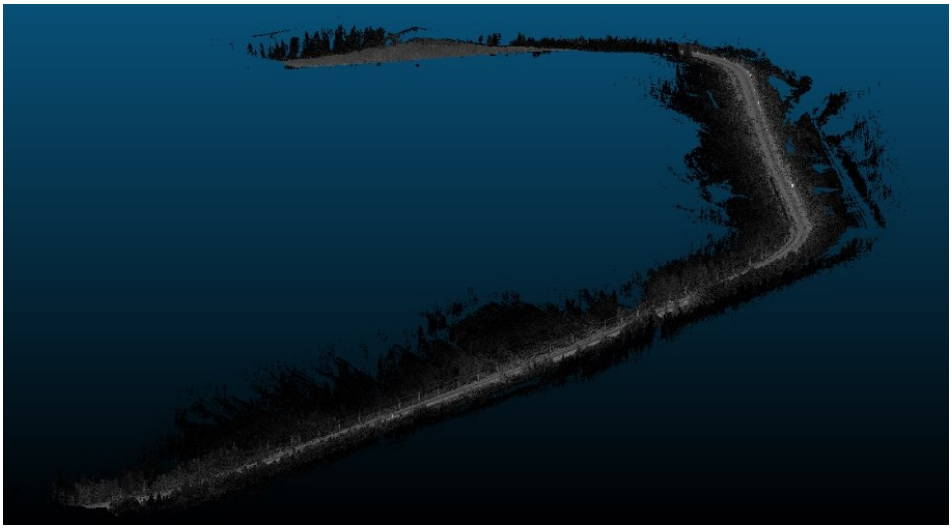

Figure 98. Road Section (10) Layout

The average data rate values from the areas close to the critical curve were calculated and compared to those at the horizontal curve locations as they would have a comparable layout. In order to quantify the influence of the horizontal curves on the data rate values, the analysis focused on the entrance and the exit of the curve. The regions just before and right after the curves, when compared to the entrance and exit of the curve, respectively, will exhibit an evident change in patterns of the data rates and will showcase the influence of occlusion on the vehicle's performance. If the roadside elements' density increases, a further rise in the data rate values at that location will be expected inside the curve itself.

4.4.1 Horizontal Curves Analysis

By examining the layout of road section (8), it was clear that the left side of the road had few roadside elements since much of that side was occupied by a waterbody, reflected as a clear space in the input point cloud. Consequently, low data rate values were recorded along the first kilometre of the road. More importantly, the data rate values started to decrease near the entrance of the horizontal curve as the occlusion started to affect the forward vision of the vehicle by hiding some road information from the sights of the vehicle. At frame 795, it was observed that only 120 m of the road ahead was visible when compared to the 180 of the backward vision. This disparity was responsible for the drop in the data rate values by 16% as the values went from $2.48x10^8\ bits/s$ to $2.13x10^8\ bits/s$. Once the vehicle was inside the curve, the vehicle could capture more of the

roadside elements, which in this case are mountains on both sides of the road, as seen in Figure 107. A slight increase in values was observed.

At the exit of the curve, the data rate values were seen to increase again because as the vehicle moves away from the curve, occlusion's influence on the sensor's backward vision dissipates, and the vehicle regains its operating vision capacity. Around frame 952, located near the curve's exit, the sensor's backward vision was limited to about 110m, while the unhindered forward vision was approximately 175m. From the graphs, the data rate values just after the curve, where the general layout of the environment was similar to that at the curve, were $2.43x10^8\ bits/s$ while the value at the exit of the curve was $2.11x10^8\ bits/s$, which means there is about a 15% drop in data rate values at this position.

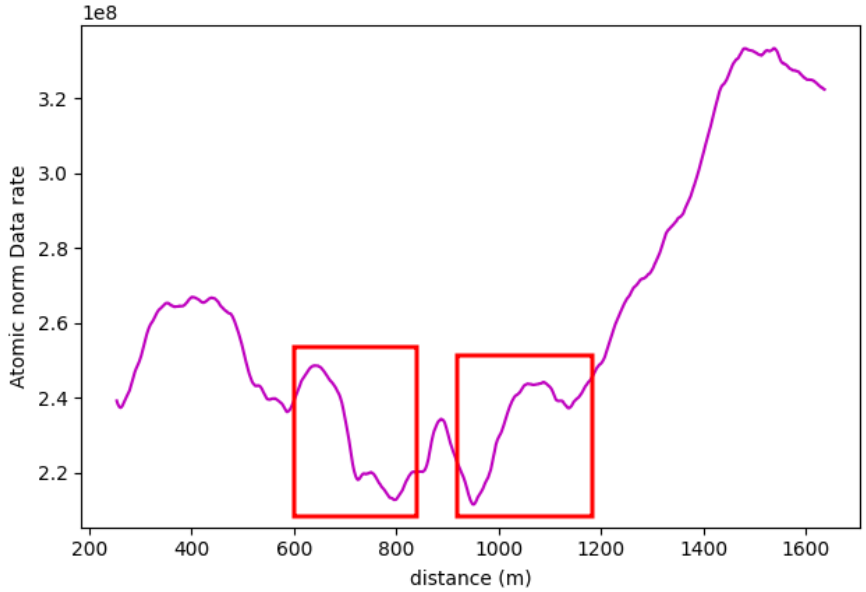

Figure 99. Data Rate Requirements for Road Section (8)

Figure 100. Water Body Along Section (8) Left Side

Figure 101. Road Section (8) - Frame 795 Location on Google Maps

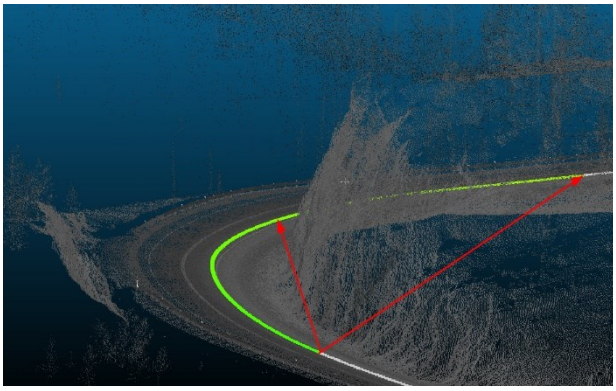

Figure 102. Road Section (8) - Frame 795 (Forward Vision Hindered)

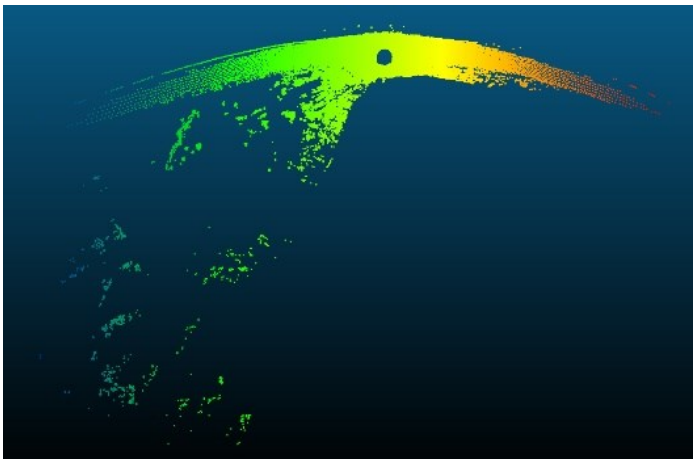

Figure 103. Road Section (8) - Frame 795 VISTA Outputs

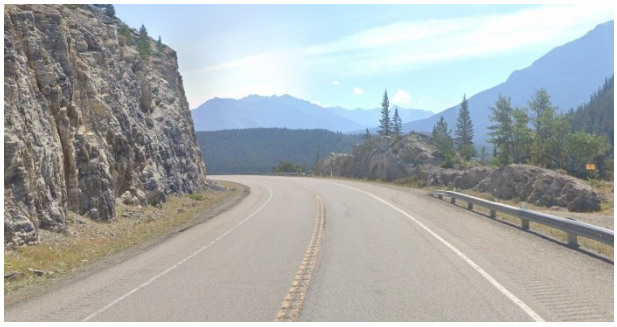

Figure 104. Road Section (8) - Frame 952 (Behind the Sensor) Location on Google Maps

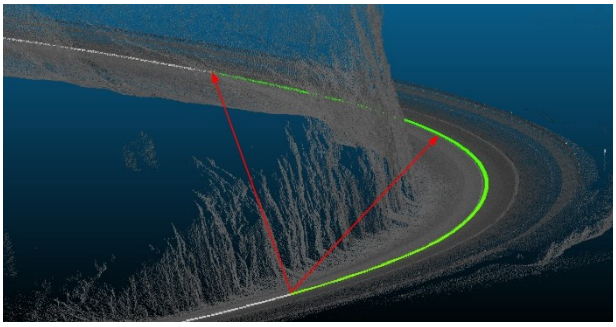

Figure 105. Road Section (8) - Frame 952 (Backward Vision Hindered)

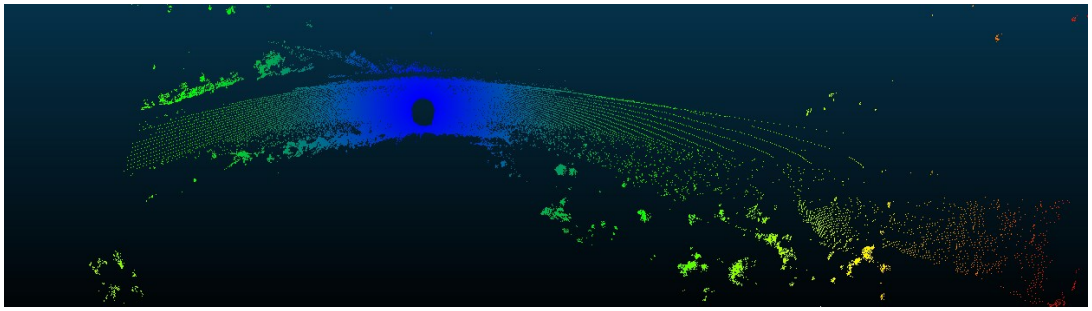

Figure 106. Road Section (8) - Frame 952 VISTA Outputs

Figure 107. Road Section (8) – Surrounding Environment Inside the Hz Curve

Similar to road section (8), which had water bodies covering some parts of it, section (9) also had water bodies covering regions on both sides of its road, as seen in Figure 112, explaining the gaps within the input point cloud. The observed pattern around the critical Hz curves was seen again when road section (9) was analyzed. This section's data rate requirements were highest just before the horizontal curve's entrance. However, the roadside mountain occluded forward vision once the vehicle entered the horizontal curve. The VISTA outputs showed that around 142m of the road ahead was visible. On the other hand, the vehicle could see around 200m behind it. This was translated into a drop in the data rate values from $2.75x10^8\ bits/s$ to $2.45x10^8\ bits/s$ which is equivalent to 12.2%. The data rate values increased slightly and stabilized for a few meters because, within the horizontal curve itself, the vehicle could see more of the mountains on both sides of the road, which naturally have high point densities. However, as the vehicle headed towards the exit, the data rate values started to decrease again and dropped to their lowest at frame 1156, where the backward vision was affected by occlusion, and the vehicle could see only 145m of the road behind it, whereas, in front of it, it could see around 210m. By comparing the data rate

values just after the horizontal curve and at its exit a drop of 12.7% was observed as the data rate values decreased from $2.32x10^8\ bits/s$ and $2.06x10^8\ bits/s$.

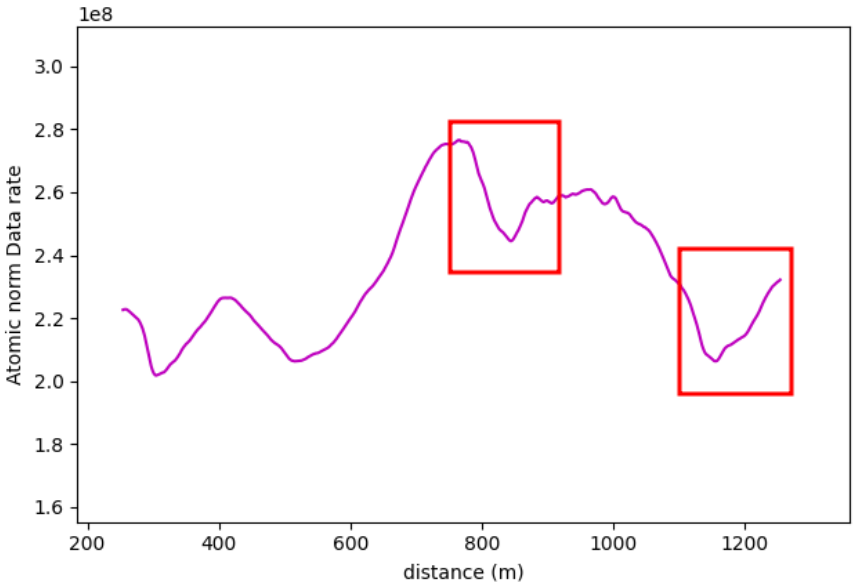

Figure 108. Data Rate Requirements for Road Section (9)

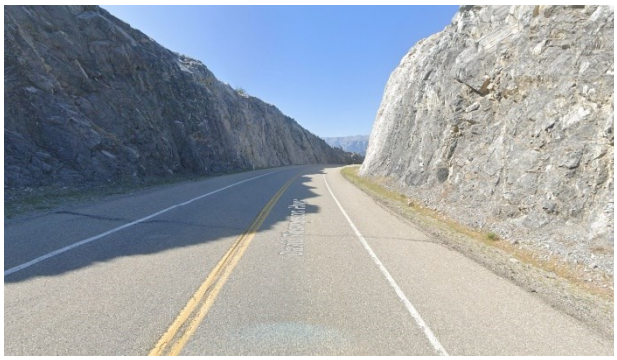

Figure 109. Road Section (9) - Frame 844 Location on Google Maps

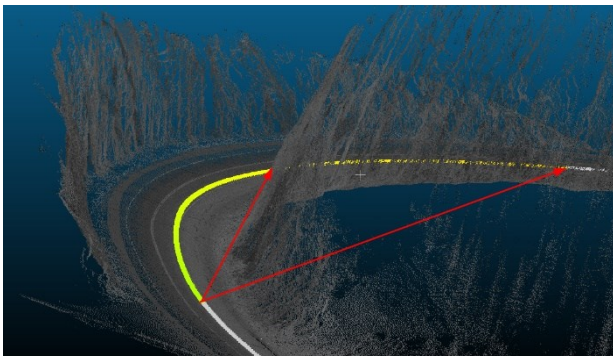

Figure 110. Road Section (9) - Frame 844 (Forward Vision Hindered)

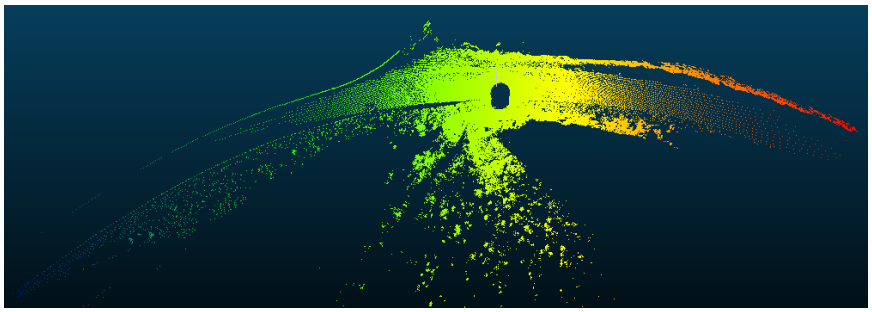

Figure 111. Road Section (9) - Frame 844 VISTA Outputs

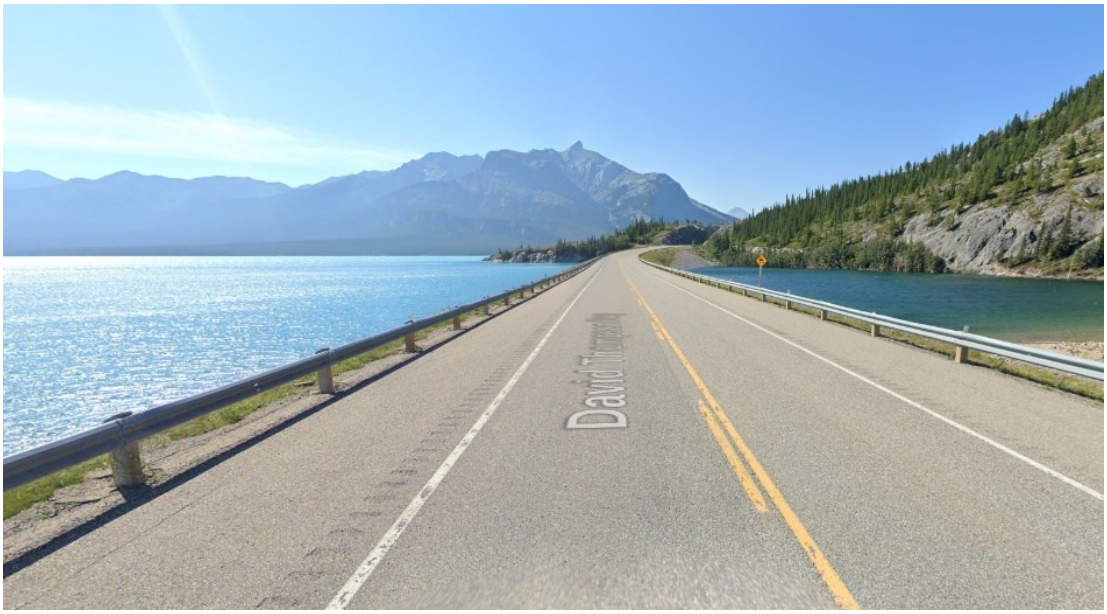

Figure 112. Waterbody Along Road Section (9)

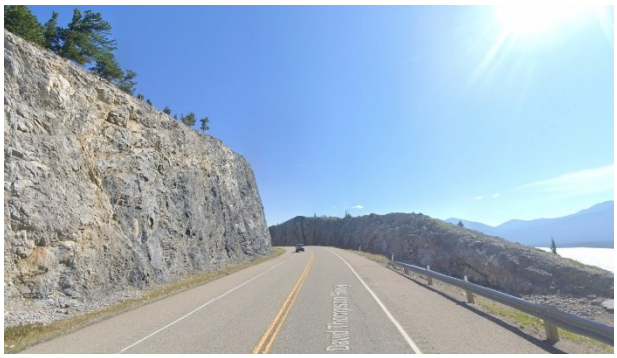

Figure 113. Road Section (9) - Frame 1156 (Behind the Sensor) Location on Google Maps

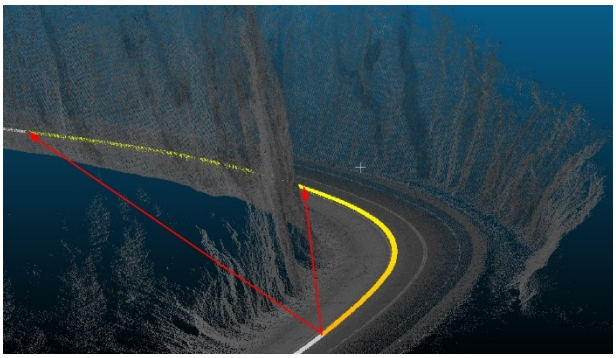

Figure 114. Road Section (9) - Frame 1156 (Backward Vision Hindered)

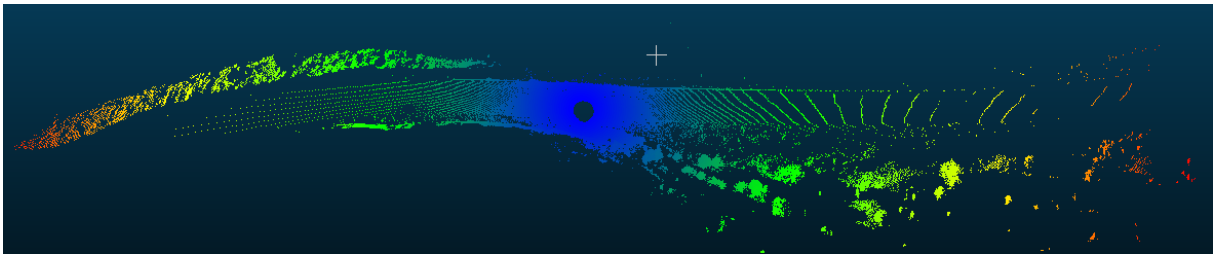

Figure 115. Road Section (9) - Frame 1156 VISTA Outputs

Figure 116. Road Section (9) - Environment Inside the Hz Curve

The last section comprised two horizontal curves, one at the start of the road section and the other near the end. The second horizontal curve was identified as critical due to roadside sight obstructions. Following a similar pattern as the previous two sections around the critical horizontal curve, at the entrance of the curve, the data rate values were seen to drop from $3.7x10^8\ bits/s$ to $3.2x10^8\ bits/s$. While at the exit of the curve, the data rate values jumped from $3.7x10^8\ bits/s$ to $4.3x10^8\ bits/s$. Inside the curve, the roadside element density increased as the vegetation turned into a rocky slope, so an increase in the data rate values at the centre of the curve was observed. At the end of the road section, a water body replaced the vegetation on one side, resulting in a sudden drop in the data rate near the section's edge. The drop seen at the entrance of the curve was around 15.6%, whereas the increase in the data rate values at the exit of the curve was approximately 16.2%. Again, the regions used for comparison are found in immediate proximity before or after the curve since the overall environment would be very similar.

