University of Nevada, Reno

Study of retrofitted system for Intelligent Compaction Analyzer, a Machine Learning approach for Quality Control of Asphalt Pavement during Construction

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in
Computer Science and Engineering

by
Shankar Poudel

Dr. Sesh Commuri and Dr. George Bebis - Thesis Advisors

August, 2023

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

Shankar Poudel

entitled

Study of retrofitted system for Intelligent Compaction Analyzer, a Machine Learning approach for Quality Control of Asphalt

be accepted in partial fulfillment of the
requirements for the degree of

Master of Science

George Bebis, Ph.D
Advisor

Sesh Commuri, Ph.D.
Co-advisor

Alireza Tavakkoli. Ph. D.
Committee Member

Jeongwon Park, Ph. D.
Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean
Graduate School

August 2023

ABSTRACT

Asphalt pavements play a vital role in transportation infrastructure, but their performance can suffer due to subpar quality resulting from improper construction practices. To tackle this issue, we introduce the Retrofit Intelligent Compaction Analyzer (RICA), a real-time compaction density estimation system for asphalt pavements during construction. RICA utilizes machine learning principles and machine learning to predict compaction density based on received vibratory patterns at different compaction levels. By leveraging the roller's spatial location and analyzing vibration patterns, RICA delivers density estimates.

In this study, we gathered data from actual construction sites, implementing RICA on a Caterpillar CB-10 Rotary dialed dual drum vibratory compactor. The density estimates from RICA were validated against densities measured from roadway cores extracted randomly on the compacted pavement. Our findings affirm the efficacy of RICA in providing reliable density estimates for asphalt pavements.

The ability of RICA to provide real-time, nondestructive compaction information to the roller operator establishes its value as a quality control tool during asphalt pavement construction. By ensuring proper compaction, RICA contributes to the construction of durable, high-quality roads while reducing the financial and environmental costs associated with construction and maintenance. The validation of RICA's estimates with percent within limits (PWL) calculations based on roadway cores further attests to its effectiveness as a Quality Assurance tool.

Keywords: Intelligent asphalt compaction analyzer, density estimation, machine learning, compaction quality, nondestructive testing, quality assurance

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors, Dr. George Bebis and Dr. Sesh Commuri, for their invaluable patience and constructive feedback throughout this research journey. Their guidance and expertise have been instrumental in shaping this thesis.

I am also grateful to my defense committee for their valuable insights and contributions to this work. Their knowledge and expertise have been crucial in refining the ideas presented in this thesis. I extend my appreciation to my friends, lab mates, librarians, and research assistants at the university. Their support, encouragement, and inspiration have been invaluable during various stages of this work.

Special thanks go to George Reed Inc. for their generous support in facilitating this research. Their provision of sites and the necessary equipment, including the rollers for field data collection, has been instrumental in the success of this project.

I am also thankful to Dr. Garrett Winkelmaier for his feedback sessions, moral support, and encouragement throughout the research process. I would not have been able to complete this endeavor without his help and assistance.

Lastly, I would like to acknowledge my family, my parents, and siblings, for their unwavering belief in me. Their love and encouragement have been a constant source of motivation throughout this journey.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Problem Statement

A significant portion of the paved roadways are asphalt pavements and serve as a vital element of transportation infrastructure. The USA has more than 6.6 million kilometers of asphalt paved road whereas the EU has more than 5 million kilometers of asphalt paved road system[1]. Hot mix asphalt (HMA) pavements are designed to maintain their performance throughout their lifespan, even when subjected to various traffic and weather conditions. The quality of the pavement is influenced by factors such as proper mix design, selection of aggregates, and the asphalt binder. However, the ultimate quality of the finished pavement relies heavily on the construction practices employed and the implementation of quality control procedures during the construction phase. It is crucial to emphasize that the desired performance of a well-designed asphalt mix can only be achieved if it is effectively compacted in the field. In the past two decades, significant efforts have been made to establish quality measures that ensure the control of the compaction process and verify compliance with specifications for all aspects of asphalt production and placement.

The escalating volume of vehicles and trucks, coupled with increasingly extreme weather conditions, places an overwhelming burden on road capacity, straining its ability to match the growing demand. Additionally, inadequate construction practices contribute to the rapid degradation and failure of asphalt pavements. In the realm of HMA pavements constructions, the critical factors for maintaining

quality are—Asphalt mix design, Gradation Analysis, and Amount of Compression. The first two are meticulously addressed prior to asphalt mix creation in the HMA production factory. However, achieving proper compaction, a key element, must occur in the field during construction. This underscores the urgent need for cutting-edge quality control tools capable of providing real-time compression analysis for HMA throughout pavement construction, not only reducing financial and environmental costs but also establishing durable roadways.

To ensure construction quality, quality control (QC) and quality assurance (QA) practices are paramount [2]. QC is pivotal for overseeing construction quality and evaluating the final product's integrity, with paving contractors in charge. Conversely, QA encompasses actions that endorse construction quality, often handled by state agencies like Departments of Transportation (DOTs). As construction dynamics evolve, the demand for innovative technology grows stronger, enabling real-time, comprehensive compression analysis for quality assurance. By closing this technological gap, we bolster construction quality, efficiency, cost-effectiveness, and environmental sustainability in one bold stride.

Compaction refers to the procedure of eliminating air voids within the HMA in the pavement, resulting in a reduction in volume and an increase in the unit weight of the mixture [3] [4]. This process also enhances the interlocking between the aggregate particles, promoting greater structural integrity. Compaction level is observed as the most significant determining factor in dense graded pavement performance [5] [6] [7] [8]. Insufficient compaction leads to several negative outcomes for pavements, including diminished stiffness, shortened fatigue life, expedited

aging and reduced durability, rutting, raveling, and increased vulnerability to moisture damage. An adequate amount of compaction in asphalt pavement is necessary to gain its designed properties and to ensure its long-term performance. Failure to reach the desired density, i.e., under or over-compaction, could result in pavement distresses during service such as rutting, raveling, cracking, and moisture damage[9]. Several research studies have highlighted the impact of air voids on various pavement properties.

Kennedy et al. (1984) [10] conducted a study and found that high air void content leads to decreased stiffness and strength in the pavement. Tensile strength, static and resilient moduli, and stability were observed to be reduced at high air void levels.

Fatigue life is significantly affected by air void content. Pell and Taylor (1969) [11], Epps and Monismith (1969) [12], and Linden et al. (1989) [13] reported a negative relationship between increased air voids and reduced fatigue life. Finn et al. (1973) [14] concluded that fatigue properties can be reduced by 30 to 40 percent for each one percent increase in air void content. Additionally, Scherocman (1984) [6] found that reducing air voids from eight percent to three percent could more than double pavement fatigue life.

Air void content also influences the aging and durability of the pavement. McLeod (1967) [15] stated that compacting a well-designed paving mixture to low air voids retards the rate of hardening of the asphalt binder, leading to longer pavement life, lower maintenance, and improved overall performance.

Raveling, which is the disintegration of the pavement surface, becomes a significant problem above approximately eight percent air voids and turns severe above 15 percent air voids, as noted by Kandhal and Koehler (1984) [16]. Rutting, characterized by permanent deformation and deformation accumulation in the pavement, is inversely proportional to the air void content (Scherocman, 1984 [6]). Rutting can be caused by vertical consolidation and lateral distortion, both of which can occur more rapidly if the HMA (Hot Mix Asphalt) air void content is too low.

Moisture damage is influenced by the presence of interconnected air voids in the HMA due to insufficient compaction. High and interconnected air voids facilitate easy water entry, increasing the likelihood of significant moisture damage (Kandhal and Koehler, 1984 [16]; Cooley et al., 2002 [17]). The relationship between permeability, nominal maximum aggregate size, and lift thickness plays a crucial role in moisture damage susceptibility and can vary significantly with changes in these parameters.

Thus, one of the primary attributes of interest for QA/QC is the air volume present within the compacted pavement, commonly measured as a percentage of air voids in relation to the total volume and denoted as "percent air voids". This calculation involves comparing the density of a test specimen with the density it would hypothetically possess if all the air voids were eliminated, known as the "theoretical maximum density" (TMD) or "Rice density. While the focus of interest in HMA QC/QA is typically on the percentage of air voids, measurements are commonly presented in terms of measured density relative to a reference density.

This is achieved by expressing the density as a percentage of TMD (or "percent Rice"). This density representation allows for straightforward conversion to air voids since any volume not occupied by asphalt binder or aggregate is assumed to be air. Exceeding a high-end threshold value of air void content in pavement leads to a predictable reduction in pavement service life, typically identified as 7-8%. According to Linden et al. (1989) [13], each 1 percent increase in air voids above a base level of 7 percent results in approximately a 10 percent loss in pavement life, equivalent to about 1 year less. Conversely, an air void content below a low-end threshold value is indicative of an unstable mixture prone to distortion and flushing. This threshold is commonly reported as around 2-3%.

Compaction density, asphalt content, and aggregate gradation are the three commonly controlled characteristics during QA/QC in USA.[18]. Among these, asphalt content and aggregate gradation are controlled in the asphalt mix production laboratory but the density is to be obtained from the field during construction. During the construction of asphalt pavement, vibratory compaction rollers are used to increase the compaction. Static force, the weight of the rollers, and dynamic force, vibration applied in the compactors, both are used to squeeze out the air voids in the pavement and increase interlocking between the pavement materials. The compaction that is achieved is highly dependent on the underlying base, the formulation of the asphalt mix used, the thickness of the pavement, and environmental conditions like rain, fog, etc. at the time of placement. In addition, the final density of the pavement is also affected by the type of rollers used and the rolling patterns used during the compaction process [19].

As of now, the prevailing approach to gauge compaction density levels in asphalt pavement construction is coring. This method employs a vacuum sealing technique, entailing the extraction of roadway cores from the finalized pavement to assess its density. While this density measurement from the cores provides the most accurate indication of quality, it does present three major shortcomings. Firstly, coring is a destructive quality check measure, making it a potential source of decreased pavement quality. Secondly, the compaction density values are obtained after a long time because the core needs to be taken to the lab able to apply vacuum sealing technology. Consequently, by the time density prediction values become available, the asphalt pavement is already dense and stiffened, making rectification costly and complex, as the only option is to remove and rebuild the pavement. The final flaw is that this mechanism relies on spot checks. Most organizations responsible for maintaining pavement quality, such as departments of transportation, mandate taking cores every 500 tons of asphalt mix, amounting to about 2 cores per km in a 2.5 cm pavement. However, given that pavement density is influenced by numerous factors that can vary significantly within short distances, this spot-check approach may not always accurately represent the overall quality of the entire pavement.