It is also worth noting that the first horizontal curve encountered in this road section did not exhibit the established pattern purely because it was not as tight as the rest of the analyzed curves. As seen in the data rate graph of Figure 117, the values were stable at the first horizontal curve location around the 2.2 km mark. No variations were observed, meaning that on safe horizontal curves with no roadside obstructions, the change in the vegetation density solely controls the data rate values. There is little to no influence of occlusion.

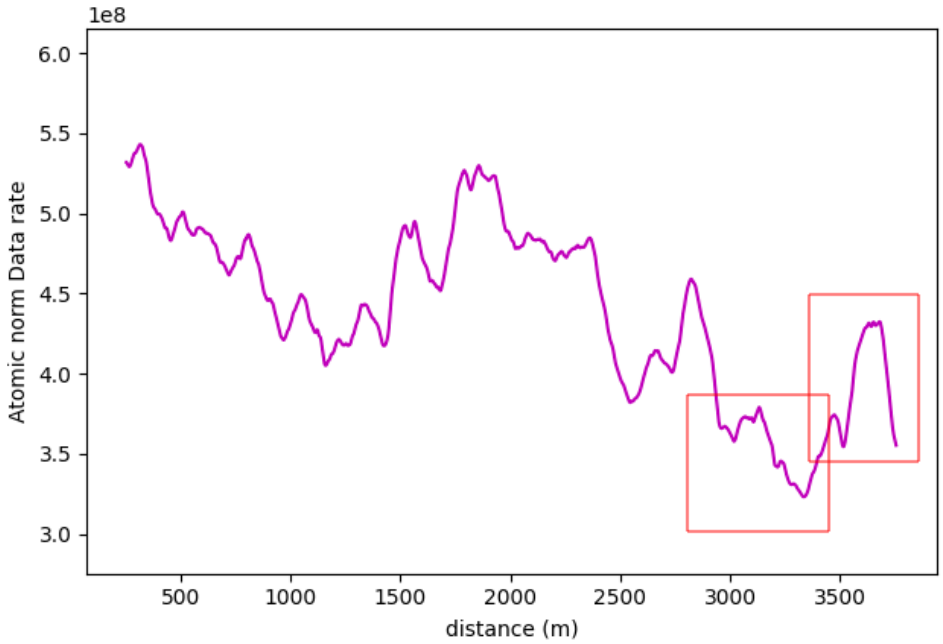

Figure 117. Data Rate Requirements for Road Section (10)

Figure 118. Road Section (10) - Frame 3336 Location on Google Maps

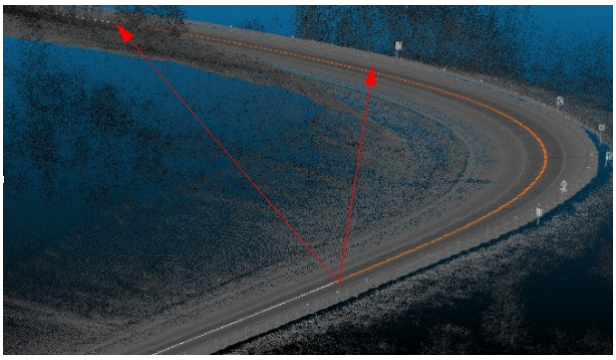

Figure 119. Road Section (10) - Frame 3336 (Forward Vision Hindered)

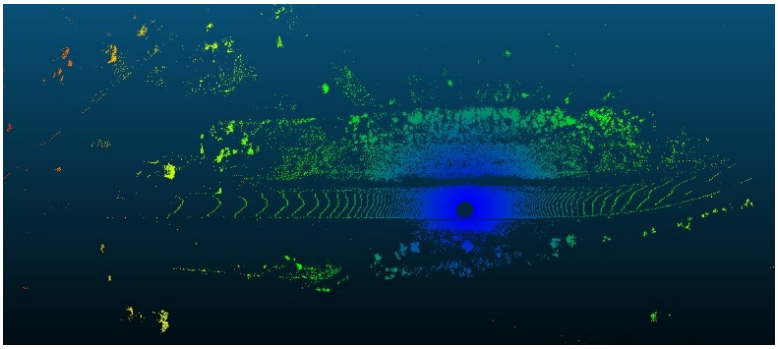

Figure 120. Road Section (10) - Frame 3336 VISTA Outputs

Figure 121. Road Section (10) - Frame 3516 (Behind the Sensor) Location on Google Maps

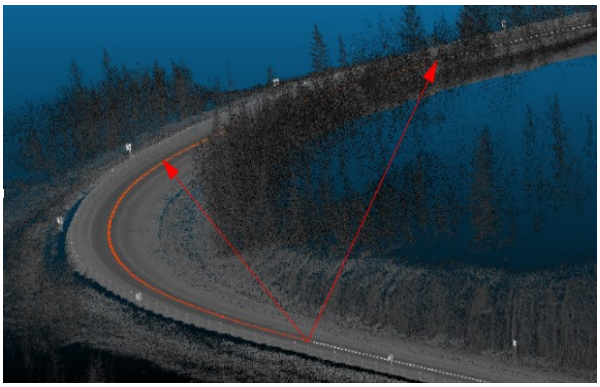

Figure 122. Road Section (10) - Frame 3516 (Backward Vision Hindered)

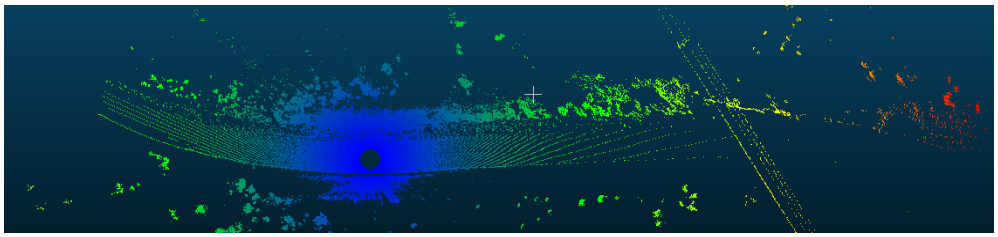

Figure 123. Road Section (10) - Frame 3516 VISTA Outputs

4.4.2 Summary

This analysis emphasizes the expected influence of horizontal curve occlusions on AV operations through dynamic environments of the same mountainous nature where the layout of the surroundings is different for each section. Generally, a decrease in the data rate requirements was observed owing to occlusions at horizontal curves, which limit the AV's vision and line of sight, both forwards and backward, indicating the loss of valuable information that could be critical for its safe navigation. The average loss of information for the three-road section, as listed in Table 9 was 12.2-15.6% at the entrance of the curve and 12.7-19% at the exit.

The mountainous terrain has proved to be challenging for the AV owing to their critical horizontal curves. The loss of information occurring at these locations can put it at risk. Not knowing what is around the corner can be detrimental to the AV’s operations. The drop in data rate values can represent the percentage of road and surrounding information for the AVs, especially at the entrance of the horizontal curves, where the forward vision is hindered. For instance, wildlife incidents will be threatening at these locations. It is expected that the AV will not have enough time to react and avoid these collisions.

Table 9. Summary For Change in Data Rate Values at Horizontal Curves.

Section	Hz Curve Entrance			Hz Curve Exit		
	Min Value	Max Value	Average Drop	Min Value	Max Value	Average drop
8	$2.1x10^8$	$2.5x10^8$	14%	$2.1x10^8$	$2.4x10^8$	19%
9	$2.45x10^8$	$2.75x10^8$	12.2%	$2.06x10^8$	$2.32x10^8$	12.7%
10	$3.2x10^8$	$3.7x10^8$	15.6%	$3.7x10^8$	$4.3x10^8$	16.2%

Table 10. Horizontal Curves Calculations for FE and RRFs analysis.

Road Section	RRFs Max Data Rate Values	FE Max Data Rate Values	% Increase
8	$1.1x10^8$	$3.3x10^8$	200%
9	$1.15x10^8$	$2.86x10^8$	149%
10	$1.12x10^8$	$5.4x10^8$	382%

4.5 SENSITIVITY ANALYSIS

This section explores how the change in different parameters of the data rates equation can influence the results. To that end, different weather conditions will be modelled by changing the signal-to-noise ratio variable in the equation. An additional sensor will also be simulated, and the results will be compared to the primary sensor used in this analysis.

4.5.1 Signal-to-Noise Ratio

This study aims to quantify the extent to which data rate requirements are affected by changes in the SNR value. As previously mentioned, the SNR delineates the ratio between the power of the reflected signal and the ambient noise. Favourable weather conditions are maintained throughout

the analysis. In this section, an additional heavy rain environment will also be modelled to showcase how much of an impact it would have on the calculations. The relationship between the $SNR_{max\,range}$ parameter and data rate requirements are inversely proportional, as equation (2) outlines. Hence, for instance, a 10% increase in the $SNR_{max\,range}$ value will be reflected as a similar 10% decrease in the data rate requirements.

Following the observations of Zhang et al. [206] and knowing that the sensor's max range is 245m, the SNR is expected to decrease by approximately 8.5 dB. Therefore, the new $SNR_{max\,range}$ will be 12-8.5 = 3.5 dB. Figure 124 shows the data rate values for the same section on the same graph. Through the analysis of road section (1) with the full environment, it can be noticed that using an $SNR_{max\,range}$ value of 3.5 dB to simulate the effects of heavy rain, the data rate values have increased by approximately 3.4 times. Worse weather can impose conditions where the signal power might be less than the ambient noise. Assuming that the noise is ten times stronger than the signal, the $SNR_{max\,range}$ will be 0.1 according to equation (3). Consequently, the data rate requirements are expected to drastically increase by 120 times. At such elevated rates, the performance of the AV is bound to be unstable as it would not be able to process such monumental data rates in real time.

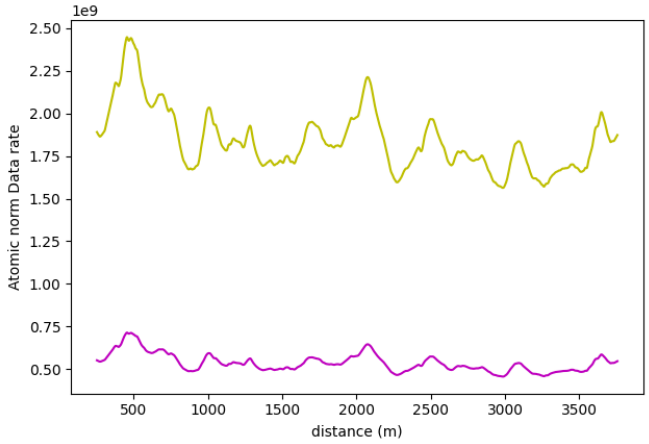

Figure 124. Road Section (1) FE Analysis. SNR=3.5 dB (Yellow Curve) Vs SNR = 12 dB (Purple Curve)

4.5.2 Alternative LiDAR Sensor

To demonstrate how variations in LiDAR sensor specifications can impact data rate calculations and affect the performance of AVs, the Velodyne-32E LiDAR will be considered an illustrative example [216]. The specifications for this sensor are detailed in Table 11. Notably, this sensor, also known as (HDL-32E) has a shorter range, limited to 100 meters, a distinct vertical field of view (FOV) distribution, and reduced vertical resolution. However, it has a slightly better range precision and matches the horizontal angle precision of the VLS-128.

The reduced vertical angular resolution of the HDL-32E LiDAR will significantly decrease the number of data points it can capture. This limitation arises because the gap between consecutive laser beams widens as the vertical angle increases. The shorter range also reduces the total number of captured data points. Consequently, the scenes captured by this sensor will generally be less complex for the onboard computer to process, containing a lower volume of information than the VLS-128 model, evident when comparing the data rate requirements for both sensors for the same road sections. In road section (2), for example, the maximum drop in the data rate requirements has been quantified as approximately 620%. This dramatic drop in data rate values means that the scenes captured by the HDL-32E LiDAR will be much less detailed than those of the VLS-128 sensor. This means that the AV might not be able to fully comprehend its environment from the scarce information it receives from the sensor and this in turn can increase the safety risks.

For comparison, the analysis will keep the default padding region of 245 meters instead of using 100 meters to focus on the same areas of interest as those considered in the VLS-128 model calculations, facilitating a direct comparison between the two LiDAR sensors.

Table 11. LiDAR Sensors Specifications.

Sensor	R	θ_h	θ_l	ϕ_h	ϕ_l	δ_R	δ_ϕ	δ_θ	F
VLS-128	245m	180°	-180°	15	-25°	0.03m	0.11°	0.11°	20hz
HDL-32E	100m	180°	-180°	10.7°	-30.7°	0.02m	1.33°	0.11°	20hz

4.5.2.1 HDL-32E Vertical Curves Analysis

As seen from Figure 125 and Figure 126 for road section (2), the established vertical curve pattern was missing from the RRFs and FE results, due in large part to the short range of the sensor, where it did not seem to be impacted by occlusion from the road geometry. In addition, the number of captured points in the far range of the sensor is expected to be significantly lesser than those captured in the long range by the VLS-128 sensor. Hence, the impact will be reduced even moreso on the occluded part of the road.

The VISTA outputs for frame 1135 are presented to highlight the disparity in scene perception by the AV using various sensors. This particular frame was selected because it yielded the highest data rate value when using the VLS-128 sensor, serving as a prime illustration of the contrast between the sensor outputs. Figure 127 and Figure 128 managed to showcase the disparity between the range of the captured points and, more importantly, the density of point cloud captured by the two sensors where the VLS-128 sensor, undoubtedly, carries much more information for the AV and hence will exert bigger burdens on the onboard computer.

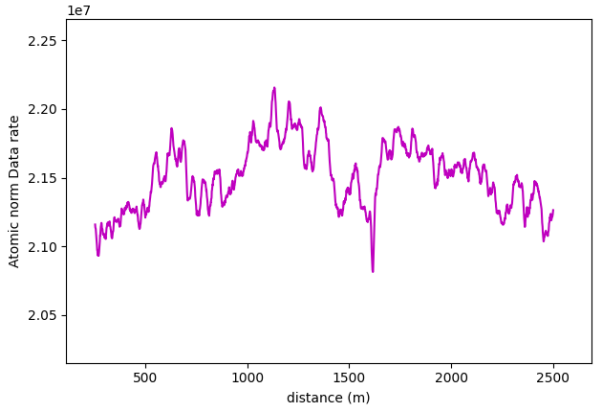

Figure 125. HDL-32E RRFs Data Rate Requirements for Road Section (2)

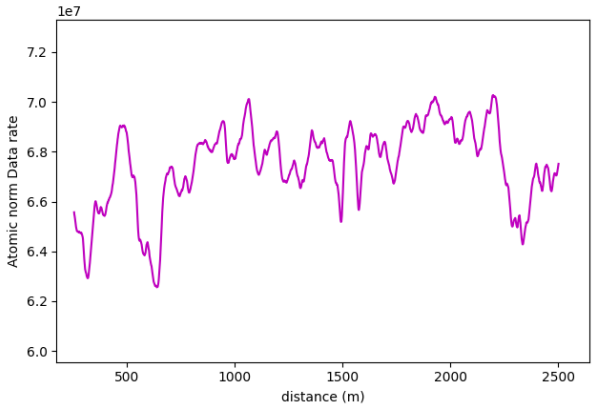

Figure 126. HDL-32E FE Data Rate Requirements for Road Section (2)

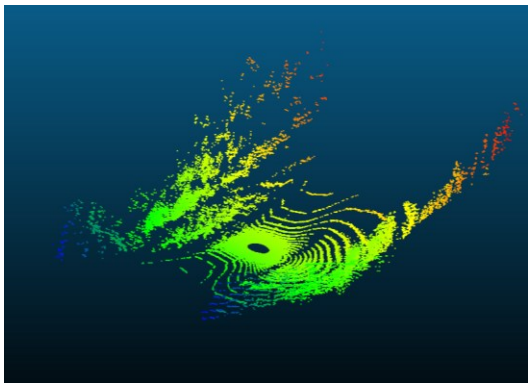

Figure 127. Road Section (2) Frame 1135 FE VISTA Outputs for HDL-32E

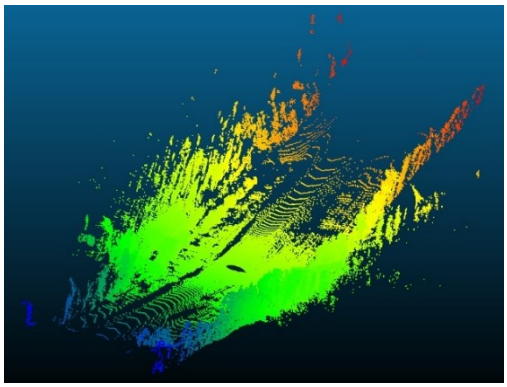

Figure 128. Road Section (2) Frame 1135 FE VISTA Outputs for VLS-128

4.5.2.2 HDL-32E Roadway Width Analysis

Through the analysis of road section (5), using the occupancy method calculations with the HDL-32E sensor, the same established pattern between the two-lane and three-lane zones is manifested. The average value in the two-lane zone is $3.55x10^7 bits/s$ whereas that of the three-lane zone is $4.01x10^7 bits/s$, thereby, there is a 12.9% change in the data rate values. However, the VLS-128 sensor's version affects a 17.4% change in values, as discussed in section 4.3.1, showing that even for short-ranged and less accurate sensors, the onboard computer will experience a surge in data rate requirements when the roadway width increases.