Another widely used mechanism used in the quality control of Asphalt Pavement is by using pulsing the radiation into the pavement and studying the backscattering called Nuclear Density Gauge (NDG). During measurements, a radioactive source (e.g., cesium), emits gamma rays to the pavement surface. A fraction of the radiated energy reflects and collected by the gauge. The amount of

back-scattered radiation energy is correlated to in situ asphalt pavement density based on lab test calibration [20]. In comparison to the core collection method, NDG density measurement is non-destructive and as its prediction is obtained before asphalt settles it can be used to fix the compression errors, if present, in almost real-time as it gives results in about 5-8 minutes for each location. NDG has been used as a non-destructive alternative to core-taking mechanisms that provide in-situ density measurements [21], however, it requires specialized training for operators, and its prediction accuracy could be lower than that of the core method, and as it uses radioactive substance which in long-term use can be harmful to the operative.

Apart from the aforementioned techniques, various studies have proposed alternative quality control methods, including the use of Fibre Bragg grating technology (Yiqiu et al., 2014 [[22]]), Portable Seismic Pavement Analyzer (PSPA) [23], Ground Penetrating Radar (GPR)[24], Non-nuclear Density Gauges (NNDG) [25], etc. However, these approaches often encounter common challenges, such as the inability to provide real-time density measurements during construction, limited accuracy, or reliance on destructive testing procedures. Consequently, there is a critical demand for the development of new technologies that can overcome these drawbacks and offer enhanced solutions for quality control in pavement construction. Intelligent Compaction (IC) emerges as a promising candidate to meet this demand.

1.2 Literature Review

Intelligent compaction (IC) is a construction technique that involves the use of specialized machinery equipped with sensors and other technology to monitor and control the compaction process in real time. In 1994, J.Y. Richard Liao and J. David Powell presented a paper titled "Intelligent Compaction: A New Approach to Asphalt Pavement Construction and Evaluation" [26] at the Transportation Research Board's 73rd Annual Meeting. This is the first time terminology Intelligent Compaction(IC) is used. This paper described the theoretical basis for the development and testing of an intelligent compaction system for soil and asphalt pavement construction and discussed the benefits and potential applications of the technology. There is no concrete method described here but states what is necessary for a technology to be IC. Since then, a lot of research has been done on these types of analyzer tools but very rarely has any IC system been used as a major QA/QC tool in construction sites.

In their study [27], Robert V. Rinehart and Michael A. Mooney investigated the development of instrumentation for eccentric mass-based vibratory roller compactors during Earthquake Compaction. Their research focused on studying the vibratory behavior of these compactors. By attaching accelerometers to the body of the compaction roller, they were able to measure the vibratory response and analyze changes in its characteristics.

The majority of the existing Intelligent Compaction (IC) specifications are designed for soil compaction and may not be suitable for asphalt compaction.

The Intelligent Compaction Measurement Value (ICMVs) is a generic term for accelerometer-based measurement systems instrumented on vibratory rollers as a key component of intelligent compaction systems[28]. The state-of-the-art ICMVs like Compaction meter value (CMV), Compaction Control Value (CCV), Vibratory modulus, Roller integrated stiffness, etc are designed using the soil IC mechanisms. During 2012-2014, Federal Highway Administration (FHWA) conducted a study examining the relationship between pass-by-pass Nuclear Density Gauge (NDG) density readings and ICMV data. The project's report [29] revealed that ICMV data correlates well with NDG readings during breakdown compaction when the temperature of the mix is high. However, it was found that ICMV does not show a strong correlation with core densities that are cut after the asphalt is totally settled. Despite this, the report concluded that IC technology can be effectively utilized for method-based acceptance, such as roller pass counts and coverage. Moreover, the report suggested that an IC-based model, calibrated with pass-by-pass NDG measurements and core density data from a test strip of a specific project, can generate predicted density values alongside other existing IC measurements. This integration can significantly enhance Quality Control (QC) during production compaction, thus improving the overall compaction process.

The relationship between the ICMV and on-spot cored density is not uniform in the literature. Studies conducted on the relationship between the ICMVs and on-spot densities show that there is no acceptable correlation between them [29] [30] [31]. Whereas other states that there is a consistent trend observed between the in-place density and ICMV value measured by presenting the IC technique [32]

[33]. So, instead of searching for the relationship between ICMV and on-spot core densities, a tool that can directly estimate density can be a better option.

Commuri and Mai in 2008 [34] demonstrated that the dynamic response of the roller-asphalt interaction during compaction can be related to the stiffness of the HMA pavement. Vibratory compaction enhances the stiffness of the pavement layer, leading to an alteration in the vibratory response of the compactor. Utilizing this approach, Commuri S. et al. [35] [36] introduced a two-step IC system referred to as the intelligent asphalt compaction analyzer (IACA). This system initially analyzes the frequency spectrum of the roller and subsequently employs a neural network for classification. The system was then trained to convert these vibration levels into a "number" indicative of the asphalt mat density at a given location. In this approach, at first Fourier transform is applied to the continuous vibration signal in the time domain into the frequency domain. The output (i.e., the power in decibels of the signal at different frequencies) was passed to the neural net to classify the vibration. Finally, the compaction analyzer post-processed the output of the NN and predicted the degree of compaction in real-time. Subsequently, in [37], performance evaluations of the IACA through case studies were conducted, demonstrating its capacity to enhance compaction uniformity when the tool is provided to the roller operators as the instant level of density check mechanism. IACA is one unique approach in the market that can directly estimate density instead of ICMVs.

The utilization of the pavement layer's characteristics and the vibration spectra of the compactor, as highlighted in [38], enables the estimation of asphalt mat

compaction. Zheng et al. [39] investigated the impact of excitation amplitude-frequency and soil parameters on the vibratory drum's response, leading to effective control of jump vibration and optimization of construction parameters for enhanced compaction quality and efficiency. Moreover, in 2019, Wan et al. [40] explored the nonlinear dynamics of the interaction between asphalt and the screed during compaction, revealing variations in resonance frequencies and density increments rates for different asphalt mixtures. Consequently, a universal IC model for estimating asphalt compaction density across all mixtures is unlikely, necessitating separate models for different asphalt mixes.

Field compaction of pavement is always executed by a different combination of static rolling, vibratory rolling, and pneumatic-tire rolling, while laboratory compaction of the mixture is conducted by various compactors, such as Marshall impactor, static compactor, vibratory compactor, and gyratory compactor. All these different roller compaction modes result in varying mechanical performance[38]. So, different rolling systems need different technology to implement IC in them.

Over the last decades, the influence of rolling temperature during construction on the mechanical strength of asphalt materials has been well-established [41] [42] [43]. Notably, Rukavina et al. [44] demonstrated the impact of temperature on asphalt viscosity, directly correlating it with the material's mechanical stiffness. As a result, any Intelligent Compaction (IC) system aiming to learn and predict compaction levels should include different models tailored for distinct temperature ranges, or at least different for breakdown and intermediate rollers even in the

same project with the same asphalt mix.

Masad et al. (2016) presented a novel constitutive model within a three-dimensional finite element system for computer-aided pavement analysis, specifically focusing on simulating hot mix asphalt field compaction [45] [46]. This model was designed based on nonlinear viscoelasticity theory to handle large deformations that occur during the compaction process. In their research, Masad et al. also explored compaction characteristics by using varying vibration frequencies to estimate density, establishing a correlation between frequency and compaction density. The finite element simulation results showed a satisfactory agreement with the actual measurements of the level of compaction and percent air voids.

In 2013, Beainy F., Commuri S., et al. [47] proposed a model to represent the interaction between the roller and asphalt called Viscoelastic-Plastic(VEP) Model. VEP is based on Burger's model [48] which defines roller-pavement interaction as a coupled system. VEP is used to represent the dynamic properties of the asphalt pavement. Using numerical simulation of the VEP model it is also shown that the response of the coupled system can be used to study the compaction of asphalt pavements. During field compaction, the vibratory roller and the underlying asphalt pavement layers form a coupled system. Thus, any changes in the stiffness and density of the asphalt mat would affect the vibratory response of the roller[49]. So, considering a limited amount of temperature change, for a given asphalt-concrete mix and for given constant roller properties like frequency and amplitude of compactor roller, the major change in the vibratory response of the roller is due to the change in stiffness of the pavement. Vibratory compaction enhances

the pavement layer's stiffness, modifying the compactor's vibratory response. The characteristics of the pavement layer and the compactor's vibration spectra can be used to estimate asphalt mat compaction.

1.3 Scope and Novelty

We are undertaking the development of an Intelligent Compaction (IC) system, which can be retrofitted into a dual-drum vibratory roller equipped with an eccentric mass. Our investigation is based on existing literature and background knowledge, which has indicated that the vibratory pattern varies with changes in compaction density. Consequently, we hypothesize that analyzing the vibratory patterns will enable us to detect different levels of compaction density.

To achieve our goal, we have created models capable of learning the various vibration patterns using unsupervised learning techniques, such as clustering. As part of our novel approach, we extract specific characteristics from the rolling patterns to establish a relationship between the vibration pattern and the compaction level. Subsequently, we convert the different compaction levels into density values using function fitting.

The study involves several phases, beginning with the creation of a retrofitting system designed to gather data from the roller. This entails the integration of sensors and processing units to capture information such as vibration, temperature, and location. Additionally, a dedicated software solution was devised for the swift acquisition and analysis of data, constituting a substantial advancement within this project.

The focus then shifted to the analysis phase, where we processed the collected data and formulated various attributes related to vibration patterns and rolling patterns from the raw data. Employing various clustering techniques, we identified

different vibration patterns present in the data. We established a conversion between the clusters and their corresponding density levels, leading to valuable insights into the compaction process.

Our research culminated in an innovative density prediction algorithm, which efficiently predicts density at each time step. Through systematic analysis and rigorous testing with multiple parameters, we have determined the effectiveness of the system in estimating compaction density.