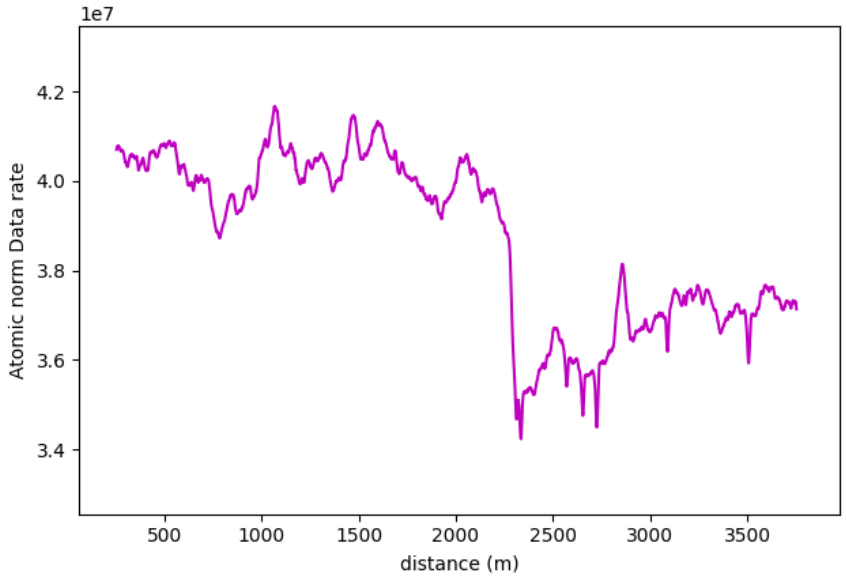

Figure 129. HDL-32E RRFs Data Rate Requirements for Road Section (5)

4.5.2.3 HDL-32E Horizontal Curve Analysis

This analysis employs the same configuration previously applied to road section (9), as detailed in 4.4. Notably, the HDL-32E sensor, as illustrated in Figure 131, exhibited a distinct pattern due to its limited range. It did not encounter any horizontal curve occlusions, and as a result, it did not experience a reduction in data rate requirements upon entering the horizontal curve (at frame 844). In contrast, this location saw a discernible increase in data rate requirements. Upon exiting the curve, an increase in values was naturally expected, but this elevation was primarily attributed to changes in roadside vegetation density.

The shifts in data rate values will be compared to showcase the difference between the HDL-32E and the VLS-128 sensor at horizontal curves. The VLS-128 sensor demonstrated a 28% increase in values, whereas the HDL-32E sensor witnessed a more moderate 15% increase. This observation highlights that while the surrounding roadside features undeniably influence changes in data rate requirements, horizontal curve occlusion also significantly shapes data rate patterns. In summary, the shifts in data rate values in the HDL-32E sensor's results are predominantly driven by changes in roadside features, whereas in the VLS-128 sensor's results, both roadside features and horizontal occlusions are pivotal factors influencing the data rate values.

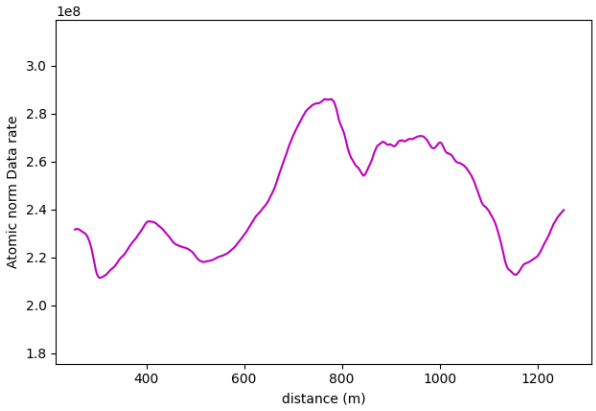

Figure 130. VLS-128 Road Section (9) Data Rate Requirements

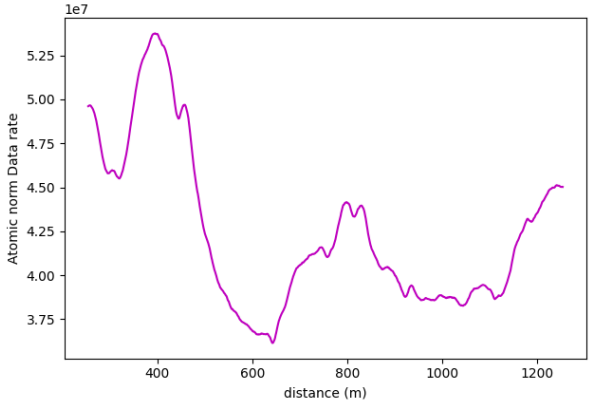

Figure 131. HDL-32E Road Section (9) Data Rate Requirements

5 DISCUSSION

Within this study's comprehensive analysis, encompassing diverse road sections set in multiple environmental contexts and varying sensor specifications, a central conclusion emerges: reporting absolute data rate requirements for static environments is exceedingly challenging. This complexity arises from many influential factors, including the quality and density of the point cloud data employed in simulations, the specifications of the utilized sensors, variables related to road terrain and occlusions, and, not to be overlooked, the prevailing weather conditions. Additionally, it is worth noting that the chosen calculation methodology, whether based on occupancy or volumetric considerations, can significantly impact the resulting data rate requirements. Consequently, the primary goal of this research is to establish patterns and trends for the different static road environment elements and study their impact on AV performance in terms of expected computational demands.

On the subject of the determined data rate values, they exhibited an intriguing characteristic where their correlation with the complexity of the surrounding environment is not as straightforward. Recognizing that an increased data rate requirement does not necessarily signify a more intricate environment is essential. To illustrate this phenomenon's counterintuitive nature, consider locations at vertical crest curves or horizontal occlusions. At first glance, their decreased data rate requirements might suggest a reduced level of complexity. However, in reality, these locations present challenges for AV operations. In such instances, it becomes apparent that the term "decrease in data rate requirements" does not fully encapsulate the situation's complexity. Instead, assessing the percentage of missing information provides a more suitable and accurate descriptor. Consequently, this approach extends beyond identifying locations with high data rate requirements. Also, locations with a notable and unexpected decline in data rate values were prioritized. These locations represent critical points in the landscape for AV operations.

Considering the prevailing literature aiming to assess the complexity of the environment surrounding AVs, it is evident that a prevailing narrative suggests that the closer an element is to the AV, the more challenging the situation becomes for the onboard computer. In alignment with this narrative, the occupancy method has been established as the primary calculation approach in this research, as it effectively mirrors this perspective. Nevertheless, it is important to note that, for long-range detections, an alternative viewpoint, represented by the volumetric method, also

exists. This perspective proposes the opposite scenario, contending that, from an operational standpoint, objects situated at a greater distance from the AV pose heightened challenges for detection. Consequently, they demand increased computational resources from the AV.

5.1 POINT CLOUD DENSITY

Broadly speaking, assessing the quality of input LiDAR point cloud data involves measuring its density, which serves as a key determinant of how closely it mirrors the actual physical environment. A higher point cloud density lends greater authenticity to virtual simulations. The density of the point cloud is directly influenced by the capabilities of the LiDAR sensor used for scanning the physical surroundings. The more advanced and precise the LiDAR sensor, the greater the point cloud density it can achieve. Consequently, data rate values are intricately linked to the density of the employed point clouds, with higher point cloud densities necessitating increased data rate requirements. Furthermore, as highlighted in Section 4.1, there is an evident fluctuation in point cloud density. This variation manifests as minor, indiscernible fluctuations in data rate requirements, primarily because points directly beneath the scanning vehicle exhibit higher density due to their proximity to the sensor. Accordingly, if simulations are conducted in an alternate lane, it is essential to recognize that the results will similarly undergo alterations.

5.2 SPECIFICATIONS OF THE USED SENSOR

Environment perception for AVs is carried out using onboard LiDAR sensors, which are essential for its navigation operations owing to their accurate object detection and tested reliability [217] [85] [86] [87]. Hence, it is used as the primary sensor in this virtual simulation. The AV will perceive its surroundings differently depending on the capabilities of the LiDAR sensor used in the simulations.

The primary sensor specifications that affect the AV's ability to view the environment are 1- Maximum range, 2- Range precision, 3- Field of view, and 4- angular resolution. Typically, longer-range sensors can capture more of the surrounding environment and will be subject to higher data rate requirements. Similarly, higher range precisions will result in more dense representations of the environment, increasing the data rate requirements. On a similar note, the bigger the field of view and the finer the angular resolutions will augment the perception of the AV, increasing the

point cloud density and, ultimately, raising the data rate values. An exercise was carried out in detail in Section 4.5.2, which compares two different LiDAR sensors and demonstrates how the AV's performance would change with changing the equipped sensor.

5.3 VOLUMETRIC APPROACH

The volumetric method is unique in its approach and exploits the spherical coordinate system used for voxelization. This method assigns more weight to objects that have larger volumes. Following the voxelization process implemented in this research, the voxel size increases the further its position is from the sensor. Hence, the biggest voxels will be located at the maximum range of the used sensor. In calculating the percentage of occupied voxels by volume to get the delta, voxels of bigger size will have more influence on the results, meaning that this method directs its attention to the distant points rather than the points near the sensor.

This approach assumes that the further the point is from the sensor, the harder it would be to process in real time due to the scarcity of points at long ranges. Additionally, since the points close to the sensor are already big in numbers, it would be easier to identify objects from them. Consequently, they will not have much processing power from the onboard computer.

Regarding the analysis of the vertical curves, the volumetric method managed to heighten the effects of the occlusions in both the RFFs and FE road layouts, while maintaining the same patterns. The percentage of increase in data rates at the sag curve was 42%, 20%, and 18.8%, respectively. In comparison, the decrease at crest locations was -40%, -18.7%, and -22.5%, respectively, for the three road sections. Similarly, in the FE analysis, the rates were +34%, +15%, and +29%, respectively, at the sag locations and -51%, -15.8, and -39.5%, respectively, at the crest locations. The higher grades of road section one increased the disparity between values on the curves and the flat road sections of the road, as evident in the recorded change percentage in the data rate values. Finally, on comparing the data rate requirements between the FE and RRFs analysis in the volumetric method, a gigantic increase was observed due to the presence of roadside features, which prompted the percentages to reach 1300%, 1950%, and 1475%, respectively, among the three road sections.

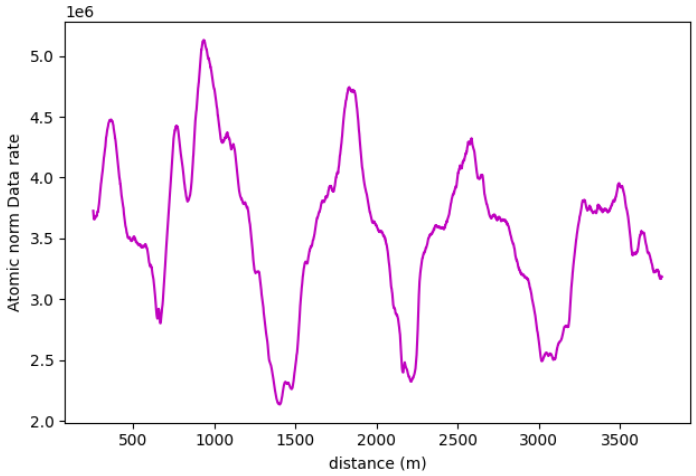

Figure 132. RRFs Data Rate Requirements for Section (1) Using Volumetric Method

The previously established pattern found in the occupancy method calculations was again observed in the roadway width analysis. It was seen to an even greater extent in both the RRFs and FE layouts. In the RRFs, the 3-lane zones exhibited a 20% increase in the data rate requirements. In contrast, the increase percentage was 30% in the FE analysis, with the visible influence of vegetation on the results. For the road sections involved in the analysis, the increase in data rate requirements between the RRFs and FE analysis ranged between 600-700%, significantly lower than that recorded in the vertical curve road sections, due to the overall lower vegetation density witnessed in these sections.

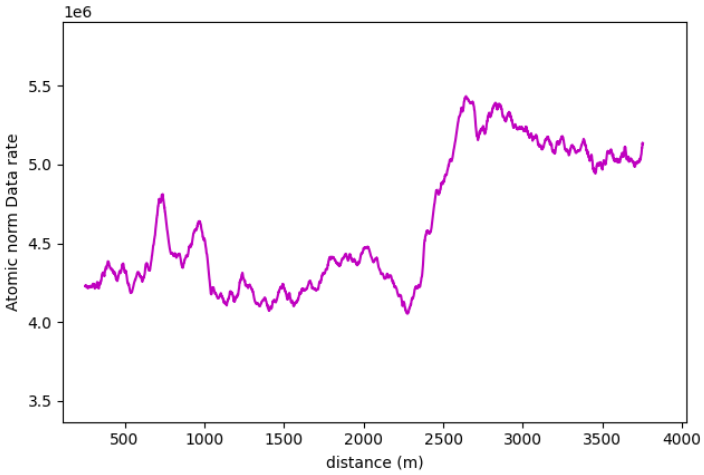

Figure 133. RRFs Data Rate Requirements for Road Section (4) Using Volumetric Method

Lastly, in the horizontal curve analysis of the modified sections, naturally, the pattern persisted but with much more significant variations in the values owing to the presence of the

roadside elements. The drop in data rate requirements reached 491% and 367% at the exit and the entrance of the curves of road sections (8-9), respectively.

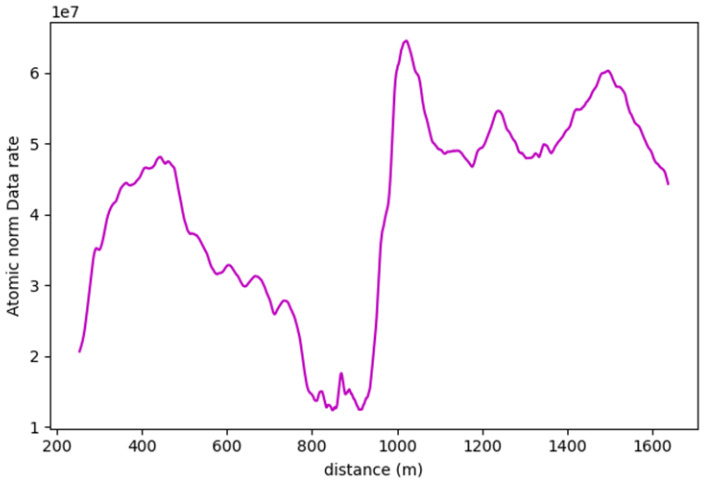

Figure 134. Data Rate Requirements for Road Section (8) Using Volumetric Method

As evident from the results of the volumetric method calculations, it is more compatible with RRFs road layouts rather than FE ones. Since it is sensitive to changes in the point densities and point distributions, the presence severely affects its results. Generally, the AV will be more concerned with far road points than it would with the vegetation found on its periphery. Hence, using this method to explain the AVs' performance will be unsuitable if the offroad elements are included in the analysis. It, nevertheless, provides robust results in the RRFs analysis as it explains, in the vertical curves, for example, the loss of road information at crest curves due to the limited range of vision or the increase in data rate requirements at sag locations owing to the augmented vision at this location. The horizontal curve analysis is a prime example of the shortcomings of the volumetric method, where it reports exaggerated percentages of loss of road information owing to the enormous influence of roadside elements. As a result, the volumetric method is not well-suited for the analysis of the horizontal curve. The attached Appendix contains an in-depth analysis of the volumetric method.

5.4 Full Environment Analysis

The objective of conducting a full environment (FE) analysis in this study is to quantify how roadside features influence the performance of AVs. Although these features may not typically affect standard AV driving operations, they remain within the scope of the vehicle's sensors, and

the onboard computer must process them for scenarios such as Wildlife Vehicle Collisions (WVCs). For instance, detecting animals concealed within roadside vegetation demands a comprehensive environmental assessment, which requires the AV to undertake more intricate computational tasks. A useful illustration to gauge the additional computational load on the onboard computer involves comparing the necessary data rates between FE and relevant road features (RRFs) analyzes, in contrast to the default computations primarily focused on driving lanes and road shoulders in case of emergency parking or manoeuvres. This comparison serves as an effective means to quantify the increased computational demands placed on the AV's onboard system.

Since the road environments are dynamic, multiple road environments with different roadside layouts have been analyzed. Rural environments with open fields and little vegetation scored the lowest data rate requirements. An example of such environments is road sections (4-7), which scored a mean value of $3.0x10^8\ bits/s$ and a maximum value of $3.3x10^8\ bits/s$. On comparing the FE to the RRFs analysis for these sections to isolate the influence of the roadside features, it was found that they increased the data requirements by an average of 138%. On the other hand, rural environments that have higher vegetation densities, such as road sections (1-3), averaged data rate values of $4.47x10^8\ bits/s$ with a maximum value of $5.25x10^8\ bits/s$. In this case, the roadside elements contributed an average 380% increase in the data rate values. Finally, two of the three studied mountainous terrain road sections (8 and 9) were situated in locations close to bodies of water, which lowered the data rate requirements considerably for most of these road sections. As a result, the average data rate requirements were found to be $2.6x10^8\ bits/s$ and had a maximum value of $3.3x10^8\ bits/s$. As for road section (10), the absence of water bodies and the presence of mountains and vegetation alongside this road prompted it to exhibit the highest data rate requirements with a maximum value of $5.4x10^8\ bits/s$.

Wildlife detection could, therefore, become a monumental task for AVs, seeing that the data rate requirements can increase fivefold. Additionally, animals are often occluded from the view of the AV, especially in rural areas with heavy vegetation [218]. One countermeasure known for its effectiveness in addressing this problem is the installation of animal fencing [66]. However, it has some drawbacks. It is naturally infeasible to fence entire road networks due to financial constraints. Additionally, they are often installed at locations with high mortality rates for wildlife.

Finally, they require frequent maintenance [219] [220]. Another proposed countermeasure is to install roadside LiDAR that, among many tasks, can monitor and detect animals [221]. However, covering an entire road network with roadside LiDAR is impractical, as it will be exceedingly costly.

5.5 STATISTICAL TESTING

To validate that the variation in the data rate requirements around the curves and at locations of varying roadway width was not due to chance or from random variables within the environment a statistical 2-tailed T-Test has been conducted. For such analysis, the null hypothesis is that the mean data rate values do not change at vertical curves, horizontal curves, or varying roadway width locations. Whereas the alternative hypothesis is that there is a variation in the data rate values at these locations. First, the t-statistic is calculated and by knowing the degrees of freedom and significance level, the critical p-value can be estimated. The null hypothesis will be rejected if the calculated critical p-Value is less than the adopted significance level.

Rejecting the null hypothesis for the vertical curve analysis would require the variations in the values between the flat and curved locations to be statistically significant on the 99% confidence level. By conducting this analysis on road section (2), the calculated p-values for crest vs. flat and sag vs. flat scenarios were below the 1% significance level. Consequently, the null hypothesis was dismissed, indicating that there are indeed statistically significant variations in data rate values at the vertical curve locations.

As for horizontal curves, road section (7) was segmented into straight sections and portions corresponding to the horizontal curves. Upon analysis, the observed *p*-value was again below the significance threshold, leading to rejecting the null hypothesis. This result indicates a 99% confidence level that the average values at the horizontal curve significantly differ from those outside, thus underscoring the impact of horizontal curve obstructions, distinct from the surrounding environment, on both data rate requirements and the vehicle's overall performance. Similar results were again seen in the test conducted on the roadway analysis between the 2-lane and 3-lane zones.