The research can be broadly divided into two major steps:

i) Field data collection: This phase involved designing and implementing a mechanism for retrofitting sensors and processing units into the roller. The hardware configuration and specialized software allowed us to collect vibration, temperature, and location data.

ii) Analysis and Prediction: In this step, we processed the collected data, engineered essential features, and employed clustering techniques to identify different vibration patterns. The clusters were then associated with corresponding density levels through conversion. The density prediction algorithm was used to successfully estimate compaction density at each time step.

Overall, our research study involved meticulous data collection and processing, the development of innovative engineering features, and the establishment of a valuable relationship between vibration patterns and compaction density. Our results imply a significant accomplishment in the field of IC and have the potential to enhance compaction processes in asphalt paving projects.

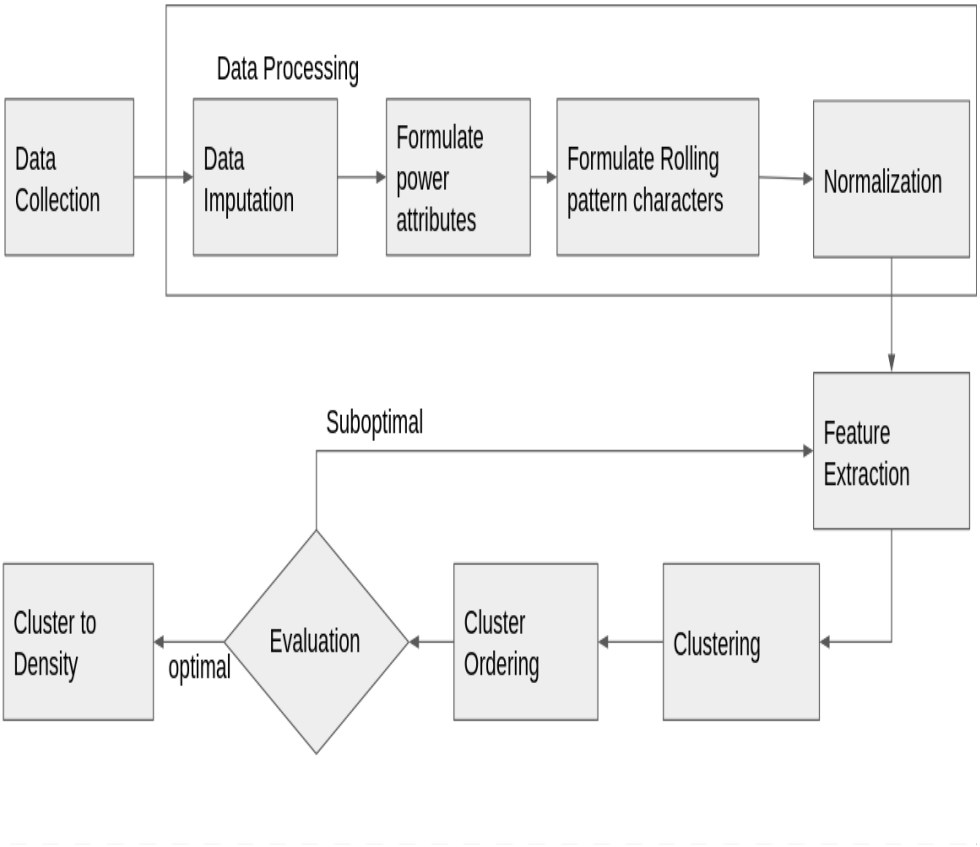


Figure 1.1: Flowchart of basic steps taken in thesis study

1.4 Organization

In the Introduction section, we presented the problem statement, review the relevant literature, and outline the scope of this research. The rest of the thesis is organized as follows:

Chapter 2: Data Collection

This chapter describes the process of data collection for the research. It discusses the instrumentation used, the retrofit system designed for the data collection on the impulse vibratory roller, the structure of the raw data collected, and the methods for collecting ground truth and calibration data.

Chapter 3: Feature Engineering

The focus of this chapter is on engineering essential features from the collected data. It addresses the problems of missing vibratory data and presents solutions. Then, it delves into the derivation of power attributes from vibration data and their importance in our system. Additionally, the formulation of rolling pattern attributes, PassID, and Average Pass Count are discussed. The chapter concludes by covering the normalization of derived attributes, denoising, and outlier removal, along with the various feature extraction mechanisms employed in our system.

Chapter 4: Density Prediction

In this chapter, we explore the density prediction process. We investigate various clustering techniques employed to identify different vibration patterns. The section discusses the clustering algorithms used in our system and emphasizes the importance of ordering the clusters for our application. Furthermore, we explore possible processes to convert the ordered clusters into density.

Chapter 5: Evaluation Metrics

This chapter defines the evaluation metrics used to assess the effectiveness of the power clustering and cluster-to-density algorithms. Two types of evaluation metrics are discussed, based on finding the cluster and as per domain knowledge. The chapter also incorporates the use of RMSE and R-squared metrics to measure cluster-to-density prediction.

Chapter 6: Results and Discussion

Here, we present the results obtained from the clustering step, including visualization and identification of the optimum clustering mechanism.

Additionally, we discuss the selection of the cluster ordering criterion and present the results of density prediction. The chapter also formalizes the algorithm that gives the optimum result for clustering.

Chapter 7: Conclusion and Future Work

In the final chapter, we summarize the key findings of the research and discuss their implications. We highlight the contributions of this work to the field of Intelligent Compaction and outline potential areas for future research and development in this domain.

2. DATA COLLECTION

From the [47] [49], we get a theoretical basis that if all other factors are kept constant, the change in stiffness of the surface changes the vibratory response of the Compaction Roller. Also [50] [27] provide us with the instrumentation base for the design of a system capable of measuring the change in the vibration in rollers using an accelerometer. Subsequently, an electrical configuration was established in tandem with software development to acquire real-time vibration data at a sampling rate of 1000Hz. Following that the vibration data is first filtered according to its fundamental frequency, and subsequently, the powers in both the fundamental frequency and its six higher harmonics are measured to generate the data utilized in this study.

2.1 Instrumentation

Initially, our objective involves developing a retrofitting system for the roller, which should possess the capabilities of collecting data and eventually presenting results to operators. As a first step, we must carefully select the essential instruments that our system will employ, ensuring they can effectively manage the required constraints.

Vibratory Rollers:

The vibratory roller compactor used for this investigation was either the Caterpillar CB-7 or the Caterpillar CB-10 double drum vibratory roller, both of

which were instrumented for the purpose of the study. Both of these rollers are solid and smooth with the capability of independent as well as tandem vibratory modes. The two rollers have drum widths of 1.49 m and 1.70 m respectively. In both rollers, the frequency and amplitude of the vibration are the be set before operation and can range from 40-65HZ and 0.25-0.8 millimeters (mm). All other specifications of the rollers are provided in Table 8.1 in Appendix A. Within each drum, an eccentric mass configuration rotating about the drum axle provides the vibratory or eccentric force:

$$f_{ecc}(t) = m_0 e_0 \omega^2 cos(\omega t) \tag{1}$$

where,

ω is the circular frequency (rad/s),

m_0 is the eccentric mass,

e_0 is the eccentricity.

The angular frequency of vibration ω or more commonly frequency $f = \omega/2\pi$ and eccentric mass $m_0 e_0$ are to be set as per the construction design parameter and using the results in the test stretches and are constant until the construction design changes for the duration of the project. For our system, we consider this to be constant so as to reduce the factors affecting vibration patterns other than the compactness of the road.

Accelerometer:

Accelerometer provides a difference in the voltage reading as per the change in the acceleration of the system it is attached in. While a system is in vibration there are continuous and abrupt changes of acceleration. We used a 2240-010 Titanium Hermetic Accelerometer from Silicon Design Inc that requires +9to+32V DC to operate and provides +0.5to+4.5 single-ended output. Accelerometer has the nominal frequency response inside 3db in 0-1000Hz. The accelerometer is rugged and sealed in a titanium case to be able to resist the harsh condition outdoors on construction sites.

Temperature Gauge:

The compaction process should be carried out at the designated nominal temperature for breakdown or intermediate rolling as per the rolling pattern designed at constrained stretch. So, having a real-time temperature is a necessary tool for compaction roller operators. RICA system is designed to have a CI Series Compact IR Temperature Sensor from Raytek. As an accelerometer, it is also rugged to be able to resist the outdoor HMA pavement construction where the attached system is vibrating continuously. It has a 150ms response time with 0.1^0C resolution and provides a linear temperature range of $-20to500^0C$. It provides two-way RS232 digital output that can be attached to the Data Acquisition system to get continuous time series signals of the temperature.

GPS System:

For the rolling pattern, we used the GPS system Vector VS1000 or VS500 GNSS RTK Receiver from Hemisphere that provides the location information at max 20HZ. It can provide output at a baud rate of 4800-115200 through RS232, ethernet, or USB. We opted to go the USB route. Using an RTK system on top of the normal GPS, the Vector system could achieve a resolution of 8-15mm. Since the system has its own Data Acquisition System, the output of the GPS hub is directly fed into the tablet, which is the processing and communication center of RICA.

Data Acquisition System (DAS):

Labjack T4 from Labjack Measurement and Automation is used as the data acquisition center for Accelerometer and Temperature Gauge. It provides 8 configurable low voltage analog inputs (0-2.5V, 12-bit resolution) that can function as digital I/O lines among other various high voltage and full digital I/O modes. It supports the LJM driver/library for simplifying device communication with the processing units. It stores 128 ms of data while sampling at 1000 Hz and its memory has to be cleared at that rate to not lose data.

Central Processing Units:

xTablet T1690 along with its rugged holders from Mobile-Demand is used as the central processing unit for RICA. It comes with 16GB SSR4 ram, 1TB SSD storage, Intel Core i5 processor, and Windows OS preinstalled; but we moved it to Linux 20 as our Operating System. A Special Mobile Network communication

module is installed in the tablet so as to connect to the internet and upload data to the cloud system as necessary. It acts as the processing unit as well as the display unit to give real-time predictions to operators. Here, we deployed the RICA software system developed that uses InfluxDb as a database tool. InfluxDB is a time-series database able to store data at 1000Hz and has a high query process rate enough to provide real-time results.

2.2 Retrofit system

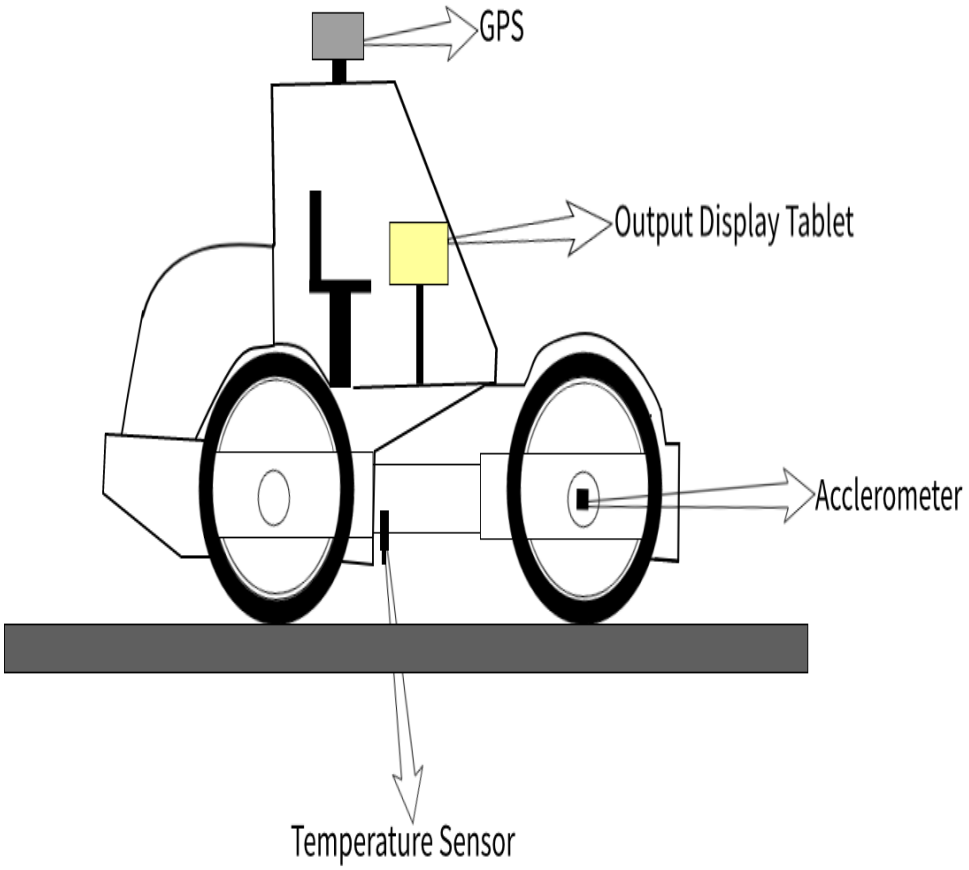


Figure 2.1: Figure depicting the positions of various instruments of RICA in Rollers

RICA has been purposefully engineered for seamless integration into the

existing market rollers, featuring an advanced data acquisition system comprising an accelerometer, temperature gauge, and Global Positioning System (GPS). These sensors have been seamlessly incorporated into the compaction roller to seamlessly gather data during real-time asphalt pavement construction. The accelerometer is strategically positioned at the axle of the roller drum to capture vibration patterns accurately. Additionally, a temperature gauge is affixed, directed towards the center of the road, while the GPS sensor system is meticulously situated at the roller's apex along the central axis, ensuring precise tracking of the roller's central coordinates. The precise arrangement of these sensors is of utmost significance, as subsequent calculations depend on these assumptions – such as the GPS sensor providing the central roller location, the accelerometer residing at the center with its measurement direction perpendicular to the road, and the temperature sensor gauging the temperature between the two rolling wheels.

2.3 Structure of raw data

Accelerometer and Temperature Gauge provided the time series data of vibration response in the drum of the compaction roller and temperature of the hot mix asphalt to be paved respectively at the rate of 1Kz. Readings received from the sensors are just the voltage reading so we need to convert the data to respective format. The temperature sensor reading is converted to temperature in Fahrenheit scale in a linear relation. The calibration of the temperature sensor is done once when the RICA is installed but can be changed if new calibration is to be done. For the readings from the accelerometer, it gives the voltage output as per the

amount of vibration it experiences so we don't need to change it to anything.

The GPS system supplies the latitude and longitude coordinates of each location, along with the speed and heading of the rollers' movement at specific points. Speed and headings, representing the angle of travel with respect to the East direction, are determined based on the current and previous GPS locations. Operating at a frequency of 1 Hertz, the sensor facilitates data collection, enabling us to capture each position at intervals of approximately 6 to 7 feet, contingent upon the roller's rolling speed.

Following the subsequent stage, the data obtained from Labjack and GPS is integrated in real-time to form a raw data table. For every 1000 data points of vibration, one corresponding GPS data point is included. Ultimately, the raw data system comprises the following components:

- Timestamp
- Vibration
- Temperature
- Latitude
- Longitude
- Speed
- Heading

Throughout the research process, all preceding steps were successfully carried out on on-road HMA (Hot Mix Asphalt) pavement sites. Visualizations of the raw data are provided in Appendix A.

2.4 Collection of ground truth

The collected raw data is unsupervised, meaning ground truth density values are not available for every location. Acquiring actual density ground truth involves cutting cores, which can initiate road degradation due to water leakage into the asphalt. Thus, obtaining ground-truth data is challenging, resulting in a small number of available ground-truth values. Consequently, we solely relied on unsupervised techniques to determine various vibratory patterns in the roller and only used a few ground-truth values for validation. Additionally, coring is performed when the roller completes its compaction process, leaving no intermediate readings for ground truth, highlighting the challenges of data collection in real construction sites compared to controlled environments.

2.5 Collection of calibration data

Furthermore, we gathered data on pavement density during the HMA laying process on the pavement by the paver. In cases where a construction involves two vibratory roller systems (Breakdown and Intermediate Rollers), we obtained density measurements before the specific roller starts vibrating. This baseline density represents the initial level of compaction density with which the roller interacts.

Additionally, during the design of any HMA mix for construction, the HMA mixing laboratory conducted tests to determine the maximum achievable level of compaction density in an isolated lab environment. This value represents the

theoretical maximum level of compaction that can be achieved.

These two values, the baseline density, and the theoretical maximum density, define the theoretical minimum and maximum levels of compaction achievable. In our RICA system, we utilize these two values to ensure that all our predictions fall within these specified limits, thus ensuring the accuracy and reliability of our results.

3. FEATURE ENGINEERING

3.1 Data Imputation

The raw vibration and temperature data are presented as time-series data, and occasionally, there were instances of missing data. To manage the data input rate, we optimized the Data Acquisition System (DAS) by regulating the data pull rate from DAQ to the CPU. Since the DAQ's cache can store a maximum of 150 temperature and accelerometer data points at once, we ensured timely data extraction to prevent reaching this threshold. Despite these efforts, there were rare cases of missing data. To address this, we applied third-degree spline interpolation, which can effectively resolve up to two continuous missing values in the vibration and temperature data. This method serves as the maximum extent of missing value resolution in our dataset.

For the GPS data, it was observed that in remote locations, chunks of data were lost when the connection to the satellite was lost. Unfortunately, there is no feasible solution to impute data in such conditions. In cases where only a small amount of data was lost, we adapted a linear interpolation method as an alternative approach.

Raw data provide us with the vibration in various locations of Rollers while paving the HMA. Next, we needed to formulate attributes that hold the different important information of vibration pattern and rolling pattern. Power attributes from vibration data and Rolling pattern attributes from the GPS location data are derived in the next step.

3.2 Derivation of power attributes from vibration data

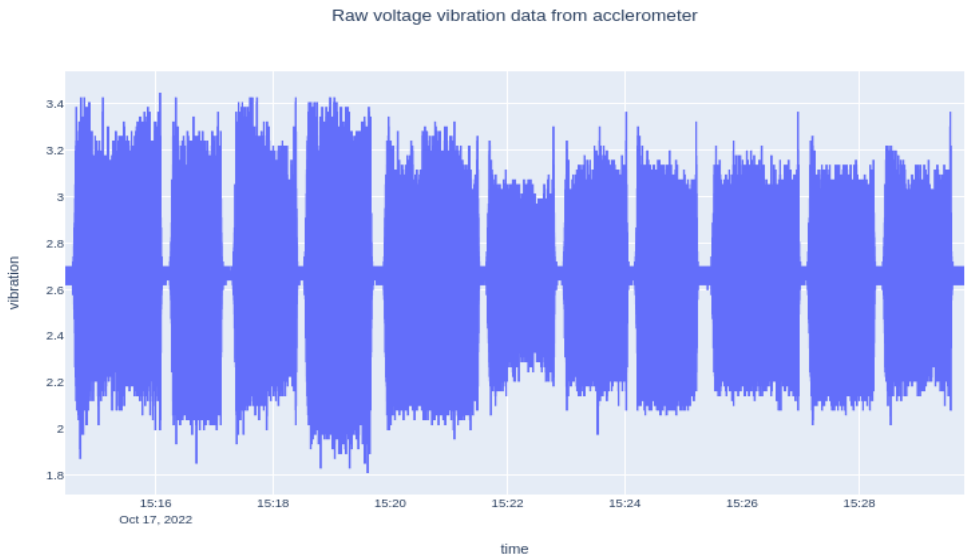


Figure 3.1: Stretch of raw vibration data

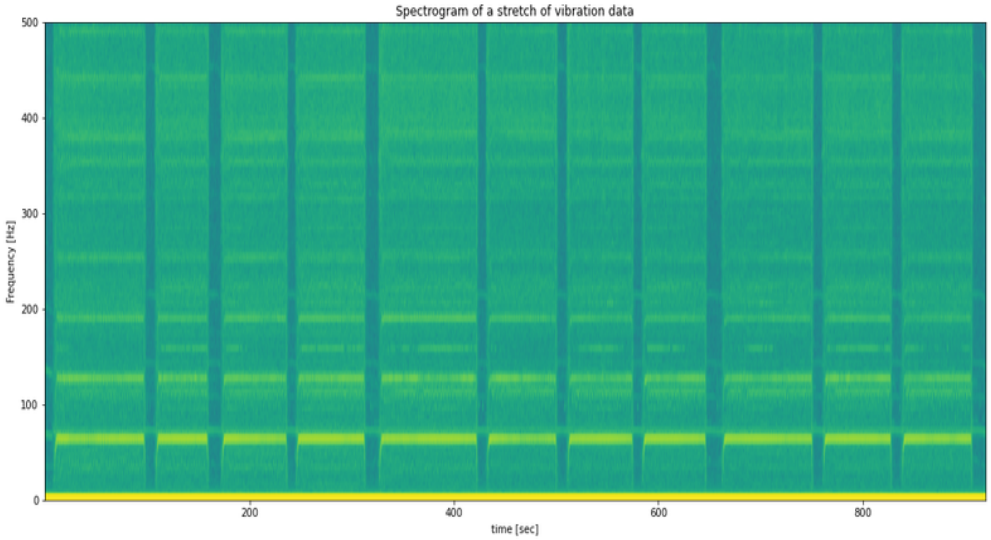


Figure 3.2: Corresponding spectrogram of vibration data from Figure 3.1

The frequency content of the continuous vibration time signal data $x(t)$ can be analyzed using the well-know Fourier Transform. A spectrogram is generated using the Discrete Fourier Transform (DFT) by first dividing a signal into short overlapping segments, then applying the DFT to each segment to convert the time-domain signal into the frequency-domain representation. The fast Fourier

transform (FFT) [51] [52] is a practical approach to the numerical computation of the DTFT for a finite length sequence and provided the power contained at each frequency in the spectrum of the signal. This process involves calculating the complex amplitudes of various frequency components present in each segment of the signal, resulting in a matrix of frequency and time bins. The magnitude or power of these complex amplitudes is often used to create the spectrogram, where each column represents a specific time segment, and each row corresponds to a frequency bin, illustrating how the signal's frequency content evolves over time.