Table 12. T-Test for Vertical Curves Analysis.

Location	Crest	Flat	Sag
Number of Samples	2015	2074	1115
Mean Value	$9.70x10^7$	$9.72x10^7$	$9.74x10^7$
Standard Deviation	$1.67x10^6$	$1.42x10^6$	$1.38x10^6$
Absolute *t*-statistic	3.2		9.3
p-Value	$1.3x10^{-3}$		0

Table 13. T-Test for Horizontal Curves Analysis.

Location	Straight	Curved
Number of Samples	1001	243
Mean Value	$2.15x10^8$	$1.94x10^8$
Standard Deviation	$2.06x10^7$	$5.95x10^6$
Absolute *t*-statistic	27.8	
p-Value	0	

Table 14. T-Test for Roadway Width Analysis.

Location	2 Lanes	3 Lanes
Number of Samples	1480	2031
Mean Value	$1.16x10^8$	$1.32x10^8$
Standard Deviation	$2.11x10^6$	$2.54x10^6$
Absolute *t*-statistic	204	
p-Value	0	

6 CONCLUSION

6.1 SUMMARY

Integrating AVs is a pivotal advancement in the ever-evolving transportation landscape, promising unparalleled safety, efficiency, and accessibility. The potential to revolutionize travel for disabled individuals and optimize traffic systems is profound. However, the path to widespread AV deployment is marked by a critical challenge: ensuring public confidence in their safety. Documented collisions involving AVs have highlighted the imperative need for rigorous testing before large-scale deployment. Virtual simulations emerge as a vital solution to meet this demand, offering a secure environment for extensive AV testing with the added benefit of cost-effectiveness. By leveraging virtual environments, researchers and developers can achieve significant milestones in enhancing the safety of AVs by conducting simulations equivalent to billions of real-world driving kilometres, which is a prerequisite to achieving the desired safety standards.

Understanding the interactions between the AVs and the dynamic physical environment is paramount for ensuring their safety. This study delves into the static surroundings of AVs, aiming to unravel complex scenarios they might encounter. Typically, AVs disregard off-road areas beyond the road shoulders, as these regions do not disrupt their navigational abilities or emergency stops. These irrelevant areas are filtered out from sensor data to streamline processing and reduce computational loads, allowing the vehicle to concentrate on pertinent road features. However, particular challenges, such as Wildlife Vehicle Collisions (WVCs), demand a holistic approach. Detecting and avoiding animal collisions requires processing the entire environment and ensuring operational safety. Therefore, the analysis is conducted twice: first, by incorporating all roadside features, and second, by excluding them. This dual approach provides a comprehensive understanding of the individual impact of each element on the environment's complexity. It also sheds light on the additional computational burdens incurred by processing the entire environment, offering invaluable insights into optimizing AV operations.

This research has developed a framework leveraging LiDAR point cloud data from various highways in Alberta, Canada, to create a virtual testing environment. Specifically, this analysis covered 34 kilometres of diverse highway sections, concentrating on two-way-two-lane rural

roads. These sections were chosen strategically to represent varied terrains, including flat, rolling, and mountainous landscapes, aiming to model a range of driving environments.

This proposed framework delves into the complexity of the surrounding environment by estimating the real time data processing requirements of AVs to maintain their navigational integrity under different weather conditions and driving environments. Other LiDAR sensors were also simulated to understand their impact on the calculations. The core principle guiding this analysis is the understanding that the complexity of the environment is directly related to the volume of data that AVs must process. This complexity can be captured by first voxelating the point cloud data surrounding the AV and then calculating the percentage of occupied voxels using the parameter delta (Δ). To ensure accuracy and reliability, the captured scenes are replicated from the vehicle's perspective along its trajectory through its sensors using the VISTA simulator.

Assessing the complexity of the environment is done through two distinct approaches in calculating (Δ). The first method prioritizes objects close to AVs and employs the occupancy method calculations. In contrast, following the volumetric method, the second approach assigns greater weight to distant objects. This weighting simulates the challenge of processing data from these remote objects instead of the closer ones, which generally have higher point densities and are easier to process.

6.2 MAJOR FINDINGS

In an effort to comprehend the intricacies of AVs' performance when engaging with the static environment and their safety implications, extensive research was conducted. This investigation involved studying variations in data rate requirements across diverse road environments and weather conditions. Numerous pivotal findings were uncovered by pursuing several objectives outlined below.

Vertical Curves

The impact of the vertical curves was investigated by analyzing three different road sections with a total of 11 km of rolling terrain. Vertical curve occlusion was found to play a role in dictating the complexity of the surroundings through the variations found in the calculated data rate requirements. A consistent pattern was evident across all analyzed sections which was amplified

when utilizing the volumetric approach. The complexity of the surroundings at sag locations can increase by up to 7%. Similarly, crest locations indicate increased complexity, quantified to reach 4.2%, as AVs face challenges due to the loss of road information in these areas.

Horizontal Curves

Critical horizontal curves with tight radii are susceptible to occlusion, especially when off-road elements like vegetation encroach into the road's right of way. While analyzing three distinct road sections containing four horizontal curves, three were identified as critical due to sensor vision impairment caused by off-road obstructions. Results demonstrated a substantial contribution to horizontal curve occlusion, with a drop in data rate values ranging between 21-29% using the occupancy method for the modified version of the analysis, while 12-19% was recorded in the original analysis.

Roadway width

The analysis of four distinct road sections, spanning 14.5 km and featuring frequent transitions between 2 and 3 lanes, unveiled a noteworthy trend: the scene surrounding the AV became notably more complex in the 3-lane zones. According to occupancy method calculations, there was an average increase in data rate requirements ranging from 12.3% to 16.5%. The volumetric method, conversely, reported a slightly higher average of 20%.

LiDAR Sensor Specifications

When utilizing less advanced sensors characterized by limited range and lower angular resolution, it was observed that, unlike their sophisticated counterparts, these sensors did not follow the established trends at both vertical and horizontal curves, as occlusions from road geometry had minimal impact on their outputs. Furthermore, the direct comparison between the data rate values of the same road section using different sensor models revealed that the outputs of the less advanced were seven times lower than those of the more advanced sensor. Such analysis concluded that equipping a sensor with inferior specifications will greatly hinder the environment perception capabilities of the AV which can jeopardize its operations and increase the safety risks.

Weather Conditions

Irrespective of the inherent complexity of the surrounding environment, weather conditions present unique challenges to the operations of AVs. Through the modelling of different signal-to-noise ratios, operating in favourable weather conditions has been shown, from the analysis results, to alleviate the burdens on AVs by decreasing the data rate requirements. For instance, when heavy rain is simulated by reducing the signal-to-noise ratio from 12 dB to 3.5 dB, data rate requirements surge by 3.4 times, underscoring the profound influence of weather on AV performance. Moreover, when noise levels exceed the signal strength by tenfold, data rate requirements skyrocket by 120 folds, representing an astronomical surge in the data that necessitates processing by the onboard computer.

Roadside Features

The omnipresence of roadside features through almost all of the driving environments necessitates the study of their impact on the performance of the AV. Especially when a problem like WVCs has to be addressed. The added complexity accompanying a higher volume of offroad elements has been quantified through different road environments. The average increase in data rate requirements in environments with low vegetation density was in the range of 140%. Whereas in high vegetation density areas, the increase was calculated to be in the ballpark of 400%. Hence, confined space areas such as tunnels are expected to be less complex for AVs than open-space environments.

Statistical Testing

The conducted T-test significantly bolsters the validity and reinforces the analysis's credibility. It demonstrates, with a 99% confidence level, that data rate requirement values at vertical curve locations differ significantly from those at flat locations. Similarly, the average data rate values between horizontal curve areas contrast starkly with those in straight areas. While external factors such as slight variations in point density, road width, or other road features might influence the results, the occlusion issue, whether due to vertical or horizontal curves, distinctly impacts the change in data rate requirements for AVs. Consequently, it also affects the complexity of the vehicle's surroundings. Likewise, the change in road width is directly related to increasing the scene complexity.

6.3 RESEARCH CONTRIBUTIONS

The primary contribution of this work is the development of a framework that leverages LiDAR point cloud data to create a virtual environment tailored for AV simulations where the quantification of real time data processing requirements under standard driving conditions is estimated by capturing the environment's dynamic nature. This investigation aimed to assess the complexity of the surrounding physical environment, thereby shedding light on scenarios presenting higher difficulty levels for AVs. By addressing this question, it is anticipated that AV safety protocols and current road design will be enhanced, leveraging the insights gained from this research.

The discoveries made in this thesis, supported by statistical testing, offer insights to AV developers, enhancing their comprehension of how the driving environment impacts AV functions. Through quantitative analysis, developers can pinpoint road segments demanding above-average processing power. This knowledge empowers them to accurately assess processing needs for distinct road sections before AV deployment, thus preventing potentially hazardous situations. Additionally, critical points prone to loss of road information or a decline in data rate requirements are presented, enabling proactive measures to avoid collisions stemming from information gaps. Moreover, comprehending the influence of weather conditions on AV performance empowers informed decision-making regarding tackling and mitigating associated challenges.

This study also allows developers to assess different commercially available LiDAR sensors, evaluate their performance, and accurately estimate the required processing power for the simulated models. Furthermore, the simulation environment developed in this research can be used as testing grounds for AV development where the data processing requirements can be optimized.

Government agencies and IOOs can benefit from this research by evaluating the preparedness of existing infrastructure for AV deployment. Currently, road networks are designed for human-operated vehicles, necessitating thorough reassessments before AV deployment [222]. Insights from this thesis shed light on the impact of road geometry, roadside features, and roadway width on AV performance. This knowledge is crucial for informing future road designs and ensuring they are optimized for AV usage.

The presented analysis indicates that simplifying the surrounding environment, such as reducing the number of lanes, results in a less complex environment for AVs. Additionally, lowering the density of roadside elements, such as environments with open fields, is recommended to enhance AV operations and minimize challenges related to wildlife detection. Furthermore, decreasing road grades and eliminating roadside obstructions at horizontal curves emerge as potent strategies to accommodate AVs effectively.

Integrating (HD) maps empowers AVs with valuable foreknowledge, enabling them to anticipate potential obstacles and proactively counteract information loss, which can be achieved by actively providing AVs with potentially occluded elements [34]. Simultaneously, establishing a robust communication network between AVs and infrastructure is a powerful solution to enhance sensor capabilities and computational limitations by facilitating seamless information exchange among various vehicles and infrastructure components. As such, AVs can dynamically adapt to their surroundings, particularly in critical areas [71]. This extended reach ensures that AVs receive pertinent information beyond their immediate proximity, significantly enhancing their overall awareness and responsiveness.

This comprehensive framework addresses fundamental questions about AV performance and provides actionable transportation and traffic engineering insights. It supports informed decision-making, promotes safety, and guides infrastructure design, making it a valuable asset for advancing AV technology within the transportation sector.

6.4 LIMITATIONS AND FUTURE RESEARCH

While the presented framework is a robust tool for understanding the interaction between AVs and their surrounding environment and anticipating complexity levels in various road and weather conditions, it has limitations in specific regions. One of the limitations lies in the assumptions endorsed by the implemented algorithm, which might not be universally applicable to all AV models. However, the methodology's flexibility allows for integrating more advanced and exhaustive algorithms, offering a deeper understanding of AV performance and environmental complexity. Additionally, to authenticate the validity of the framework’s results, this will require the deployment of an AV in the field which is both costly and has safety risks.

For future endeavors, this framework can be modified to incorporate other sensory information such as cameras and radars. Furthermore, a more holistic approach is necessary to comprehensively model the intricacies of dynamic agent interactions with AVs, especially in urban environments where such dynamic objects abound. Additionally, parametric models can be developed to accurately quantify and predict the influence of the different vertical grades and the radii of the horizontal curve on the data rate requirements.

This framework can be further enhanced by adopting a more nuanced approach where the input sensor specification variables can be assumed to follow a probability distribution instead of using deterministic values. Moreover, modelling different sensor heights can also be conducted in the future to assess how it can impact the environment perception and consequently, the data rate requirements of the AVs.

Last but not least, this work can also be expanded to perform network-level studies, where the most critical locations can be identified to facilitate the optimization process of assigning roadside sensors which can relay real time information for the AVs on the road to assist them in mitigating potentially critical situations.

In light of the “Responsible and Ethical Use of Large Language Models (LLMs) in Academic Settings” guidelines, the authors of this research hereby declare that they have used ChatGPT for proofreading and editing purposes only. Furthermore, all ideas presented in this work are original and belong solely to the authors.

REFERENCES

[1] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, "Applications of Artificial Intelligence in Transport: An Overview," *Sustainability*, vol. 11, no. 1, p. 189, Jan. 2019, doi: 10.3390/su11010189.

[2] M. P. McKay and F. Vaca, "Commentary: Pedestrian roadway fatalities: Profiling the problem," *Annals of Emergency Medicine*, vol. 42, no. 4, pp. 480–482, 2003.

[3] World Health Organization, *Global status report on road safety 2018*. Geneva: World Health Organization, 2018. Accessed: Apr. 27, 2023. [Online]. Available: https://apps.who.int/iris/handle/10665/276462

[4] "Motor Vehicle - Introduction," Injury Facts. Accessed: Oct. 18, 2023. [Online]. Available: https://injuryfacts.nsc.org/motor-vehicle/overview/introduction/

[5] T. Canada, "Canadian Motor Vehicle Traffic Collision Statistics: 2021," Transport Canada. Accessed: Dec. 04, 2023. [Online]. Available: https://tc.canada.ca/en/road-transportation/statistics-data/canadian-motor-vehicle-traffic-collision-statistics-2021

[6] J. P. Michael, N. M. Wells, L. Shahum, H. N. Bidigare-Curtis, S. F. Greenberg, and T. Xu, "Roadway safety, design & equity: A paradigm shift," *Journal of Transport & Health*, vol. 23, p. 101260, Dec. 2021, doi: 10.1016/j.jth.2021.101260.

[7] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," Washington, DC, USA, Tech. Rep. DOT HS 812 115, 2015.

[8] J. Cui, G. Sabaliauskaite, L. S. Liew, F. Zhou, and B. Zhang, "Collaborative analysis framework of safety and security for autonomous vehicles," *IEEE Access*, vol. 7, pp. 148672–148683, 2019, doi: 10.1109/ACCESS.2019.2946632.

[9] J. M. Anderson, N. Kalra, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola, *Autonomous vehicle technology: a guide for policymakers*. Santa Monica, CA: Rand Corporation, 2014.

[10] W. D. Montgomery, R. Mudge, E. L. Groshen, S. Helper, J. P. MacDuffie, and C. Carson, "America's workforce and the self-driving future: Realizing productivity gains and spurring economic growth," 2018.

[11] M. Gouda, J. Mirza, J. Weiß, A. Ribeiro Castro, and K. El-Basyouny, "Octree-based point cloud simulation to assess the readiness of highway infrastructure for autonomous vehicles," *Computer-Aided Civil and Infrastructure Engineering*, vol. 36, no. 7, pp. 922–940, 2021.

[12] D. Elliott, W. Keen, and L. Miao, "Recent advances in connected and automated vehicles," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 6, no. 2, pp. 109–131, 2019.

[13] T. Litman, "Autonomous Vehicle Implementation Predictions: Implications for Transport Planning," *Victoria, Canada: Victoria Transport Policy Institute*.

[14] H. Bradshaw-Martin and C. Easton, "Autonomous or 'driverless' cars and disability: a legal and ethical analysis.," *European Journal of Current Legal Issues*, vol. 20, no. 3, 2014.

[15] H. Surden and M.-A. Williams, "How Self-Driving Cars Work," *SSRN Journal*, 2016, doi: 10.2139/ssrn.2784465.

[16] J. Kemeny and K. Turner, "Ground-based lidar: rock slope mapping and assessment," United States. Federal Highway Administration. Central Federal Lands Highway …, 2008.

[17] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, “Sensor and sensor fusion technology in autonomous vehicles: A review,” *Sensors*, vol. 21, no. 6, p. 2140, Mar. 2021, doi: 10.3390/s21062140.

[18] B. Shahian Jahromi, T. Tulabandhula, and S. Cetin, “Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles,” *Sensors*, vol. 19, no. 20, p. 4357, 2019.

[19] J. Kocić, N. Jovičić, and V. Drndarević, “Sensors and sensor fusion in autonomous vehicles,” in *2018 26th Telecommunications Forum (TELFOR)*, IEEE, 2018, pp. 420–425.

[20] J. Khoury, 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*, vol. 2019, pp. 1–10, May 2019, doi: 10.1155/2019/6126408.

[21] T. He, L. Zhang, F. Kong, and A. Salekin, “Exploring inherent sensor redundancy for automotive anomaly d\etection,” in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, Jul. 2020, pp. 1–6. doi: 10.1109/DAC18072.2020.9218557.

[22] Y. Li and H. Shi, Eds., *Advanced driver assistance systems and autonomous vehicles: From fundamentals to applications*. Singapore: Springer Nature Singapore, 2022. doi: 10.1007/978-981-19-5053-7.

[23] H. Zhu, K.-V. Yuen, L. Mihaylova, and H. Leung, “Overview of environment perception for intelligent vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 10, pp. 2584–2601, 2017.

[24] C. D. Prakash, F. Akhbari, and L. J. Karam, “Robust obstacle detection for advanced driver assistance systems using distortions of inverse perspective mapping of a monocular camera,” *Robotics and Autonomous Systems*, vol. 114, pp. 172–186, 2019.

[25] C. Michaelis *et al.*, “Benchmarking robustness in object detection: Autonomous Driving When Winter is Coming,” *arXiv preprint arXiv:1907.07484*, 2019.

[26] D. Neven, B. D. Brabandere, S. Georgoulis, M. Proesmans, and L. V. Gool, “Towards end-to-end lane detection: An instance segmentation approach,” in *2018 IEEE Intelligent Vehicles Symposium (IV)*, Jun. 2018, pp. 286–291. doi: 10.1109/IVS.2018.8500547.

[27] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” *arXiv preprint arXiv:1804.02767*, 2018.

[28] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “Yolov4: Optimal speed and accuracy of object detection,” *arXiv preprint arXiv:2004.10934*, 2020.

[29] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, “Spatial as deep: Spatial cnn for traffic scene understanding,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.

[30] S. Lee *et al.*, “Vpgnet: Vanishing point guided network for lane and road marking detection and recognition,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 1947–1955.

[31] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. Mccullough, and A. Mouzakitis, “A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 829–846, 2018.

[32] K. Jiang, D. Yang, C. Liu, T. Zhang, and Z. Xiao, “A flexible multi-layer map model designed for lane-level route planning in autonomous vehicles,” *Engineering*, vol. 5, no. 2, pp. 305–318, 2019.

[33] C. Urmson *et al.*, “High speed navigation of unrehearsed terrain: Red team technology for grand challenge 2004,” *Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-04-37*, 2004.

[34] Z. Bao, S. Hossain, H. Lang, and X. Lin, “A review of high-definition map creation methods for autonomous driving,” *Engineering Applications of Artificial Intelligence*, vol. 122, p. 106125, Jun. 2023, doi: 10.1016/j.engappai.2023.106125.

[35] S. Das, R. Kumari, and S. Deepak Kumar, “A review on applications of simultaneous localization and mapping method in autonomous vehicles,” *Advances in Interdisciplinary Engineering: Select Proceedings of FLAME 2020*, pp. 367–375, 2021.