During periods when there is no compaction vibration (i.e. when the drum with eccentric weights inside the roller remains inactive), the spectrogram shows no significant power at any frequency level. However, upon turning the compaction vibration on, the accelerometer registers a substantial power shift in the vibration data. The spectrogram reveals a concentration of power in specific frequencies, while most frequencies display little change compared to the non-vibration state. The first frequency of such concentration is the fundamental frequency(f0) and observed that the f0 of the roller-pavement system is the same frequency as that of the roller drum.

All possible frequency bands up to the Nyquist Frequency, which, in the case of data collected at 1 KHz, is 500 Hz, are considered, and changes in power level within each band are analyzed for any significant variations. Significantly, the frequencies that display noteworthy power changes are the fundamental frequency, at which the vibratory roller operates, and its harmonics. As the roller maintains vibration in the same location, a substantial power shift occurs around the fundamental

frequency and its harmonics, while the power at higher harmonics undergoes minimal change, with remarkable shifts being predominantly concentrated in the lower frequencies. Notably, at high frequencies, noise has a more considerable impact on power levels than changes in pavement stiffness. Therefore, the power content around the fundamental frequency and its first six harmonics is calculated and utilized as a vibratory feature of the system.

In pursuit of real-time results, a time-step of 128ms is implemented, and all manipulations are performed on data acquired every 128ms. Within each time-step data, features representing the power levels in the vicinity of the fundamental frequency (f0) and its multiples are extracted for input to the prediction model. To achieve this, fourth-order band-pass Butterworth filters are designed.

In the realm of lower frequencies, the alteration in power manifests within a relatively confined area. However, as the harmonics elevate, the dispersion of power extends to encompass larger regions. With heightened compaction levels, the collective power across the system augments; nonetheless, this doesn't uniformly apply to power across individual frequency tiers. Subsequent to this, the vibration data undergoes filtration, pinpointing frequency ranges of 3Hz encompassing the fundamental frequency, 5Hz surrounding the first harmonic, and 11Hz surrounding the subsequent five higher harmonics. In each of these filtered frequency ranges, the corresponding power is calculated and designated as "p0," "p1," ..."p6," respectively.

Finally, a single value representing the total power in each step is calculated. But, we wanted this attribute to have similar weights from each of p0-p6. The

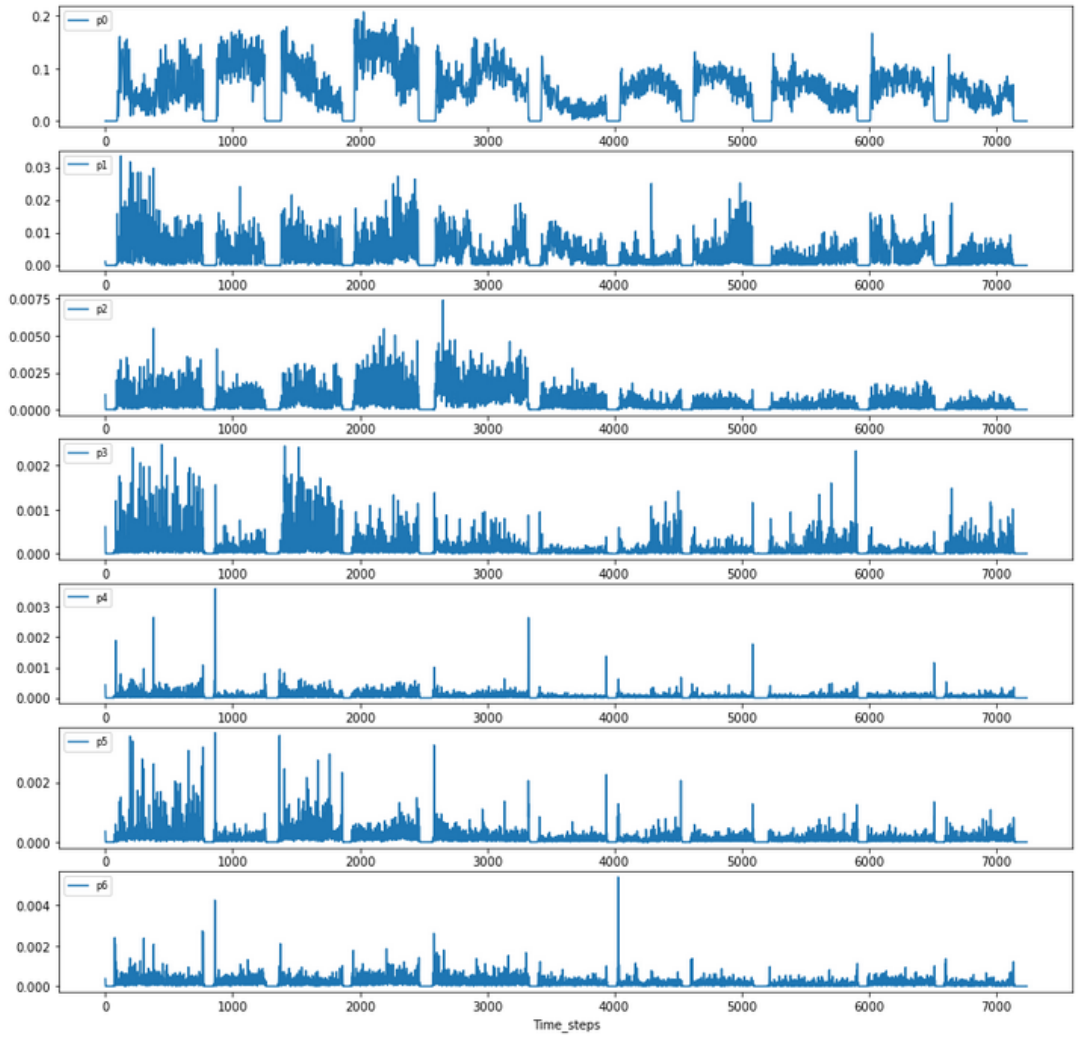


Figure 3.3: Various power attributes created in each time-step

calculated values of p0 when the roller is in vibration is in $1e^{-}2$ whereas values of p5/p6 are in order of $1e^{-}4/1e^{-}5$. So, simple summing does not give the desired result. Thus, we formulated a single parameter that weights the power value with the frequency it belongs to power to the harmonic number of the frequency called total_power:

$$total\_power(P) = \sum_{x=0}^{6} p_x f_x^{(x+1)} \tag{2}$$

where,

x=harmonic number(fundamental frequency has harmonic value of 0)

p_x=power in xth harmonic

f_x=frequency of xth harmonic

This power value gave us a single parameter to compare and contrast various intermediate results throughout the study.

3.3 Formulation of rolling pattern attributes

The rolling pattern encompasses the recommended technique for operating a vibratory drum roller, derived from multiple tests performed on a controlled stretch. Prior to commencing any Hot Mix Asphalt (HMA) construction project, a designated section of the site is prepared as a controlled stretch. Various tests are conducted to ascertain the optimal rolling pattern based on in-situ density measurements. The quality control/quality assurance (QC/QA) process involves conducting frequent checks at various points after each pass of the compaction roller. This enables the operator to determine the number of times a specific location should be rolled over to achieve adequate compaction.

The rolling pattern holds significant importance for roller operators, their supervisors, and QC/QA personnel, often serving as the initial step in HMA pavement construction. In certain instances, particularly for HMA pavement with a thickness of less than 1.5 inches, the construction site may prioritize assessing the roller's rolling pattern over taking cores and conducting additional spot checks. This approach is favored because vibration characteristics are substantially influenced by the underlying surface rather than the HMA layer and its compaction level. As a result, it's worth noting that the RICA system's effectiveness is constrained by

a requirement for the HMA pavement to have a thickness exceeding 2 inches for proper performance evaluation.

The compaction process involves the roller making multiple passes over a given area while applying vibration to achieve the desired density. These passes occur at speeds of approximately 3-5 mph. To capture the evolving state of each pass during operation, an attribute known as "Pass-ID" is formulated. This continuous, incrementing attribute represents each sequential pass made by the roller. It facilitates the tracking of changes in pass direction, with an increase in the Pass-ID recorded whenever the roller changes direction, stops, and then resumes movement in the opposite direction.

In a given stretch of pavement, it is commonly assumed that an increase in passId results in a corresponding increase in compaction density. However, as demonstrated in a controlled experiment in [34], this is not always the case, and there can be various reasons for such variations. One prominent factor is the different ratios of compacted and non-compacted regions with each increasing pass in the same stretch. The roller's width typically ranges from 6-7 feet, while a lane of road paved at a time is usually 10-12 feet wide. As a result, during each pass, the roller may not consistently follow the same path over the entire width.