[36] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, “Simultaneous localization and mapping: A survey of current trends in autonomous driving,” *IEEE Transactions on Intelligent Vehicles*, vol. 2, no. 3, pp. 194–220, 2017.

[37] J. Zhang and S. Singh, “LOAM: Lidar odometry and mapping in real-time.,” in *Robotics: Science and systems*, Berkeley, CA, 2014, pp. 1–9.

[38] J.-E. Deschaud, “IMLS-SLAM: Scan-to-model matching based on 3D data,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2018, pp. 2480–2485.

[39] L. Liu *et al.*, “Computing Systems for Autonomous Driving: State of the Art and Challenges,” *IEEE Internet Things J.*, vol. 8, no. 8, pp. 6469–6486, Apr. 2021, doi: 10.1109/JIOT.2020.3043716.

[40] S. Sumikura, M. Shibuya, and K. Sakurada, “OpenVSLAM: A versatile visual SLAM framework,” in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 2292–2295.

[41] D. Schlegel, M. Colosi, and G. Grisetti, “Proslam: Graph SLAM from a programmer’s Perspective,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2018, pp. 3833–3840.

[42] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, “A survey of autonomous driving: Common practices and emerging technologies,” *IEEE Access*, vol. 8, pp. 58443–58469, 2020, doi: 10.1109/ACCESS.2020.2983149.

[43] X. Geng, H. Liang, B. Yu, P. Zhao, L. He, and R. Huang, “A scenario-adaptive driving behavior prediction approach to urban autonomous driving,” *Applied Sciences*, vol. 7, no. 4, p. 426, 2017.

[44] D. González, J. Pérez, V. Milanés, and F. Nashashibi, “A review of motion planning techniques for automated vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 1135–1145, 2015.

[45] S. Allou, Y. Zennir, and A. Belmeguenai, “Fuzzy logic controller for autonomous vehicle path tracking,” in *2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (sta)*, IEEE, 2017, pp. 328–333.

[46] A. Baskaran, A. Talebpour, and S. Bhattacharyya, “End-to-end drive by-wire PID lateral control of an autonomous vehicle,” in *Proceedings of the Future Technologies Conference*, Springer, 2019, pp. 365–376.

[47] H. Nam, W. Choi, and C. Ahn, “Model predictive control for evasive steering of an autonomous vehicle,” *International Journal of Automotive Technology*, vol. 20, pp. 1033–1042, 2019.

[48] Y. Xiao, F. Codevilla, A. Gurram, O. Urfalioglu, and A. M. López, "Multimodal end-to-end autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 537–547, 2020.

[49] L. Le Mero, D. Yi, M. Dianati, and A. Mouzakitis, "A survey on imitation learning techniques for end-to-end autonomous vehicles," *IEEE Trans. Intell. Transport. Syst.*, vol. 23, no. 9, pp. 14128–14147, Sep. 2022, doi: 10.1109/TITS.2022.3144867.

[50] C. Chen, A. Seff, A. Kornhauser, and J. Xiao, "DeepDriving: Learning affordance for direct perception in autonomous driving," in *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile: IEEE, Dec. 2015, pp. 2722–2730. doi: 10.1109/ICCV.2015.312.

[51] K. Muhammad, A. Ullah, J. Lloret, J. D. Ser, and V. H. C. de Albuquerque, "Deep learning for safe autonomous driving: Current challenges and future directions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4316–4336, Jul. 2021, doi: 10.1109/TITS.2020.3032227.

[52] J. Janai, F. Güney, A. Behl, and A. Geiger, "Computer vision for autonomous vehicles: Problems, datasets and state of the art," *Foundations and Trends® in Computer Graphics and Vision*, vol. 12, no. 1–3, pp. 1–308, 2020.

[53] David_P, "Preparing Infrastructure for Automated Vehicles," ITF. [Online]. Available: https://www.itf-oecd.org/preparing-infrastructure-automated-vehicles

[54] B. Lundgren, "Safety requirements vs. crashing ethically: what matters most for policies on autonomous vehicles," *AI & SOCIETY*, vol. 36, no. 2, pp. 405–415, 2021.

[55] D. J. Hicks, "The safety of autonomous vehicles: lessons from philosophy of science," *IEEE Technology and Society Magazine*, vol. 37, no. 1, pp. 62–69, Mar. 2018, doi: 10.1109/MTS.2018.2795123.

[56] L. D. Burns and C. Shulgan, *Autonomy: The Quest to Build the Driverless Car—And How It Will Reshape Our World*. NY, USA: HarperCollins, 2018.

[57] E. Ackerman, "Toyota's Gill Pratt on self-driving cars and the reality of full autonomy," *IEEE Spectrum*, vol. 23, 2021, [Online]. Available: https://spectrum.ieee.org/toyota-gill-pratt-on-the-reality-of-full-autonomy

[58] D. Lavrinc, "This is how bad self-driving cars suck in rain," *Technology Report*, 2014.

[59] A. Davies, "Google's self-driving car caused its first crash," in *Wired*, 2016.

[60] "US report: Nearly 400 crashes of automated tech vehicles," AP NEWS. Accessed: May 05, 2023. [Online]. Available: https://apnews.com/article/self-driving-car-crash-data-ae87cadec79966a9ba56e99b4110b8d6

[61] D. Wakabayashi, "Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam," *The New York Times*, Mar. 19, 2018. Accessed: May 05, 2023. [Online]. Available: https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html

[62] M. McFarland, "Who's responsible when an autonomous car crashes," *Accessed: Jul*, vol. 24, p. 2020, 2016.

[63] R. L. McCarthy, "Autonomous vehicle accident data analysis: California ol 316 reports: 2015–2020," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 8, no. 3, p. 034502, 2022.

[64] F. Cugurullo, R. A. Acheampong, M. Gueriau, and I. Dusparic, “The transition to autonomous cars, the redesign of cities and the future of urban sustainability,” *Urban Geography*, vol. 42, no. 6, pp. 833–859, 2021.
[65] M. R. Conover, “Numbers of human fatalities, injuries, and illnesses in the United States due to wildlife,” *Human–Wildlife Interactions*, vol. 13, no. 2, p. 12, 2019.
[66] F. Ascensão, D. R. Yogui, M. H. Alves, A. C. Alves, F. Abra, and A. L. J. Desbiez, “Preventing wildlife roadkill can offset mitigation investments in short-medium term,” *Biological Conservation*, vol. 253, p. 108902, Jan. 2021, doi: 10.1016/j.biocon.2020.108902.
[67] M. P. Huijser *et al.*, “Wildlife-vehicle collision reduction study: Report to Congress,” 2017.
[68] F. Shilling, D. Waetjen, K. Harrold, and P. Farman, “Impact of Wildlife-Vehicle Conflict on California Drivers and Animals,” *UC Davis Road Ecology Center Retrieved from https://roadecology. ucdavis. edu/files/content/news/CA_WVC_Hotspots_2018. pdf*, 2018.
[69] M. Gren and A. Jägerbrand, “Calculating the costs of animal-vehicle accidents involving ungulate in Sweden,” *Transportation research part D: transport and environment*, vol. 70, pp. 112–122, 2019.
[70] Government of Alberta, “Alberta wildlife watch an animal-vehicle collision safety program.” Accessed: Dec. 04, 2023. [Online]. Available: https://open.alberta.ca/publications/alberta-wildlife-watch-animal-vehicle-collision-safety-program
[71] H. U. Ahmed, Y. Huang, P. Lu, and R. Bridgelall, “Technology developments and impacts of connected and autonomous vehicles: An overview,” *Smart Cities*, vol. 5, no. 1, pp. 382–404, Mar. 2022, doi: 10.3390/smartcities5010022.
[72] F. Rosique, P. J. Navarro, C. Fernández, and A. Padilla, “A systematic review of perception system and simulators for autonomous vehicles research,” *Sensors*, vol. 19, no. 3, p. 648, Feb. 2019, doi: 10.3390/s19030648.
[73] S. Gupta, D. Chand, and I. Kavati, “Computer vision based animal collision avoidance framework for autonomous vehicles,” in *Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part III 5*, Springer, 2021, pp. 237–248.
[74] A. Saxena, D. K. Gupta, and S. Singh, “An animal detection and collision avoidance system using deep learning,” in *Advances in Communication and Computational Technology: Select Proceedings of ICACCT 2019*, Springer, 2021, pp. 1069–1084.
[75] G. Li *et al.*, “Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios,” *Transportation research part C: emerging technologies*, vol. 122, p. 102820, 2021.
[76] I. Silva and J. Calabrese, “Emerging opportunities for wildlife with sustainable autonomous transportation,” Physical Sciences and Mathematics, preprint, Sep. 2021. doi: 10.32942/OSF.IO/BD38F.
[77] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator,” in *Conference on robot learning*, PMLR, 2017, pp. 1–16.
[78] M. Deitke *et al.*, “Robothor: An open simulation-to-real embodied ai platform,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3164–3174.

[79] Ł. Sobczak, K. Filus, A. Domański, and J. Domańska, “LiDAR point cloud generation for SLAM algorithm evaluation,” *Sensors*, vol. 21, no. 10, p. 3313, 2021.

[80] J. Jung, M. J. Olsen, D. S. Hurwitz, A. G. Kashani, and K. Buker, “3D virtual intersection sight distance analysis using lidar data,” *Transportation research part C: emerging technologies*, vol. 86, pp. 563–579, 2018.

[81] Y. Ma, Y. Zheng, S. Easa, Y. D. Wong, and K. El-Basyouny, “Virtual analysis of urban road visibility using mobile laser scanning data and deep learning,” *Automation in Construction*, vol. 133, p. 104014, 2022.

[82] N. Kalra and S. M. Paddock, “Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?,” *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 182–193, 2016.

[83] K. Williams, M. J. Olsen, G. V. Roe, and C. Glennie, “Synthesis of transportation applications of mobile LiDAR,” *Remote Sensing*, vol. 5, no. 9, pp. 4652–4692, 2013.

[84] B. Rezvani, B. Hassibi, and L. Burns, “Zero Roadway Deaths Means Seeing Everything Sooner, Clearer and Farther,” Neural Propulsion Systems, Inc, Pleasanton, CA, USA, Feb. 2022. [Online]. Available: https://nps.ai

[85] S. Manivasagam *et al.*, “LiDARsim: Realistic lidar simulation by leveraging the real world,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11167–11176.

[86] M. Gschwandtner, R. Kwitt, A. Uhl, and W. Pree, “BlenSor: Blender Sensor Simulation Toolbox,” in *Advances in Visual Computing*, vol. 6939, G. Bebis, R. Boyle, B. Parvin, D. Koracin, S. Wang, K. Kyungnam, B. Benes, K. Moreland, C. Borst, S. DiVerdi, C. Yi-Jen, and J. Ming, Eds., in Lecture Notes in Computer Science, vol. 6939. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 199–208. doi: 10.1007/978-3-642-24031-7_20.

[87] X. Yue, B. Wu, S. A. Seshia, K. Keutzer, and A. L. Sangiovanni-Vincentelli, “A lidar point cloud generator: from a virtual world to autonomous driving,” in *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*, 2018, pp. 458–464.

[88] S. Gilroy, E. Jones, and M. Glavin, “Overcoming occlusion in the automotive environment—A review,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 23–35, Jan. 2021, doi: 10.1109/TITS.2019.2956813.

[89] A. Shetty, M. Yu, A. Kurzhanskiy, O. Grembek, H. Tavafoghi, and P. Varaiya, “Safety challenges for autonomous vehicles in the absence of connectivity,” *Transportation research part C: emerging technologies*, vol. 128, p. 103133, 2021.

[90] J. Fang *et al.*, “Simulating LIDAR point cloud for autonomous driving using real-world scenes and traffic flows,” *arXiv preprint arXiv:1811.07112*, vol. 1, 2018.

[91] L. G. Galvao, M. Abbod, T. Kalganova, V. Palade, and M. N. Huda, “Pedestrian and vehicle detection in autonomous vehicle perception systems—A review,” *Sensors (14248220)*, vol. 21, no. 21, p. 7267, Nov. 2021, doi: 10.3390/s21217267.

[92] M. Sha and A. Boukerche, “Performance evaluation of CNN-based pedestrian detectors for autonomous vehicles,” *Ad Hoc Networks*, vol. 128, p. 102784, Apr. 2022, doi: 10.1016/j.adhoc.2022.102784.

[93] D. Parekh *et al.*, “A review on autonomous vehicles: Progress, methods and challenges,” *Electronics*, vol. 11, no. 14, p. 2162, Jul. 2022, doi: 10.3390/electronics11142162.

[94] K. Bimbraw, “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)*, Jul. 2015, pp. 191–198.
[95] R. Behringer and N. Muller, “Autonomous road vehicle guidance from autobahnen to narrow curves,” *IEEE Transactions on Robotics and Automation*, vol. 14, no. 5, pp. 810–815, 1998.
[96] S. Chandran, L. S. Davis, D. Dementhon, S. J. Dickenson, and S. Gajulapalli, “An overview of vision-based navigation for autonomous land vehicles 1986.,” 1987.
[97] P. Resende, E. Pollard, H. Li, and F. Nashashibi, “Low Speed Automation: technical feasibility of the driving sharing in urban areas,” in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, IEEE, 2013, pp. 442–448.
[98] J. Wenger, “Automotive radar-status and perspectives,” in *IEEE Compound Semiconductor Integrated Circuit Symposium, 2005. CSIC’05.*, Ieee, 2005, p. 4 pp.
[99] C. Rouff and M. Hinchey, *Experience from the DARPA Urban Challenge*. Springer Science & Business Media, 2012.
[100] M. Buehler, K. Iagnemma, and S. Singh, *The 2005 DARPA grand challenge: The great robot race*, vol. 36. Springer, 2007.
[101] M. Campbell, M. Egerstedt, J. P. How, and R. M. Murray, “Autonomous driving in urban environments: approaches, lessons and challenges,” *Phil. Trans. R. Soc. A.*, vol. 368, no. 1928, pp. 4649–4672, Oct. 2010, doi: 10.1098/rsta.2010.0110.
[102] A. Broggi, P. Cerri, M. Felisa, M. C. Laghi, L. Mazzei, and P. P. Porta, “The VisLab intercontinental autonomous challenge: An extensive test for a platoon of intelligent vehicles,” *IJVAS*, vol. 10, no. 3, p. 147, 2012, doi: 10.1504/IJVAS.2012.051250.
[103] J. Van Brummelen, M. O’Brien, D. Gruyer, and H. Najjaran, “Autonomous vehicle perception: The technology of today and tomorrow,” *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 384–406, Apr. 2018, doi: 10.1016/j.trc.2018.02.012.
[104] Y. Li, Y. Cai, R. Malekian, H. Wang, M. A. Sotelo, and Z. Li, “Creating navigation map in semi-open scenarios for intelligent vehicle localization using multi-sensor fusion,” *Expert Systems with Applications*, vol. 184, p. 115543, Dec. 2021, doi: 10.1016/j.eswa.2021.115543.
[105] M. N. Ahangar, Q. Z. Ahmed, F. A. Khan, and M. Hafeez, “A survey of autonomous vehicles: Enabling communication technologies and challenges,” *Sensors*, vol. 21, no. 3, p. 706, Jan. 2021, doi: 10.3390/s21030706.
[106] “J3016_202104: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles - SAE International.” Accessed: May 01, 2023. [Online]. Available: https://www.sae.org/standards/content/j3016_202104/
[107] M. R. Hafner, D. Cunningham, L. Caminiti, and D. Del Vecchio, “Cooperative collision avoidance at intersections: Algorithms and experiments,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1162–1175, 2013.
[108] J. Lee *et al.*, “Emergency pull-over algorithm for Level 4 autonomous vehicles based on model-free adaptive feedback control with sensitivity and learning approaches,” *IEEE Access*, vol. 10, pp. 27014–27030, 2022, doi: 10.1109/ACCESS.2022.3156275.

[109] A. Bartels, T. Ruchatz, and S. Brosig, “Intelligence in the automobile of the future,” *Smart Mobile In-Vehicle Systems: Next Generation Advancements*, pp. 35–46, 2014.
[110] E. Guizzo, “Toyota’s semi-autonomous car will keep you safe,” *IEEE Spectrum Automation*, p. 8, 2013.
[111] L. Ulrich, “Top 10 tech cars: slenderized,” *IEEE Spectrum*, vol. 50, no. 4, pp. 34–41, Apr. 2013, doi: 10.1109/MSPEC.2013.6481696.
[112] D. Kim, R. R. L. Mendoza, K. F. R. Chua, M. A. A. Chavez, R. S. Concepcion, and R. R. P. Vicerra, “A systematic analysis on the trends and challenges in autonomous vehicles and the proposed solutions for level 5 automation,” in *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Manila, Philippines: IEEE, Nov. 2021, pp. 1–6. doi: 10.1109/HNICEM54116.2021.9731982.
[113] J. M. Gitlin, “Assists, autopilot, and more: Ars talks about autonomous driving with Audi,” Ars Technica. Accessed: Sep. 06, 2023. [Online]. Available: https://arstechnica.com/cars/2016/01/assists-autopilot-and-more-ars-talks-about-autonomous-driving-with-audi/
[114] M. Martínez-Díaz and F. Soriguera, “Autonomous vehicles: theoretical and practical challenges,” *Transportation Research Procedia*, vol. 33, pp. 275–282, 2018, doi: 10.1016/j.trpro.2018.10.103.
[115] Venkata Satya Rahul Kosuru and Ashwin Kavasseri Venkitaraman, “Advancements and challenges in achieving fully autonomous self-driving vehicles,” *World J. Adv. Res. Rev.*, vol. 18, no. 1, pp. 161–167, Apr. 2023, doi: 10.30574/wjarr.2023.18.1.0568.
[116] J. Zmud, G. Goodin, M. Moran, N. Kalra, and E. Thorn, *Advancing automated and connected vehicles: policy and planning strategies for state and local transportation agencies*, no. Project 20-102 (01). 2017.
[117] F. Kuhnert, C. Stürmer, and A. Koster, “Five trends transforming the Automotive Industry,” *PricewaterhouseCoopers GmbH Wirtschaftsprüfungsgesellschaft: Berlin, Germany*, vol. 1, no. 1, pp. 1–48, 2018.
[118] S. Shaheen, H. Totte, and A. Stocker, “Future of mobility white paper,” *UC Berkeley*, 2018.
[119] S. Taghavi and W. Shi, “EdgeMask: An edge-based privacy preserving service for video data sharing,” in *2020 IEEE/ACM Symposium on Edge Computing (SEC)*, IEEE, 2020, pp. 382–387.
[120] K. C. Zeng *et al.*, “All your {GPS} are belong to us: Towards stealthy manipulation of road navigation systems,” in *27th USENIX security symposium (USENIX security 18)*, 2018, pp. 1527–1544.
[121] C. Ilas, “Electronic sensing technologies for autonomous ground vehicles: A review,” in *2013 8th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, IEEE, 2013, pp. 1–6.
[122] T. B. Lee, “Autopilot was active when a Tesla crashed into a truck, killing driver,” *Ars Technica*, 2019.
[123] J. Wang, L. Zhang, Y. Huang, J. Zhao, and F. Bella, “Safety of autonomous vehicles,” *Journal of Advanced Transportation*, vol. 2020, pp. 1–13, Sep. 2020, doi: 10.1155/2020/8867757.
[124] L. Collingwood, “Privacy implications and liability issues of autonomous vehicles,” *Information & Communications Technology Law*, vol. 26, no. 1, pp. 32–45, 2017.

[125] M. Alawadhi, J. Almazrouie, M. Kamil, and K. A. Khalil, “Review and analysis of the importance of autonomous vehicles liability: a systematic literature review,” *International Journal of System Assurance Engineering and Management*, vol. 11, no. 6, pp. 1227–1249, Dec. 2020, doi: 10.1007/s13198-020-00978-9.
[126] D. A. Pomerleau, “Alvinn: An autonomous land vehicle in a neural network,” *Advances in Neural Information Processing Systems*, vol. 1, 1988.
[127] F. Neuhaus, D. Dillenberger, J. Pellenz, and D. Paulus, “Terrain drivability analysis in 3D laser range data for autonomous robot navigation in unstructured environments,” in *2009 IEEE Conference on Emerging Technologies & Factory Automation*, IEEE, 2009, pp. 1–4.
[128] W. Li *et al.*, “AADS: Augmented autonomous driving simulation using data-driven algorithms,” *Science Robotics*, vol. 4, no. 28, p. eaaw0863, 2019.
[129] J. Fang *et al.*, “Augmented LiDAR simulator for autonomous driving,” *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1931–1938, Apr. 2020, doi: 10.1109/LRA.2020.2969927.
[130] F. Wang, Y. Zhuang, H. Gu, and H. Hu, “Automatic generation of synthetic LiDAR point clouds for 3-D data analysis,” *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 7, pp. 2671–2673, Jul. 2019, doi: 10.1109/TIM.2019.2906416.
[131] S. Shah, D. Dey, C. Lovett, and A. Kapoor, “Airsim: High-fidelity visual and physical simulation for autonomous vehicles,” in *Field and Service Robotics: Results of the 11th International Conference*, Springer, 2018, pp. 621–635.
[132] W. Qiu and A. Yuille, “Unrealcv: Connecting computer vision to unreal engine,” in *Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14*, Springer, 2016, pp. 909–916.
[133] U. G. Engine, “Unity game engine-official site,” *Online][Cited: October 9, 2008.] http://unity3d. com*, pp. 1534–4320, 2008.
[134] F. Di Stefano, S. Chiappini, A. Gorreja, M. Balestra, and R. Pierdicca, “Mobile 3D scan LiDAR: a literature review,” *Geomatics, Natural Hazards and Risk*, vol. 12, no. 1, pp. 2387–2429, Jan. 2021, doi: 10.1080/19475705.2021.1964617.
[135] P. D. Groves, “Principles of GNSS, inertial, and multisensor integrated navigation systems, [Book review],” *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 2, pp. 26–27, 2015.
[136] L. Chang, X. Niu, T. Liu, J. Tang, and C. Qian, “GNSS/INS/LiDAR-SLAM integrated navigation system based on graph optimization,” *Remote Sensing*, vol. 11, no. 9, p. 1009, 2019.
[137] Q. Wu, H. Yang, M. Wei, O. Remil, B. Wang, and J. Wang, “Automatic 3D reconstruction of electrical substation scene from LiDAR point cloud,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 143, pp. 57–71, Sep. 2018, doi: 10.1016/j.isprsjprs.2018.04.024.
[138] H. Guan, J. Li, S. Cao, and Y. Yu, “Use of mobile LiDAR in road information inventory: A review,” *International Journal of Image and Data Fusion*, vol. 7, no. 3, pp. 219–242, 2016.
[139] S. A. Gargoum and K. El Basyouny, “A literature synthesis of LiDAR applications in transportation: feature extraction and geometric assessments of highways,” *GIScience & Remote Sensing*, vol. 56, no. 6, pp. 864–893, Aug. 2019, doi: 10.1080/15481603.2019.1581475.
[140] C. G. Duffell and D. M. Rudrum, “Remote sensing techniques for highway earthworks assessment,” in *Site Characterization and Modelling*, 2005, pp. 1–13.

[141] J. Landa and D. Prochazka, “Automatic road inventory using LiDAR,” *Procedia Economics and Finance*, vol. 12, pp. 363–370, 2014.
[142] J. Lam, K. Kusevic, P. Mrstik, R. Harrap, and M. Greenspan, “Urban scene extraction from mobile ground based lidar data,” in *Proceedings of 3DPVT*, 2010, pp. 1–8.
[143] S. Hu, H. Chen, B. Wang, J. Gong, and Y. Ma, “LiDAR-based road extraction for UGV in high definition map,” in *2020 3rd International Conference on Unmanned Systems (ICUS)*, IEEE, 2020, pp. 303–308.
[144] H. Wang *et al.*, “Automatic road extraction from mobile laser scanning data,” in *2012 International Conference on Computer Vision in Remote Sensing*, IEEE, 2012, pp. 136–139.
[145] M. Yadav, A. K. Singh, and B. Lohani, “Extraction of road surface from mobile LiDAR data of complex road environment,” *International Journal of Remote Sensing*, vol. 38, no. 16, pp. 4655–4682, 2017.
[146] M. Yadav, B. Lohani, A. K. Singh, and A. Husain, “Identification of pole-like structures from mobile lidar data of complex road environment,” *International Journal of Remote Sensing*, vol. 37, no. 20, pp. 4748–4777, 2016.
[147] M. Lehtomäki *et al.*, “Object classification and recognition from mobile laser scanning point clouds in a road environment,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 1226–1239, 2015.
[148] B. Wu *et al.*, “A voxel-based method for automated identification and morphological parameters estimation of individual street trees from mobile laser scanning data,” *Remote Sensing*, vol. 5, no. 2, pp. 584–611, 2013.
[149] C. Cabo, C. Ordoñez, S. García-Cortés, and J. Martínez, “An algorithm for automatic detection of pole-like street furniture objects from Mobile Laser Scanner point clouds,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 47–56, 2014.
[150] Y. Yu, J. Li, H. Guan, and C. Wang, “Automated extraction of urban road facilities using mobile laser scanning data,” *IEEE Transactions on intelligent transportation systems*, vol. 16, no. 4, pp. 2167–2181, 2015.
[151] X. Mi, B. Yang, Z. Dong, C. Chen, and J. Gu, “Automated 3D road boundary extraction and vectorization using MLS point clouds,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5287–5297, 2021.
[152] L. Zhou and Z. Deng, “LIDAR and vision-based real-time traffic sign detection and recognition algorithm for intelligent vehicle,” in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2014, pp. 578–583.
[153] W. Zhang, “Lidar-based road and road-edge detection,” in *2010 IEEE Intelligent Vehicles Symposium*, IEEE, 2010, pp. 845–848.
[154] A. Serna and B. Marcotegui, “Urban accessibility diagnosis from mobile laser scanning data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 84, pp. 23–32, 2013.
[155] Y. Zhang, J. Wang, X. Wang, and J. M. Dolan, “Road-segmentation-based curb detection method for self-driving via a 3D-LiDAR sensor,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 12, pp. 3981–3991, 2018.
[156] A. Jaakkola, J. Hyyppä, H. Hyyppä, and A. Kukko, “Retrieval algorithms for road surface modelling using laser-based mobile mapping,” *Sensors*, vol. 8, no. 9, pp. 5238–5249, 2008.

[157] P. Kumar, C. P. McElhinney, P. Lewis, and T. McCarthy, “Automated road markings extraction from mobile laser scanning data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 32, pp. 125–137, 2014.
[158] H. Guan, J. Li, Y. Yu, C. Wang, M. Chapman, and B. Yang, “Using mobile laser scanning data for automated extraction of road markings,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 93–107, 2014.
[159] D. Kukolj, I. Marinović, and S. Nemet, “Road edge detection based on combined deep learning and spatial statistics of LiDAR data,” *Journal of Spatial Science*, pp. 1–15, 2021.
[160] L. Zhao, L. Yan, and X. Meng, “The extraction of street curbs from mobile laser scanning data in urban areas,” *Remote Sensing*, vol. 13, no. 12, p. 2407, 2021.
[161] D. Hutchison *et al.*, “Superpixels and Supervoxels in an Energy Optimization Framework,” in *Computer Vision – ECCV 2010*, vol. 6315, K. Daniilidis, P. Maragos, and N. Paragios, Eds., in Lecture Notes in Computer Science, vol. 6315. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 211–224. doi: 10.1007/978-3-642-15555-0_16.
[162] M. Gouda, A. Epp, R. Tilroe, and K. El-Basyouny, “Traffic sign extraction using deep hierarchical feature learning and mobile light detection and ranging (LiDAR) data on rural highways,” *Journal of Intelligent Transportation Systems*, vol. 27, no. 5, pp. 643–664, 2023.
[163] C. Ai and Y.-C. J. Tsai, “Critical assessment of an enhanced traffic sign detection method using mobile LiDAR and INS technologies,” *Journal of Transportation Engineering*, vol. 141, no. 5, p. 04014096, 2015.
[164] S. Gargoum, K. El-Basyouny, J. Sabbagh, and K. Froese, “Automated highway sign extraction using lidar data,” *Transportation Research Record*, vol. 2643, no. 1, pp. 1–8, Jan. 2017, doi: 10.3141/2643-01.
[165] H. Guan, W. Yan, Y. Yu, L. Zhong, and D. Li, “Robust traffic-sign detection and classification using mobile LiDAR data with digital images,” *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 11, no. 5, pp. 1715–1724, May 2018, doi: 10.1109/JSTARS.2018.2810143.
[166] M. Javanmardi, Z. Song, and X. Qi, “Automated traffic sign and light pole detection in mobile LiDAR scanning data,” *IET Intelligent Transport Systems*, vol. 13, no. 5, pp. 803–815, 2019, doi: 10.1049/iet-its.2018.5360.
[167] M. Gouda, A. Shalkamy, X. Li, and K. El-Basyouny, “Fully automated algorithm for light pole detection and mapping in rural highway environment using mobile light detection and ranging point clouds,” *Transportation Research Record*, vol. 2676, no. 7, pp. 617–629, Jul. 2022, doi: 10.1177/03611981221082531.
[168] Y. Tsai, C. Ai, Z. Wang, and E. Pitts, “Mobile cross-slope measurement method using lidar technology,” *Transportation Research Record*, vol. 2367, no. 1, pp. 53–59, 2013.
[169] A. Holgado-Barco, B. Riveiro, D. González-Aguilera, and P. Arias, “Automatic inventory of road cross-sections from mobile laser scanning system,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 1, pp. 3–17, 2017.
[170] S. A. Gargoum, K. El-Basyouny, K. Froese, and A. Gadowski, “A fully automated approach to extract and assess road cross sections from mobile lidar data,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 11, pp. 3507–3516, 2018.

[171] R. Souleyrette, S. Hallmark, S. Pattnaik, M. O'Brien, and D. Veneziano, "Grade and cross slope estimation from LiDAR-based surface models," The National Consortium of Remote Sensing in Transportation, 2003.
[172] K. Zhang and H. C. Frey, "Road grade estimation for on-road vehicle emissions modelling using light detection and ranging data," *Journal of the Air & Waste Management Association*, vol. 56, no. 6, pp. 777–788, 2006.
[173] S. Higuera de Frutos and M. Castro, "A method to identify and classify the vertical alignment of existing roads," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 11, pp. 952–963, 2017.
[174] P. Di Mascio, M. Di Vito, G. Loprencipe, and A. Ragnoli, "Procedure to determine the geometry of road alignment using GPS data," *Procedia-Social and Behavioral Sciences*, vol. 53, pp. 1202–1215, 2012.
[175] A. Holgado-Barco, D. González-Aguilera, P. Arias-Sanchez, and J. Martinez-Sanchez, "Semiautomatic extraction of road horizontal alignment from a mobile LiDAR system," *Computer-Aided Civil and Infrastructure Engineering*, vol. 30, no. 3, pp. 217–228, 2015.
[176] S. Gargoum, K. El-Basyouny, and J. Sabbagh, "Automated extraction of horizontal curve attributes using LiDAR data," *Transportation Research Record*, vol. 2672, no. 39, pp. 98–106, 2018.
[177] A. Shalkamy, L. Karsten, S. Gargoum, and K. El-Basyouny, "A framework to detect horizontal curves and assess their geometric properties from remotely sensed point clouds," *International Journal of Remote Sensing*, vol. 41, no. 21, pp. 8328–8351, Nov. 2020, doi: 10.1080/01431161.2020.1771792.
[178] M. Castro, L. Iglesias, J. A. Sánchez, and L. Ambrosio, "Sight distance analysis of highways using GIS tools," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 997–1005, 2011.
[179] M. Castro, J. A. Anta, L. Iglesias, and J. A. Sánchez, "GIS-based system for sight distance analysis of highways," *Journal of computing in civil engineering*, vol. 28, no. 3, p. 04014005, 2014.
[180] S. A. Gargoum, L. Karsten, K. El-Basyouny, and J. C. Koch, "Automated assessment of vertical clearance on highways scanned using mobile LiDAR technology," *Automation in Construction*, vol. 95, pp. 260–274, 2018.
[181] K. Wang, X. Feng, H. Li, and Y. Ren, "Exploring influential factors affecting the severity of urban expressway collisions: A study based on collision data," *International Journal of Environmental Research and Public Health*, vol. 19, no. 14, p. 8362, 2022.
[182] Y. Wang, K. Li, Y. Hu, and H. Chen, "Modelling and quantitative assessment of environment complexity for autonomous vehicles," in *2020 Chinese Control and Decision Conference (CCDC)*, Hefei, China: IEEE, Aug. 2020, pp. 2124–2129. doi: 10.1109/CCDC49329.2020.9164239.
[183] J. Wang, C. Zhang, Y. Liu, and Q. Zhang, "Traffic sensory data classification by quantifying scenario complexity," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, Changshu: IEEE, Jun. 2018, pp. 1543–1548. doi: 10.1109/IVS.2018.8500669.
[184] H. C. Zhang, "Research on complexity of road traffic environment based on gravitation model," *MS thesis, Dept. Transp. Eng., Beijing Inst. Technol., Beijing, China,* 2016.

[185] S. Yang, L. Gao, Y. Zhao, and X. Li, “Research on the quantitative evaluation of the traffic environment complexity for unmanned vehicles in urban roads,” *IEEE Access*, vol. 9, pp. 23139–23152, 2021, doi: 10.1109/ACCESS.2021.3054882.
[186] R. Yu, Y. Zheng, and X. Qu, “Dynamic driving environment complexity quantification method and its verification,” *Transportation Research Part C: Emerging Technologies*, vol. 127, p. 103051, 2021.
[187] M. T. Wolf and J. W. Burdick, “Artificial potential functions for highway driving with collision avoidance,” in *2008 IEEE International Conference on Robotics and Automation*, IEEE, 2008, pp. 3731–3736.
[188] J. Wang, J. Wu, and Y. Li, “The driving safety field based on driver–vehicle–road interactions,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 2203–2214, Aug. 2015, doi: 10.1109/TITS.2015.2401837.
[189] Q. Tu, H. Chen, and J. Li, “A potential field based lateral planning method for autonomous vehicles,” *SAE International Journal of Passenger Cars-Electronic and Electrical Systems*, vol. 10, no. 2016-01–1874, pp. 24–34, 2016.
[190] Y. Rasekhipour, A. Khajepour, S.-K. Chen, and B. Litkouhi, “A potential field-based model predictive path-planning controller for autonomous road vehicles,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1255–1267, May 2017, doi: 10.1109/TITS.2016.2604240.
[191] Y. Cheng, Z. Liu, L. Gao, Y. Zhao, and T. Gao, “Traffic risk environment impact analysis and complexity assessment of autonomous vehicles based on the potential field method,” *IJERPH*, vol. 19, no. 16, p. 10337, Aug. 2022, doi: 10.3390/ijerph191610337.
[192] Z. Yang, Y. Hu, and Y. Zhang, “Path-planning strategy for lane changing based on adaptive-grid risk-fields of autonomous vehicles,” *World Electric Vehicle Journal*, vol. 13, no. 10, p. 175, 2022.
[193] P. Kalvoda, J. Nosek, M. Kuruc, and T. Volarik, “Accuracy evaluation of RIEGL VMX-450 mobile mapping system,” *International Multidisciplinary Scientific GeoConference: SGEM*, vol. 20, no. 2.2, pp. 165–174, 2020.
[194] M. J. Olsen *et al.*, *NCHRP 15-44 Guidelines for the Use of Mobile LiDAR in Transportation Applications, 2013*. 2013.
[195] “US DOT FHWA (United States Department of Transportation Federal Highway Administration), Clear Zones.” Accessed: Oct. 02, 2023. [Online]. Available: https://highways.dot.gov/safety/rwd/provide-safe-recovery/clear-zones/clear-zones#zones
[196] P. Aryal, *Optimization of geometric road design for autonomous vehicles. Master’s Thesis.* Stockholm, Sweden: KTH Royal Institute of Technology, 2020.
[197] K. Othman, “Impact of autonomous vehicles on the physical infrastructure: Changes and challenges,” *Designs*, vol. 5, no. 3, p. 40, 2021.
[198] A. Amini *et al.*, “VISTA 2.0: an open, data-driven simulator for multimodal sensing and policy learning for autonomous vehicles.” arXiv, Nov. 23, 2021. Accessed: May 09, 2023. [Online]. Available: http://arxiv.org/abs/2111.12083
[199] Y. Li and Y. Chi, “Off-the-grid line spectrum denoising and estimation with multiple measurement vectors,” *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016, doi: 10.1109/TSP.2015.2496294.

[200] Z. Yang and L. Xie, “Exact joint sparse frequency recovery via optimization methods,” *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 5145–5157, Oct. 2016, doi: 10.1109/TSP.2016.2576422.
[201] C. Kim, Y. Jung, and S. Lee, “FMCW LiDAR system to reduce hardware complexity and post-processing techniques to improve distance resolution,” *Sensors*, vol. 20, no. 22, p. 6676, 2020.
[202] X. Li, B. Yang, X. Xie, D. Li, and L. Xu, “Influence of waveform characteristics on LiDAR ranging accuracy and precision,” *Sensors*, vol. 18, no. 4, p. 1156, 2018.
[203] V. Kilic, D. Hegde, V. Sindagi, A. B. Cooper, M. A. Foster, and V. M. Patel, “Lidar Light Scattering Augmentation (LISA): Physics-based Simulation of Adverse Weather Conditions for 3D Object Detection.” arXiv, Jul. 14, 2021. Accessed: Jul. 30, 2023. [Online]. Available: http://arxiv.org/abs/2107.07004
[204] Y. Li, P. Duthon, M. Colomb, and J. Ibanez-Guzman, “What happens for a ToF LiDAR in fog?,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 6670–6681, Nov. 2021, doi: 10.1109/TITS.2020.2998077.
[205] X. Li and P. Luo, “Boosting ranging performance of LiDAR using multi-pulse coherent average,” *IEEE Sensors J.*, vol. 19, no. 15, pp. 6270–6278, Aug. 2019, doi: 10.1109/JSEN.2019.2910561.
[206] Y. Zhang, A. Carballo, H. Yang, and K. Takeda, “Perception and sensing for autonomous vehicles under adverse weather conditions: A survey,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 196, pp. 146–177, Feb. 2023, doi: 10.1016/j.isprsjprs.2022.12.021.
[207] B. Rezvani, B. Hassibi, F. Brännström, and M. Manteghi, “Letting robocars see around corners: Using several bands of radar at once can give cars a kind of second sight,” *IEEE Spectrum*, vol. 59, no. 2, pp. 36–41, Feb. 2022, doi: 10.1109/MSPEC.2022.9706401.
[208] M. Ijaz, Z. Ghassemlooy, H. Le Minh, S. Rajbhandari, and J. Perez, “Analysis of fog and smoke attenuation in a free space optical communication link under controlled laboratory conditions,” in *2012 International Workshop on Optical Wireless Communications (IWOW)*, Oct. 2012, pp. 1–3. doi: 10.1109/IWOW.2012.6349680.
[209] O. Kilani, M. Gouda, J. Weiß, and K. El-Basyouny, “Safety assessment of urban intersection sight distance using mobile LiDAR data,” *Sustainability*, vol. 13, no. 16, p. 9259, 2021.
[210] L. Fan, F. Wang, N. Wang, and Z. Zhang, “Fully sparse 3D object detection,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 351–363, 2022.
[211] S. N. Sridhara, E. Pavez, and A. Ortega, “Cylindrical coordinates for lidar point cloud compression,” in *2021 IEEE International Conference on Image Processing (ICIP)*, Anchorage, AK, USA: IEEE, Sep. 2021, pp. 3083–3087. doi: 10.1109/ICIP42928.2021.9506448.
[212] E. Camuffo, D. Mari, and S. Milani, “Recent advancements in learning algorithms for point clouds: An updated overview,” *Sensors*, vol. 22, no. 4, p. 1357, Feb. 2022, doi: 10.3390/s22041357.
[213] H. G. Seif and X. Hu, “Autonomous driving in the iCity—HD maps as a key challenge of the automotive industry,” *Engineering*, vol. 2, no. 2, pp. 159–162, Jun. 2016, doi: 10.1016/J.ENG.2016.02.010.