Consequently, one pass might exhibit significant overlap with a previously compacted pavement region, resulting in a high compaction level. On the other hand, another pass may predominantly traverse freshly laid-over asphalt mix or uncompacted pavement on a different side of the constructed lane. It is also observed that the start of one pass may significantly overlap with the compacted

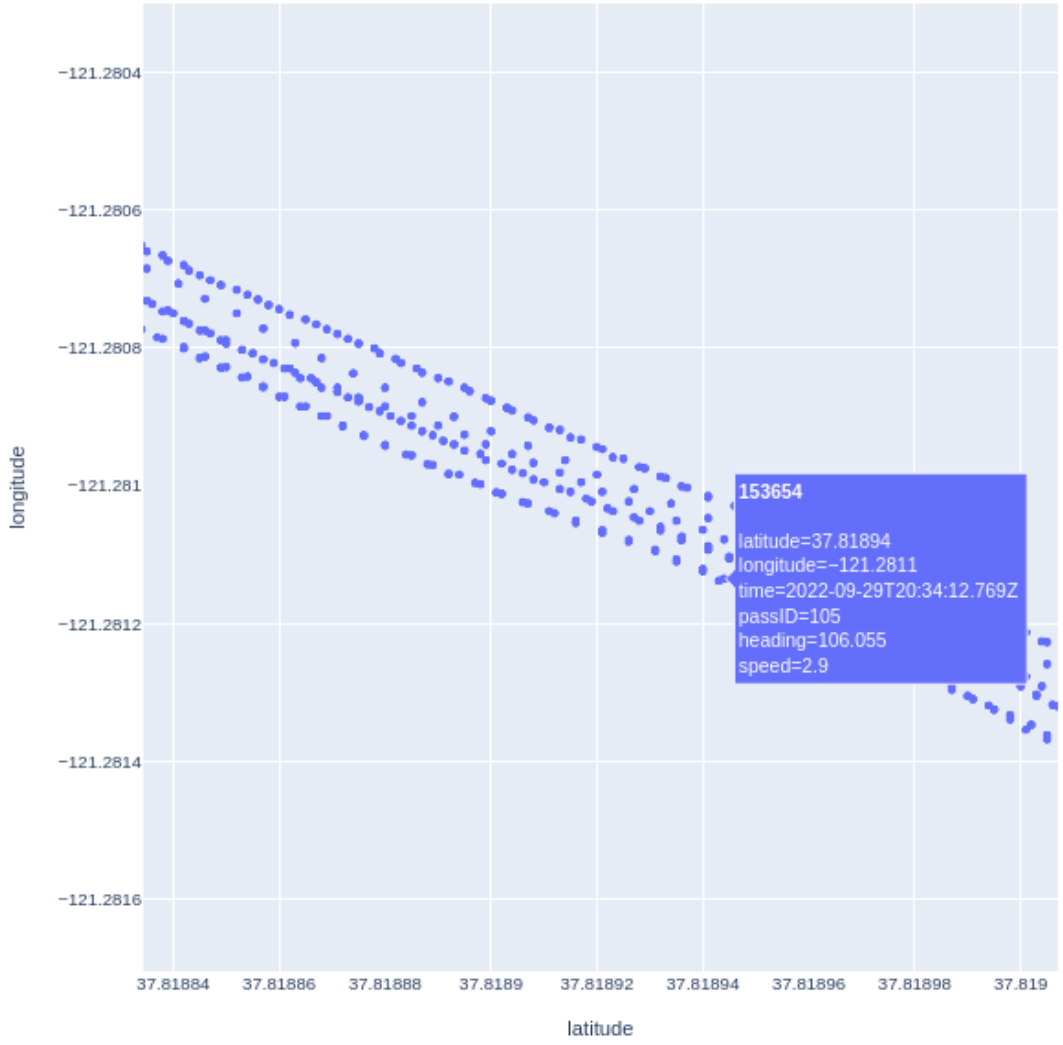


Figure 3.4: Visualization of location data of compaction roller during HMV pavement construction

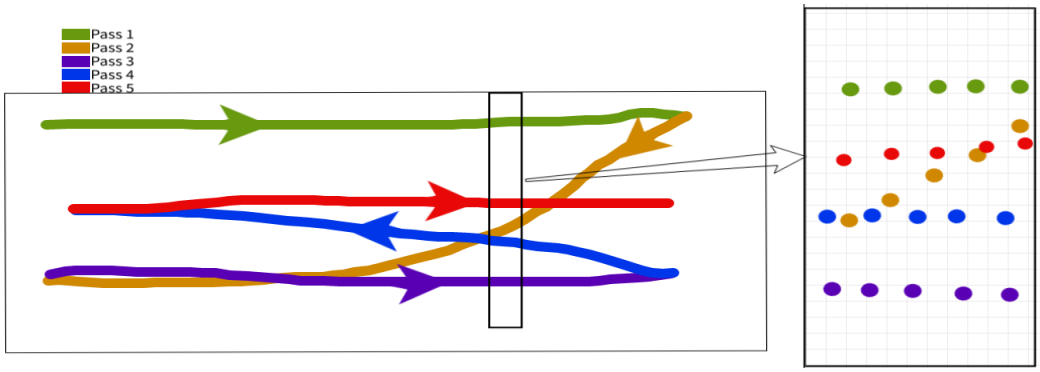
pavement region whereas the end of the pass reaches the virgin, uncompacted region leading to a significant difference in compaction density in the same pass. To address this issue, a new attribute called Average Pass Count (APC) is introduced. It calculates and signifies what portion of the pavement area the roller has previously passed through in the last time-step. APC represents the average number of passes the roller has made in the area it most recently traversed, offering valuable insights into the compaction process.

To calculate the Average Pass Count (APC), a stretch of pavement is taken into account, and a grid of $1/2ft\ X\ 1/2ft$ is constructed in the latitude-longitude space. The size of the grid in latitude-longitude space varies based on the project site's location and the relation between distance and latitude-longitude. Each grid cell possesses a pass-count property, representing the number of times the roller has passed over it.

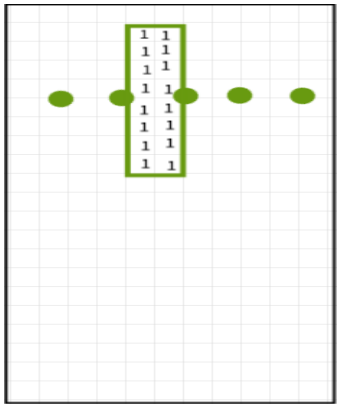
For each time step, a rectangular area is formed with a length equal to the distance traveled during the last time step and a width equivalent to the roller's width, which is in a direction perpendicular to the roller's heading. Subsequently, the pass count of all grid cells within this rectangular area is incremented by one. Finally, the APC for the given time step is calculated as the average pass count within the previously defined area. This process enables a quantification of the roller's previous passes at a given location and so each time step will have characteristics of a rolling pattern that is dependent only in its vicinity.

The complete process can be observed in the example stretch shown in Figure 3.5. In this scenario, the roller makes five passes in a specific location data stretch. For a given time_step between two GPS data points, we create a rectangular area. In this example, the rectangle is oriented perpendicular to the left/right direction for simplicity, rather than being perpendicular to the heading. During each pass, we update the pass_count for each grid area that the rectangle touches. The Average Pass Count (APC) for the specific time step is determined by computing the average pass count within the grid areas intersected by the rectangle.

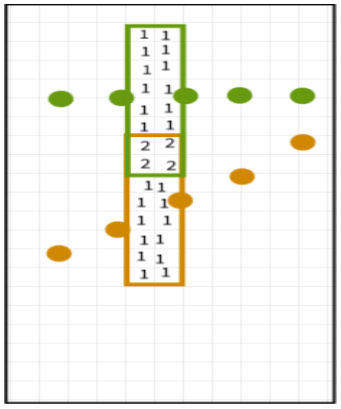
For instance, in the example stretch, the APC for the time_step between the



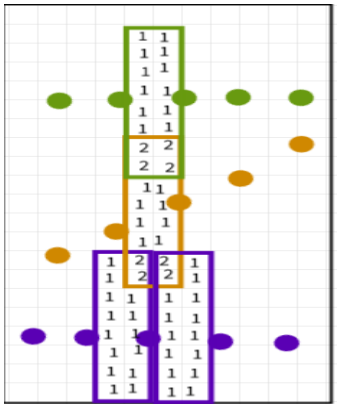
(a) An example stretch of passes of the roller when paving the road and zoom in at a location of the stretch.



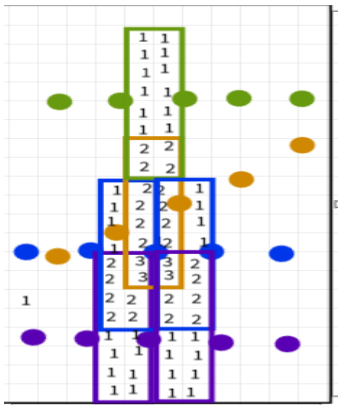
(b) Update of pass_count parameter of each grid in the first pass.



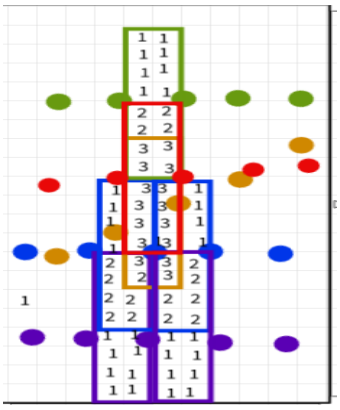
(c) Update of pass_count parameter of each grid in second pass pass.



(d) Update of pass_count parameter of each grid in third pass pass.



(e) Update of pass_count parameter of each grid in fourth pass pass.



(f) Updation of pass_count parameter of each grid in fifth pass pass.

Figure 3.5: Visualization of calculation of Average Pass Count(APC) attribute

second and third GPS locations from the left in pass 1 data is 1. The APC for the time_steps in the nearest location (between second and third gps readings from left in zoomed-in section) in pass 2 data is 1.25. We get it by averaging 2,2,2,2,1,1,1,1,1,1,1,1,1,1,1. Similarly, APC in the nearest location in pass 3 data is 1.125, in pass 4 data is 1.875, and in pass 5 data is 2.75.

To calculate the APC, access to a continuous stretch of data is necessary, making it impractical for real-time systems. Therefore, the APC is to be exclusively employed for the training and evaluation stages but may not be applicable in real time after the completion of training.

3.4 Normalization

From Figure 3.3, it is observed that the value of power in the seven frequency bands varies a lot. We can observe that power in fundamental frequency is in the range of 1e-1 whereas the power in the top two harmonics(p5,p6) are below 1e-5. We are working on the presumption that each of these seven features has equal weight but since the numerical values are different in the order of 1e3, we need to normalize them. We choose to normalize the data using mean and standard deviation.

$$\bar{P_x} = \frac{1}{N}\sum_{i=1}^{N} P_{x,ti} \tag{3}$$

$$\hat{\sigma}_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N_1}(P_{x,ti} - \bar{P_x})^2} \tag{4}$$

$$P_{x,ti}|norm = \frac{P_{x,ti} - \bar{P_x}}{\hat{\sigma}_x} \tag{5}$$

where:

N = Number of data points in training data

$P_{x,ti}$ = Power in x^{th} harmonics at time ti

$P_{x,ti}|norm$ = Normalized power x^{th} harmonics at time ti

$\bar{P_x}$ = Mean power of x^{th} harmonics

$\hat{\sigma}_x$ = Standard deviation of power at x^{th} harmonics

Using the above equation we can get the normalized values of all the powers in the form of deviation of power from mean in terms of multiple of standard deviation. Values of $\bar{P_x}$ and $\hat{\sigma}_x$ are saved in the calibration file from the training step. These values are used to normalize data from each time step in the real-time operation of the system.

3.5 Removal of outliers

From the raw power data in Figure 3.3, we observe significant disturbances in power at various time steps. Notably, the start and stop times of the vibration in the roller consistently yield very high power readings for the accelerometer, but this data is considered an outlier for our system. Furthermore, we notice similar non-normal power responses whenever the roller encounters roadside curves. In such cases, we expect all the data that falls outside the range of $+/-2\hat{\sigma}_x$ to be outliers of the system and clip everything to fall in this range.

These spikes in power levels occur when the roller hits curves, other hardened substances, or large aggregates beyond the prescribed range. These occurrences are relatively common on construction sites, but such power levels do not indicate

the compaction process. Therefore, while making predictions, we need to clip these extreme values and exclude such data during the training of the prediction model.

After normalization and outlier removals of data, we get the feature that we are going to use in the learning step. The normalized powers of the unnormalized powers in Figure 3.3 can be observed in Figure3.6. Here, we can see that each feature ranges from -2 to +2, and anything larger than +2 is truncated to the value of +2, and smaller than -2 is pushed up to -2.

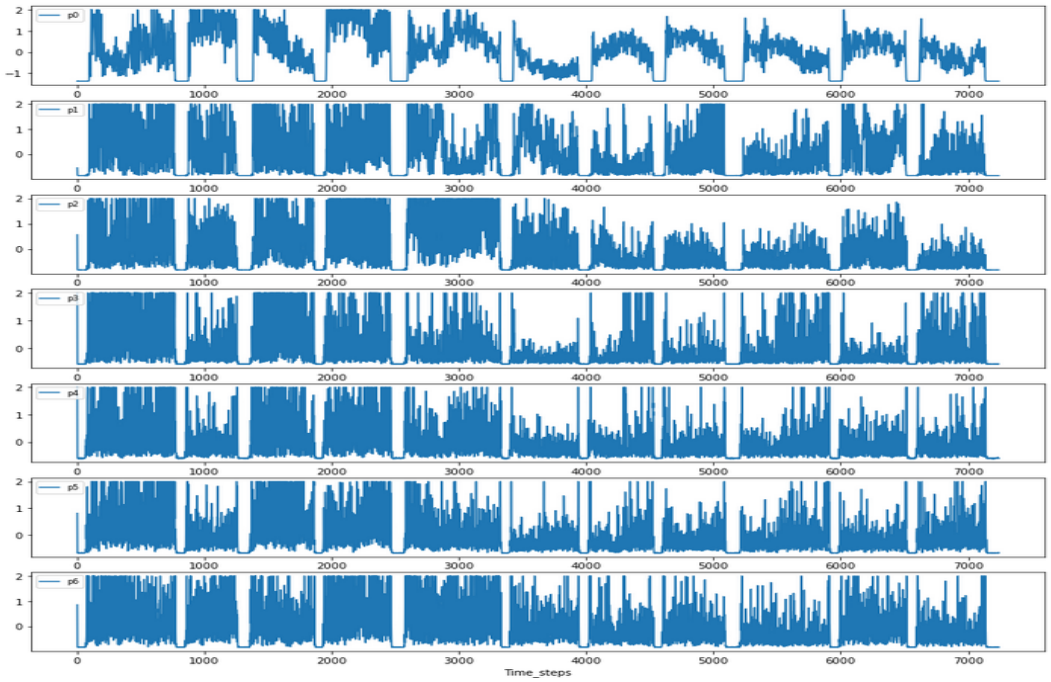


Figure 3.6: Various power features created in each timestep

From Figure 3.1 as well as Figure 3.2, we see that there are a lot of times with no vibration and those have no vibratory relation to the amount of compaction and also the system we are designing cannot use those data to process. These are the noises that can affect the learning mechanism in the system. So, remove them in the training using the total power calculated. we always see that the power calculated for the no-vibration data is always less than 1, and removing any data

below this threshold resolved the issue.

3.6 Feature Extraction

Once normalized features have been generated to quantify variations in vibration and rolling characteristics, the next step involves assessing the significance of these variations in determining distinct compaction densities. The objective of feature extraction within unsupervised learning is to convert the initial raw data into a novel representation that captures crucial information and inherent patterns within the data. This procedure trims down the data's dimensions, filters out noise, and emphasizes meaningful structures, thereby rendering it more suitable for subsequent tasks such as clustering, visualization, or other unsupervised learning endeavors. Feature extraction contributes to refining algorithm performance by concentrating on pertinent attributes and curbing computational complexity. Consequently, we transformed power features across various linear and non-linear frameworks, and subsequently cherry-picked essential features using evaluation metrics to make density predictions.

In this study, we tested various feature extraction algorithms. In this application, we don't have any one-to-one data to density ground truth data so we need to apply only unsupervised methods of feature extraction. We experimented with the original seven power data along with feature extracted using Principal Component Analysis(PCA) [53], Kernel PCA (K-PCA) [54] and Isomap [55] to train various clustering models and analyzed them.

Principle Component Analysis(PCA): Principal Component Analysis, often referred to as PCA, is a valuable method for reducing the dimensionality of data with high dimensions. It achieves this by transforming a large set of variables into a smaller one while retaining most of the essential information present in the original data set. PCA encompasses both the steps of feature extraction to a new space and the reduction of features within that space. In our study, we performed data transformation to the PCA space and evaluated the various feature reduction in the metrics defined in Section 5.1 for clustered data.

Channel PCA (K-PCA): Channel-PCA or Kernel-PCA is nonlinear feature extraction and then reduction algorithm. K-PCA effectively computes principle components in higher dimension feature spaces related to the input space using Kernel methods [56]. Evaluation of feature reduction was done on the basis of the metrics in Section 5.1.

Isomap: Isomap [57] is a dimensionality reduction technique that constructs a lower-dimensional representation of data by preserving geodesic distances (shortest path distances) between data points on a manifold. Its ability to capture nonlinear relationships and maintain the local and global structure of data makes it a valuable tool for exploring complex datasets and gaining insights into the underlying patterns and clusters within the data. We evaluated the use of Isomap in our application using the clustering evaluation metrics defined in Section 5.1.

When working with data transformed into various spaces, we aimed to avoid

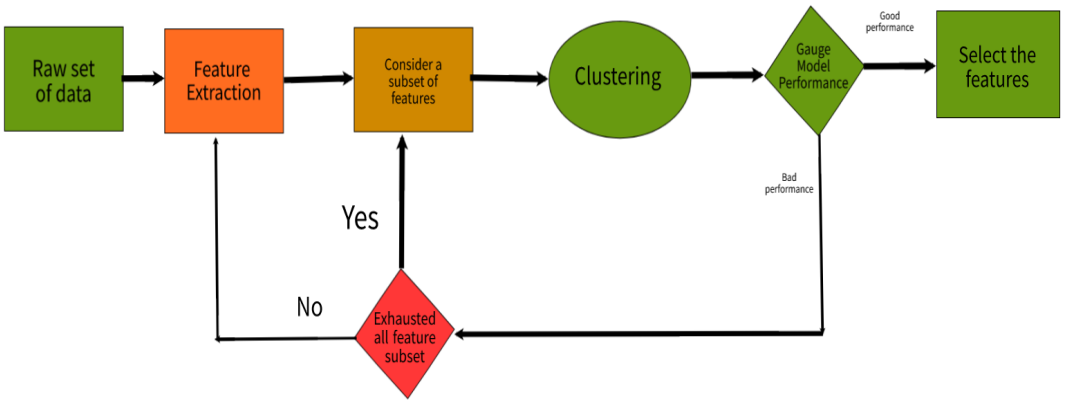


Figure 3.7: The flowchart depicting the method of Feature Extraction and Selection utilized.

using all the high-dimensional data. To address the curse of dimensionality, we experimented with reducing the number of features and examined the outcomes. We adopted a Wrapper-based feature selection approach, utilizing the evaluation metrics defined in Section 5.1, to assess the effectiveness of feature extraction and selection for our application. Wrapper-based feature selection involves applying the selected features to the model and analyzing the evaluation metric on the verification data.

As illustrated in Figure 3.7, we implemented a two-level iteration to explore different feature extraction and selection processes. While this method is time-consuming, it offers a comprehensive assessment of feature selection and provides exhaustive results.

3.7 Data Visualization

Feature extraction is a mechanism that is not only used to decrease the dimensionality of data but there are some feature extraction algorithms used for

visualization but rarely for dimensionality reduction. We have used two feature extraction algorithms in the visualization of raw data viz: t-SNE and UMAP.

t-Distributed Stochastic Neighbor Embedding (t-SNE): t-SNE (t-distributed Stochastic Neighbor Embedding) [58] is a dimensionality reduction technique that minimizes the divergence between probability distributions to map high-dimensional data points into a lower-dimensional space while preserving the pairwise similarities between the data points. t-SNE primarily serves as a powerful visualization technique for high-dimensional data, assigning each data point a position in a two or three-dimensional map. While it can technically be used for dimensionality reduction, as explored in [58], we discovered that it is best not employed as a feature extractor. However, t-SNE proved invaluable for visualizations, offering valuable insights and a better understanding of the data at hand.