[214] “Alpha Prime,” Velodyne Lidar. Accessed: Aug. 14, 2023. [Online]. Available: https://velodynelidar.com/products/alpha-prime/
[215] D. Girardeau-Montaut, “CloudCompare,” *France: EDF R&D Telecom ParisTech*, vol. 11, 2016.
[216] “Velodyne HDL-32E | Proven 3D LiDAR sensor that delivers unsurpassed resolution,” Mapix technologies. Accessed: Sep. 21, 2023. [Online]. Available: https://www.mapix.com/lidar-scanner-sensors/velodyne/velodyne-hdl-32e/
[217] Y. Cao, S. H. Bhupathiraju, P. Naghavi, T. Sugawara, Z. M. Mao, and S. Rampazzi, “You can’t see me: Physical removal attacks on {LiDAR-based} autonomous vehicles driving frameworks,” in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023, pp. 2993–3010.
[218] M. Gouda, Z. Pawliuk, J. Mirza, and K. El-Basyouny, “Using convex hulls with octree/voxel representations of point clouds to assess road and roadside geometric design for automated vehicles,” *Automation in Construction*, vol. 154, p. 104967, Oct. 2023, doi: 10.1016/j.autcon.2023.104967.
[219] M. P. Huijser, A. V. Kociolek, T. D. Allen, P. McGowen, P. C. Cramer, and M. Venner, “Construction guidelines for wildlife fencing and associated escape and lateral access control measures,” *NCHRP Project 25-25, Task 84, National Cooperative Highway Research Program. Transportation Research Board of the National Academies, Washington D.C., USA*, 2015.
[220] A. G. Spanowicz, F. Z. Teixeira, and J. A. Jaeger, “An adaptive plan for prioritizing road sections for fencing to reduce animal mortality,” *Conservation Biology*, vol. 34, no. 5, pp. 1210–1220, 2020.
[221] J. Chen, S. Tian, H. Xu, R. Yue, Y. Sun, and Y. Cui, “Architecture of vehicle trajectories extraction with roadside LiDAR serving connected vehicles,” *Ieee Access*, vol. 7, pp. 100406–100415, 2019.
[222] M. Gouda, I. Chowdhury, J. Weiß, A. Epp, and K. El-Basyouny, “Automated assessment of infrastructure preparedness for autonomous vehicles,” *Automation in Construction*, vol. 129, p. 103820, Sep. 2021, doi: 10.1016/j.autcon.2021.103820.

APPENDIX

VOLUMETRIC METHOD CALCULATIONS

The volumetric method is unique in its approach and exploits the spherical coordinate system used for voxelization. This method assigns more weight to objects that have larger volumes. Following the voxelization process implemented in this research, the voxel size increases the further its position is from the sensor. Hence, the biggest voxels will be located at the maximum range of the used sensor. In calculating the percentage of occupied voxels by volume to get the delta, voxels of bigger size will have more influence on the results, meaning that this method directs its attention to distant points rather than points in the vicinity of the sensor.

This approach assumes that the further the point is from the sensor, the harder it would be to process in real time due to the scarcity of points at long ranges. Additionally, since the points close to the sensor are already big in numbers, it would be easier to identify objects from them. Consequently, they will not have much processing power from the onboard computer.

That said, this method can be rather unstable in FE analysis because the scattered residue of roadside elements, like vegetation, in the outskirts of the sensor vision can significantly influence the data rate values while not being necessary for the AV navigations. However, it could used to intelligibly highlight the occlusion problems in the RRFs analysis, witnessed in locations that have vertical or horizontal curves. Although the missing points will be few owing to the ring artifact of the LiDAR sensor, they will have more influence on the anticipated performance of the AV.

Finally, it should be noted that overall lower magnitudes are generally expected in volumetric method calculations when compared to the occupancy method because the number of occupied voxels situated at the far range of the sensor, and consequently possessing larger volumes, is considerably smaller compared to the occupied voxels located closer to the sensor, which tends to have smaller volumes, was discussed. Therefore, it is anticipated that lower data rate values will be achieved on the whole when the volumetric method is employed. On average, it was noticed that the data rate requirements using the volumetric calculations are lower than that of the occupancy method.

Vertical Curves

The volumetric calculations for road sections (1-3) will be demonstrated in this section. For both RRFs and FE considerations, to illustrate how this approach influences different road layouts.

Volumetric Method (RFFs) Vertical Curves Analysis

The volumetric method can better accentuate the differences between sag and crest vertical curves. As seen in the data rate requirements results for the three road sections, the graphs are more defined, and the value difference is more discernable.

As illustrated in Table A 1, the biggest disparity between the values at sag or crest curves and the average value at flat road sections was reported in the first road section, where there is a 40% decrease in data rate values at the crest location and a 42% increase in values at the sag location. For the second and third road sections, the maximum sag locations increased by 2x% and 18.8%, respectively, whereas the minimum crest locations decreased by 16.7% and 22.5%, respectively. The main reason behind the difference in these values is the grade of the vertical curve. The grade of the first road sections is steeper than the other grades, and hence, the degree of increase or decrease in road information is amplified.

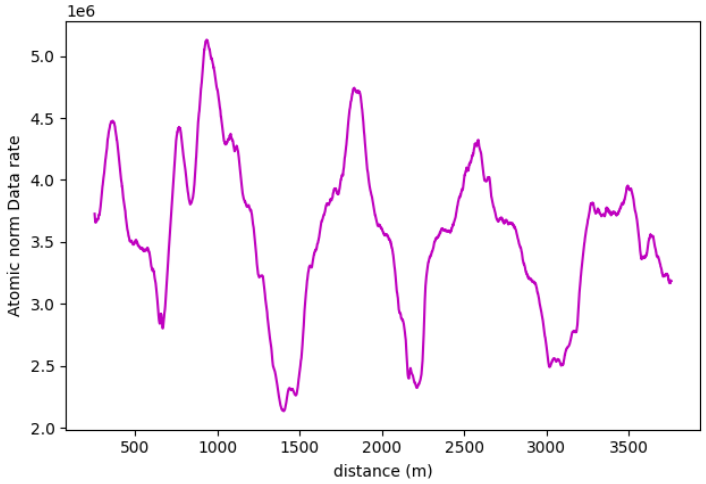

Figure A 1. Data Rate Requirements for Section (1) Using Volumetric Method

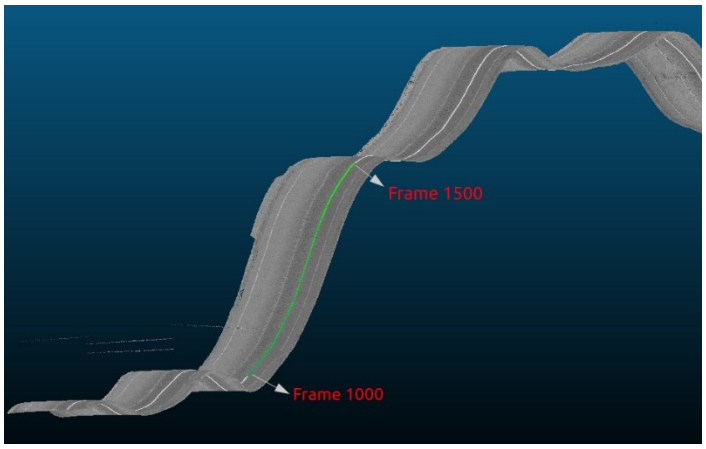

Figure A 2. Road Section (1) Frame 1500 (Lowest Value on Crest Curve) vs Frame 1000 (Highest Value on Sag Curve)

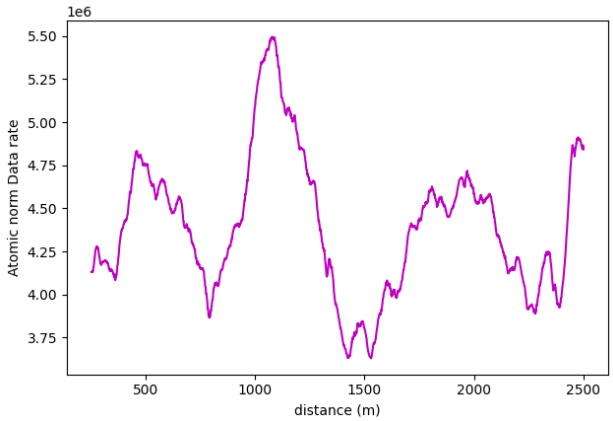

Figure A 3. Data Rate Requirements for Section (2) Using Volumetric Method

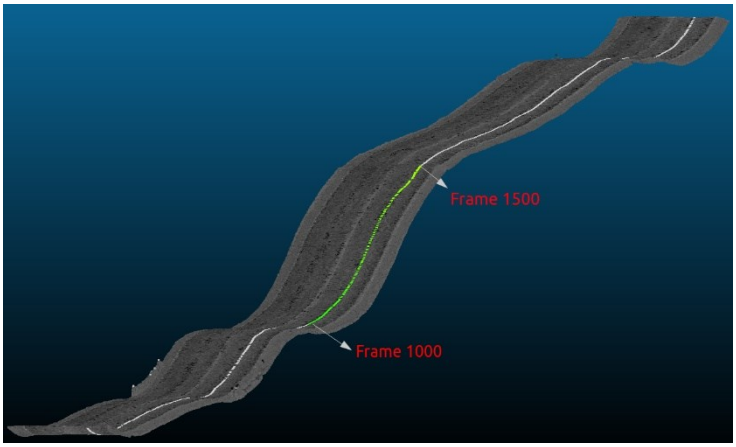

Figure A 4. Road Section (2) Frame 1500 (Lowest Value on Crest Curve) vs Frame 1000 (Highest Value on Sag Curve)

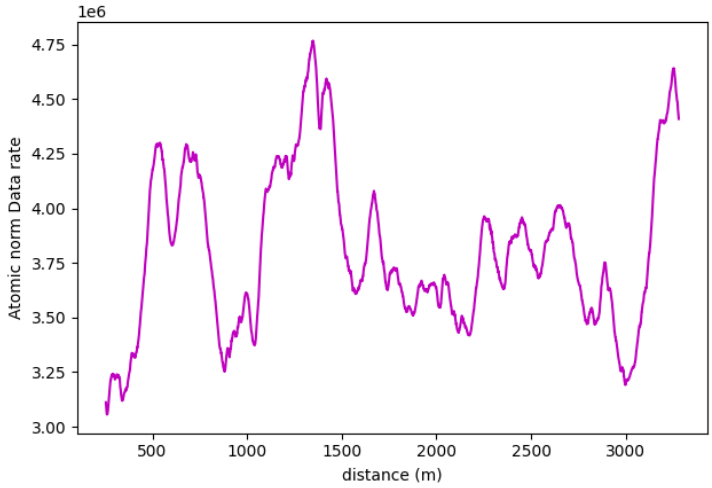

Figure A 5. Data Rate Requirements for Section (3) Using Volumetric Method

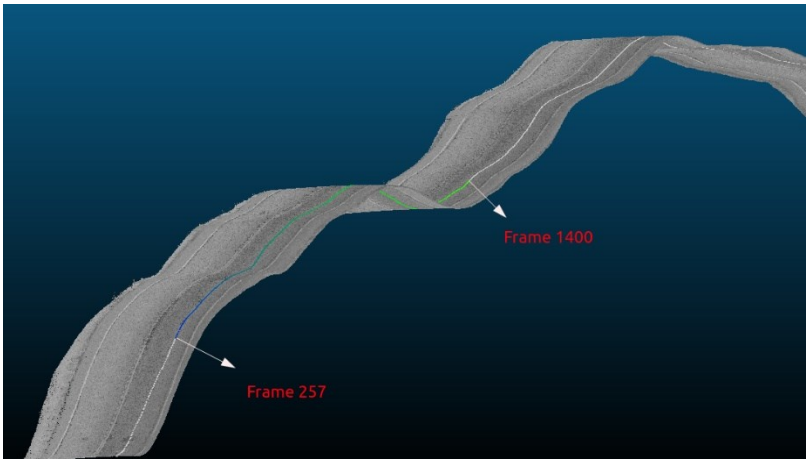

Figure A 6. Road Section (3) Frame 257 (Lowest Value on Crest Curve) vs Frame 1400 (Highest Value on Sag Curve)

Table A 1. Volumetric Method RRFs Change in Data Requirements for Vertical Curves.

Section	Min Value (Crest)	Average Value (Flat)	Max Value (Sag)	% Change Relative to Flat Values
1	2.1×10^6	3.5×10^6	5.0×10^6	-40% → +42%
2	3.75×10^6	4.5×10^6	5.5×10^6	-16.7% → +22%
3	3.1×10^6	4.0×10^6	4.75×10^6	-22.5% → +18.8%

Volumetric Method (FE) Vertical Curves Analysis

As previously outlined in Section 4.2.2, the data rate requirements in the FE analysis are highly dictated by the presence of roadside features. Additionally, since the volumetric method is already sensitive to the density and distribution of points in the point cloud, it is expected that the geometry of the road will have little control over the data rate values. Furthermore, comparing the volumetric calculations of the RRFs analysis with the FE analysis, the jump in data rate requirements is expected to be much higher than that witnessed in the occupancy method calculations. The maximum increase in data rate requirements as per the volumetric method calculations, as listed in Table A 3, are 1300%, 1950%, and 1475% for the three road sections, respectively, meaning that the data rate values will increase by an average of 16.7 times in case the AV attempts to process the entire environment. This increase is much higher than that recorded in the occupancy method calculations, shown in Table 6, which averaged only 4.8 times more processing power.

As for the established pattern for the data rate values in this analysis, it was observed that it is still visible at multiple locations throughout the road sections. However, some deviations still emphasize the control of roadside features in this type of analysis. For example, road section (1) had similar patterns as in the RRFs analysis, unlike sections (2) and (3), where more deviation has been noticed in accordance with the different distribution of vegetation. Nevertheless, since this method puts more emphasis on the far objects rather than the closer ones, it is natural that its results for vertical curves are consistent regardless of the road layout, as in the FE analysis, the crest locations will naturally have lower point densities at the far end in contrast to sag locations which will help the sensor capture more points at long distances.

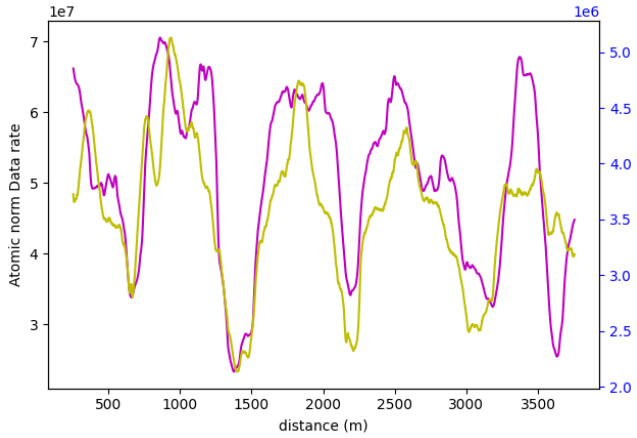

Figure A 7. Road Section (1) Volumetric Method FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

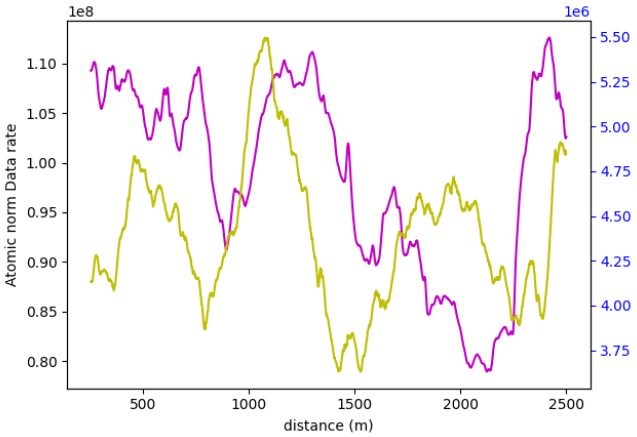

Figure A 8. Road Section (2) Volumetric Method FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

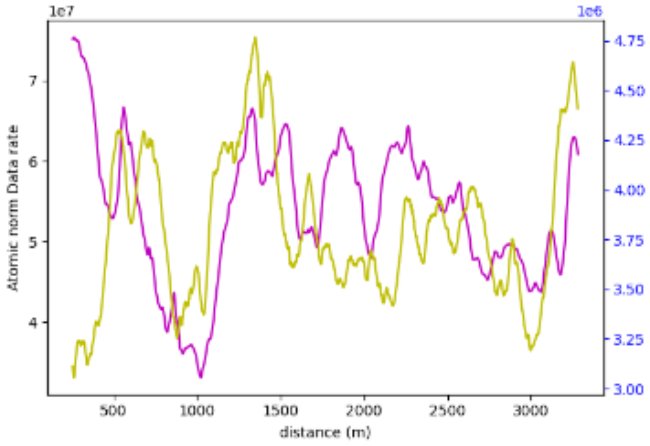

Figure A 9. Road Section (3) Volumetric Method FE (Purple + Left Axis) vs RRFs (Yellow + Right Axis)

Table A 2. Volumetric Method FE Change in Data Requirements for Vertical Curves.

Section	Min Value (Crest)	Average Value (Flat)	Max Value (Sag)	% Change Relative to Flat Values
1	2.5×10^7	5.2×10^7	7.0×10^7	-51% → +34%
2	0.8×10^8	0.95×10^8	1.1×10^8	-15.8% → +15%
3	3.5×10^7	5.8×10^7	7.5×10^7	-39.5% → +29%

Table A 3. Vertical Curves Volumetric Method Calculations for FE and RRFs Analysis.

Section	Max RRFs Values	Max FE Values	% Increase
1	5.0×10^6	7.0×10^7	1300%
2	5.5×10^6	1.1×10^8	1950%
3	4.75×10^6	7.5×10^7	1475%

Roadway width

The analysis of the variation of the roadway width using the volumetric method while considering RRFs and FE will be demonstrated in this section.

Volumetric Method (RRFs) Roadway Width Analysis

Using the volumetric method approach in quantifying the expected increase in data rate values upon increasing the number of lanes for road section (4) is illustrated in Figure 133. The same pattern of the jump in the data rate values in the region where the number of lanes is increased to three instead of two was witnessed again. The average data rate value in the two-lane region was $4.4x10^6 bits/s$ and that at the three-lane region was $5.3x10^6 bits/s$, meaning that on average, there is a 20% increase in the data rate values.

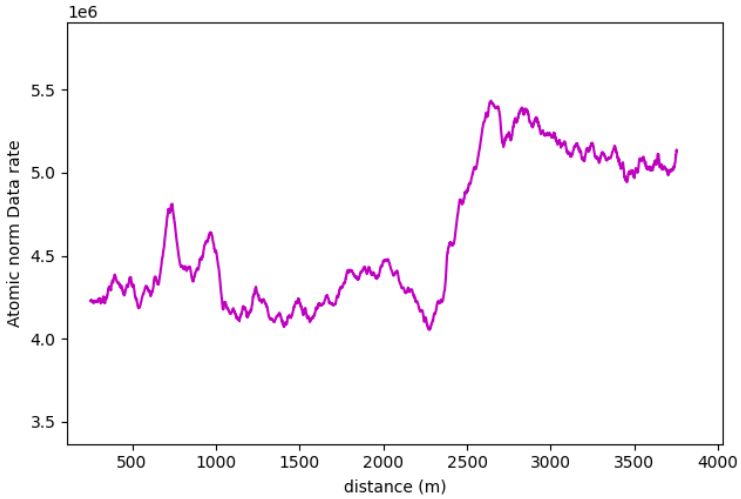

Figure A 10. RRFs Data Rate Requirements for Road Section (4) Using Volumetric Method

Volumetric Method (FE) Roadway Width Analysis

In the FE analysis, the established pattern has been significantly attenuated due to the masking effect of the roadside features. As seen in **Error! Reference source not found.**, the difference between the 2-lane and 3-lane regions is somewhat obscured, unlike in the RRFs analysis. However, except for frame 1000, where the roadside features had much higher density than the rest of the road, the 3-lane zone still reports about 30% more data rate values than its 2-lane zone counterpart. Figure A 12 and Figure A 13 showcase the difference in the surrounding environment between the highest and lowest frames in the road section. More importantly, on comparing the percentage increase in data rate values if the roadside elements were to be considered in the

analysis, about 600-700% increase is expected using the volumetric method calculations, which translates to around an 8-fold increase in the data rate values.

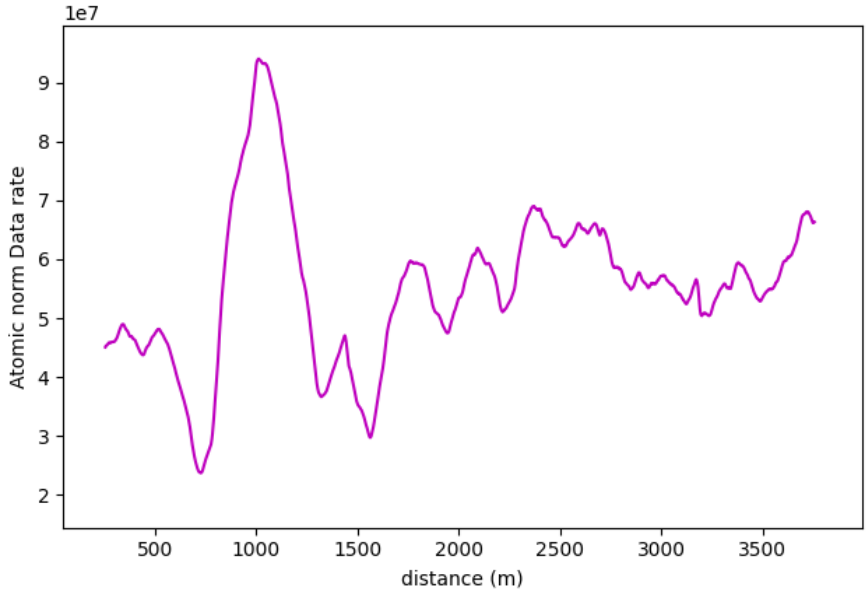

Figure A 11. FE Data Rate Requirements for Road Section (4) Using Volumetric Method

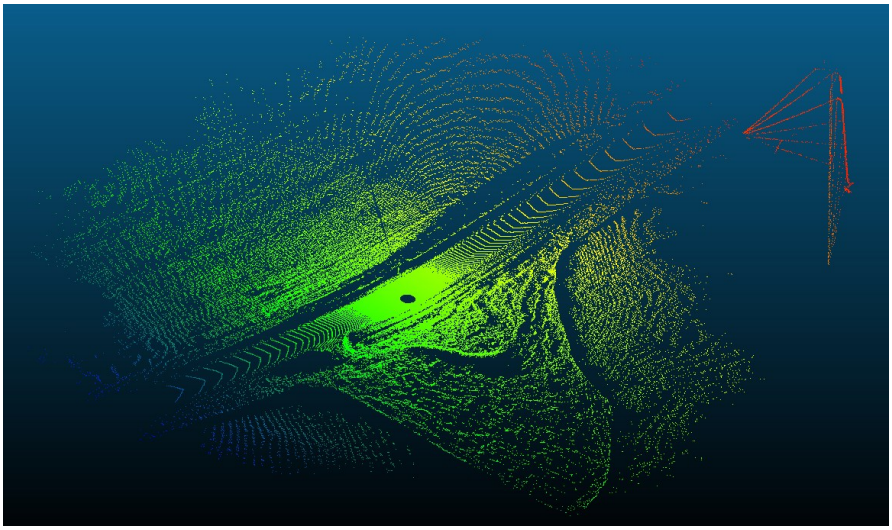

Figure A 12. Road Section (4) Frame 1000 (Highest Data Rate Value) VISTA Outputs

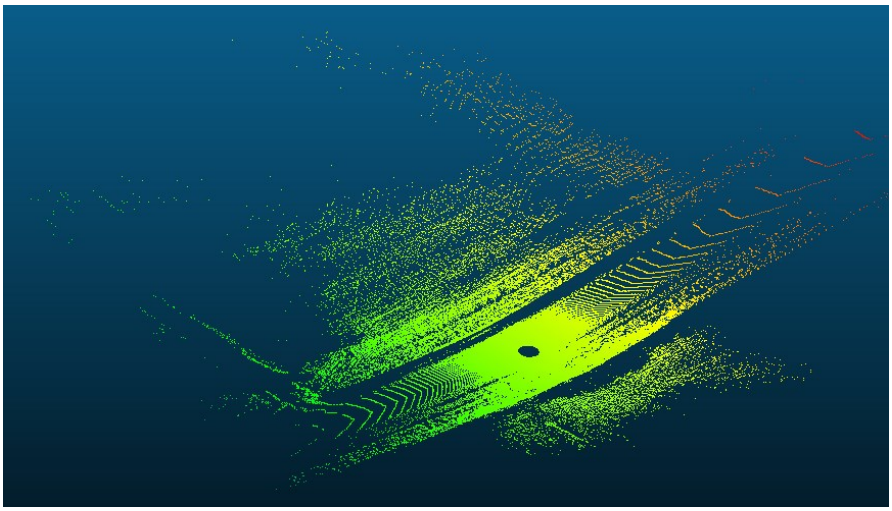

Figure A 13. Road Section (4) Frame 750 (Lowest Data Rate Value) VISTA Outputs

Horizontal Curves Analysis

Given that this analysis is being conducted on the entire environment surrounding the AV, it is essential to note that a significant influence on the data rate results will be exerted by stray points in the cloud that pertain to various roadside elements, primarily vegetation. Due to the inconsistency of the roadside features and their continuously changing density, the volumetric method is expected to yield erratic data rate values and experience high fluctuations. Nonetheless, the ability to capture shifts in data rate values at critical horizontal curves is still retained. To demonstrate this capability, the data rate values for sections 8 and 9 are presented in the figures below.

In road section (8), the previously observed patterns in the occupancy method calculations at the entrance and exit of the horizontal curve are once again visible in the volumetric method calculations. Notably, there is a 127% drop in values at the entrance of the curve, while the drop in values at the exit of the curve is more dramatic, with an approximate decrease of 491% since the road environment ahead of the curve was very rich in roadside elements. Similarly, in road section (9), a similar trend around the horizontal curve is observable, with a significant drop in values of 367% at the entrance of the road section and a 33% decrease at the curve's exit.

The primary factor controlling the resulting data rate values is the roadside elements, which is apparent when comparing the Vista outputs of frames 800 and 1000, as shown in Figure A 16 and Figure A 17, which represent the lowest and highest data rate values in section (8). The

distribution and density of points between these two frames explain the stark contrast in values. Frame 1000 contains a greater number of points than frame 800, and more importantly, there is an increased concentration of points at the extremes of the frame, resulting in higher data rate values due to greater contributions from large voxels. The same pattern can also be explained in road section (9) from the Vista outputs, depicted in Figure A 18 and Figure A 19, of frames 600 and 1000 to show the difference between the lowest and highest values. Finally, from the listed figures, it is apparent that frame 600 of road section (9) will have the highest data rate values in both road sections, judging from the density and distribution of its points.

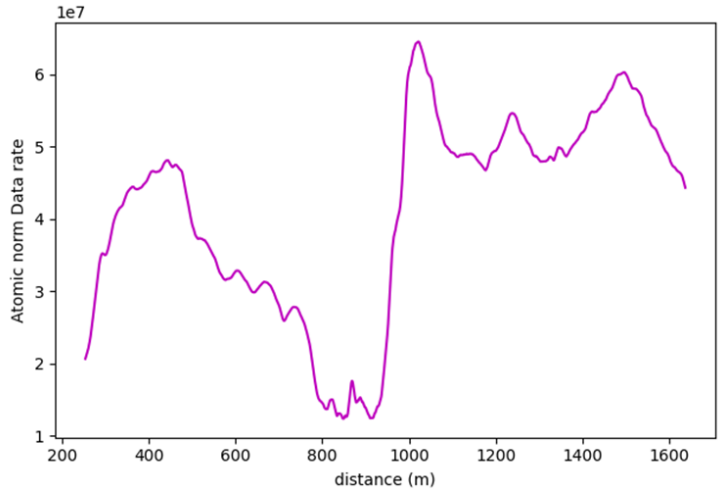

Figure A 14. Data Rate Requirements for Road Section (8) Using Volumetric Method

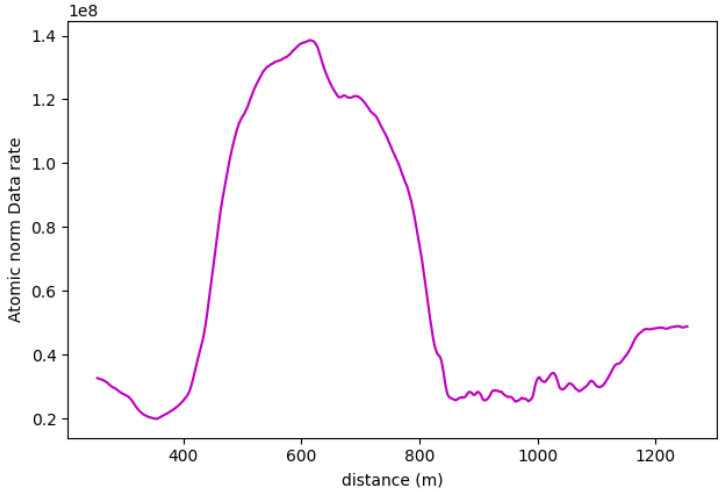

Figure A 15. Data Rate Requirements for Road Section (9) Using Volumetric Method

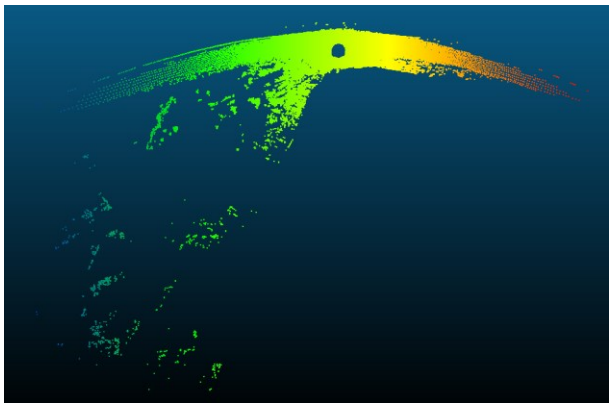

Figure A 16. Section (8) Frame 800 VISTA Outputs

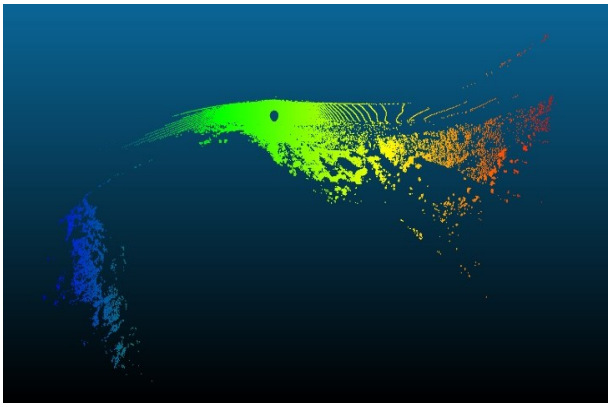

Figure A 17. Section (8) Frame 1000 VISTA Outputs

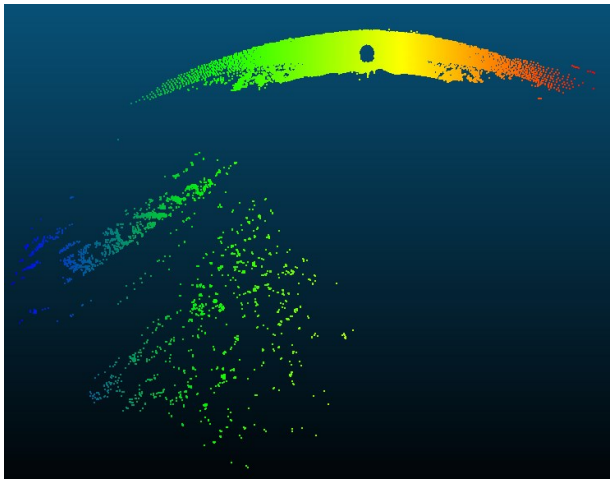

Figure A 18. Section (9) Frame 1000 VISTA Outputs

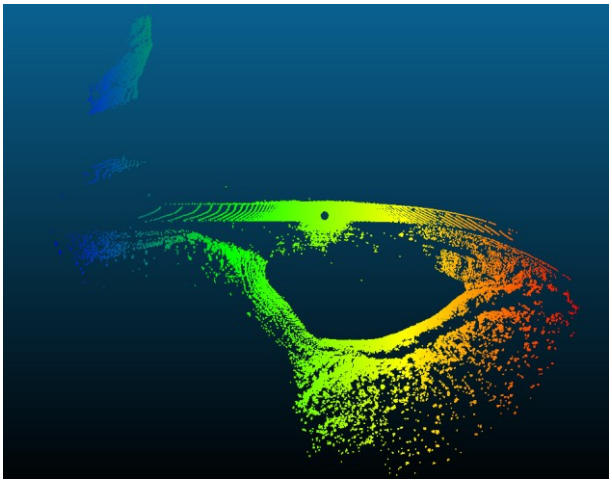

Figure A 19. Section (9) Frame 600 VISTA Outputs

MODIFIED HORIZONTAL CURVE ANALYSIS

Since the goal is to ultimately isolate the horizontal curves themselves to better understand their impact, this analysis attempts to attenuate the influence of the roadside elements on the data rate values and get a better estimation of the contribution of horizontal occlusions. An investigation was conducted where the roadside features were kept only on the side of the road that caused the occlusions. To illustrate the change in the values, Road sections (8) and (9) were reanalyzed based on the aforementioned criteria. The previously recorded shift in values for section (8) of 15 and 16% at the horizontal curve's entrance and exit, respectively, have now increased to 21 and 29%.

Similarly, the change in values for section (9) is now 23 and 28% instead of 12.2 and 12.7% for the entrance and exit of the curve.

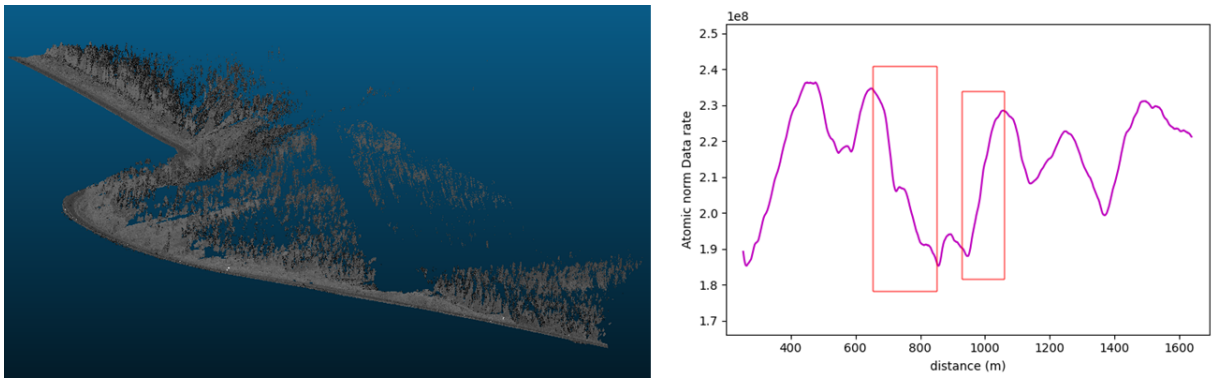

Figure A 20. Road Section (8) Modified Section Data Rate Requirements

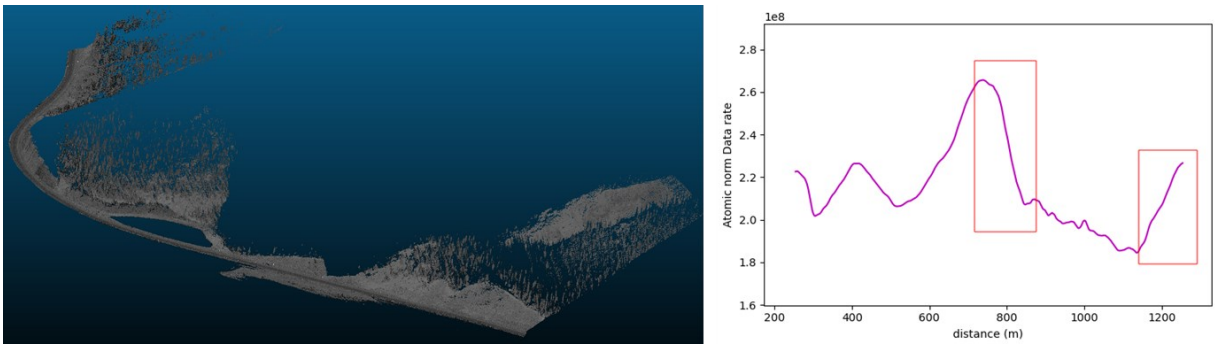

Figure A 21. Road Section (9) Modified Section Data Rate Requirements