Uniform Manifold Approximation and Projection (UMAP): UMAP (Uniform Manifold Approximation and Projection) [59] is a dimensionality reduction technique that constructs a low-dimensional representation of high-dimensional data by preserving local and global data structures while minimizing the spread of points in the embedding space. UMAP is robust to variations in hyperparameters, providing consistent and stable results across different runs, and enhancing its reliability for data analysis and exploration. In this research, we have used it only for visualization purposes to observe if there are clear clusters or not.

4. DENSITY PREDICTION

After all the feature engineering process in the data we collected, we delve into finding the pattern of power and thus the vibration for different level of compaction to predict density levels.

4.1 Learning mechanism

In our continuous system, regression would be the ideal technique to solve the application problem if we had sufficient one-to-one ground truth data. However, due to the absence of supervised data, we turned to unsupervised learning methods. Clustering, an unsupervised technique, became our choice to group similar data points together based on their inherent characteristics or similarities. Its main objective is to unveil distinctive patterns and structures within the data, forming separate clusters with shared properties yet distinct from one another. By leveraging clustering, we can gain valuable insights into the underlying relationships and trends in the data, allowing us to make informed decisions and predictions despite the lack of labeled data.

In the Introduction section, we established that the vibration of the Asphalt-Roller system directly depends on the compaction density of the asphalt. Hence, our primary goal was to identify various power patterns and clusters representing different compaction densities. To achieve this, clustering emerged as a prominent solution. By employing clustering techniques, we aimed to group similar vibration patterns together and uncover distinct clusters that correspond to different levels

of compaction density in the asphalt. This approach allowed us to analyze and understand the variations in vibration data and ultimately gain valuable insights into the compaction process.

Clustering Algorithms

In our study, we extensively explored a variety of clustering algorithms to identify the most suitable approach for our application. Our investigation encompassed diverse techniques, including density-based, distribution-based, centroid-based, and hierarchical-based clustering methods. For each of these methods, we carefully assessed and fine-tuned the hyperparameters to obtain optimal results, which were further visualized for comprehensive analysis. Some of the major algorithms we tried in the study are described below.

◇ K-Means Clustering

K-Means clustering is a popular unsupervised machine-learning technique used for partitioning data into distinct clusters based on their similarities. The fundamental concept behind K-Means clustering is to divide a dataset into 'K' clusters, where 'K' represents the number of desired clusters specified by the user or determined through optimization techniques. The algorithm iteratively assigns data points to the nearest cluster center and recalculates the centroids of these clusters until convergence is achieved, ensuring that the data points within each cluster are as close to the centroid as possible.

The process begins with the random selection of 'K' initial centroids, which

act as the center points for each cluster. Then, each data point is assigned to the nearest centroid based on a distance metric, like Euclidean distance, cosine similarity, Manhattan distance, or others as per the application. This assignment creates 'K' distinct clusters, and the mean of all data points within each cluster becomes the new centroid. The algorithm repeats this assignment and centroid recalculation process until the centroids stabilize, and no further reassignment is necessary.

K-Means clustering aims to minimize the within-cluster sum of squared distances, often referred to as the sum of squared errors (SSE) or distortion. The objective is to find the optimal partitioning of data points into clusters that minimize the overall SSE. The objective function for partitioning data $x_1, x_2, ...x_n$ into k sets $S = S_1, S_2, ...S_k$ is to minimize the SSE loss (J_{SSE}). For each iteration, if a data x_i is assigned in cluster set S_i,

$$J_{SSE} = \sum_{i=1}^{k} \sum_{x_i \in S_i} ||x_i - \mu_i||^2 \tag{6}$$

where, x_i is all data belonging to cluster S_i

μ_i is the mean of S_i

However, a significant drawback of K-Means is that it is sensitive to the initial centroid selection, leading to potential convergence to local optima. To mitigate this issue, the algorithm is often run multiple times with different random initialization, and the result with the lowest SSE is chosen as the final clustering.

It is essential to determine the appropriate value of 'K' for the K-Means

algorithm, which significantly impacts the quality of the clustering result. Several techniques, such as the elbow method and silhouette analysis, are used to identify the optimal 'K' value by evaluating the trade-off between the number of clusters and the SSE.

K-Means clustering is known for its simplicity, efficiency, and scalability. However, it assumes clusters with spherical shapes and is sensitive to outliers, noise, and non-linearly separable data. The two major hyperparameters to tune in K-Means are the number of clusters(K) and the distance parameter to use.

In summary, K-Means clustering is a powerful and widely used algorithm in unsupervised machine learning, providing valuable insights into the underlying structure and patterns within data sets. Its ease of implementation and efficiency make it an attractive choice for various clustering tasks, enabling us to gain valuable insights and make data-driven decisions.

- **Gaussian Mixture Model(GMM)**

The Gaussian Mixture Model (GMM) is a probabilistic model used for clustering data points into multiple clusters. Unlike the K-Means algorithm, which assigns data points to a single cluster with a hard assignment, GMM allows for soft assignment, meaning that each data point can have a probability of belonging to multiple clusters.

In GMM, each cluster is represented by a Gaussian distribution, also known as a component. The model assumes that each data point is generated by one of

these Gaussian distributions, and the goal is to estimate the parameters of these distributions to best fit the data.

The GMM [60] algorithm works in an iterative manner to find the optimal parameters. The general steps of the GMM algorithm are as follows:

1. Initialization: Randomly initialize the parameters of the Gaussian distributions, including the mean, covariance, and mixing coefficients.

2. Expectation Step (E-step): In this step, the algorithm estimates the probability that each data point belongs to each of the Gaussian distributions. This is done using Bayes' theorem to calculate the posterior probability.

3. Maximization Step (M-step): In this step, the algorithm updates the parameters of the Gaussian distributions based on the probabilities obtained in the E-step. Specifically, it calculates the new means, co-variances, and mixing coefficients that maximize the likelihood of the data given in the model.

4. Convergence Check: After the E-step and M-step are performed, the algorithm checks for convergence. If the parameters have not significantly changed from the previous iteration, the algorithm terminates.

5. Iteration: If the convergence criteria are not met, the algorithm goes back to the E-step and continues the iteration until convergence is achieved.

GMM provides a more flexible clustering approach than K-Means, as it can

handle clusters of different shapes and sizes that are normally distributed. It is particularly useful when the underlying data distribution is not well-separated and contains overlapping clusters. Additionally, GMM can handle missing or noisy data points and can be used for data imputation tasks.

However, GMM has some limitations, such as the sensitivity to the initial parameter values and the potential for overfitting when the number of components is large. Therefore, it is important to carefully tune the hyperparameters and choose an appropriate number of components for the GMM model. Hyperparameters are parameters that are set before the model is trained and can influence the performance and behavior of the algorithm. Here are the main hyperparameters in GMM:

- Number of Components ($n - components$): The number of Gaussian distributions (components) to be used in the model. This hyperparameter determines the number of clusters that GMM will try to identify in the data. If this value is set too low, some clusters may be merged, leading to underfitting. On the other hand, if it is set too high, some clusters may overfit, and the model may become too complex and less interpretable.

- Covariance Type: This hyperparameter determines the type of covariance matrices used for each component. The three common options are: 'full': Each component has its own general covariance matrix. It allows the model to capture correlations between features but can be computationally expensive, especially for high-dimensional data. 'tied': All components share the same covariance matrix. It reduces the number of parameters to estimate and

is useful when the data has similar covariance structures. 'diag': Each component has its own diagonal covariance matrix. It assumes that the features are uncorrelated, making it more computationally efficient and useful for high-dimensional data.

- Initialization Method: GMM is sensitive to the initial parameter values. Different initialization methods can affect the convergence and performance of the algorithm. Common initialization methods include random initialization and k-means initialization.

- Convergence Criteria: This hyperparameter sets the threshold for determining convergence in the E-step and M-step of the algorithm. It defines how close the parameters need to be in consecutive iterations to consider the model as converged.

- Maximum Number of Iterations: This hyperparameter sets the maximum number of iterations that the algorithm will run, regardless of whether convergence is achieved. Setting this value too low may result in a premature termination before the model has converged.

- Regularization: GMM may suffer from overfitting, especially when the number of components is large. Regularization techniques, such as adding a small constant to the diagonal of the covariance matrix, can help control overfitting and improve model generalization.

The choice of hyperparameters in GMM depends on the characteristics of the data and the specific task at hand. Tuning these hyperparameters requires

experimentation and validation on a held-out dataset to find the best configuration that optimizes the model's performance and generalization ability. Grid search, cross-validation, and model selection techniques are commonly used to find the optimal hyperparameter values.

In summary, the Gaussian Mixture Model is a powerful and versatile clustering algorithm that can effectively model complex data distributions and provide valuable insights into the underlying structure of the data. Its probabilistic nature makes it well-suited for various applications, including pattern recognition, anomaly detection, and data generation.

◇ **Adaptive Gaussian Mixture Model**

Adaptive Gaussian Mixture Model (GMM) is an extension of the traditional GMM that incorporates adaptability to handle data with varying complexities and underlying structures. Unlike the standard GMM, which assumes a fixed number of components and static hyperparameters throughout the clustering process, the adaptive GMM dynamically adjusts its model complexity based on the characteristics of the data. This adaptability is achieved through a data-driven approach that automatically determines the optimal number of components and covariance structures to better capture the underlying patterns in the data. The adaptive GMM iteratively refines its model by considering both local and global data characteristics, allowing it to adapt to varying densities, non-uniform cluster shapes, and noisy regions in the data. By dynamically adjusting the model complexity, the adaptive GMM offers improved flexibility and robustness, making

it well-suited for clustering tasks in diverse and challenging datasets.

The core idea behind the adaptive GMM lies in its ability to automatically determine the number of components and covariance structures that best fit the data distribution. To achieve this, the adaptive GMM employs data-driven approaches, such as the Expectation-Maximization (EM) algorithm, to iteratively refine the model based on the input data. In each iteration, the adaptive GMM updates its parameters, including the means, covariances, and mixing coefficients, to better capture the underlying patterns in the data.

One approach to adaptivity in the GMM is the use of model selection criteria, such as the Bayesian Information Criterion (BIC) or the Akaike Information Criterion (AIC), to estimate the optimal number of clusters. These criteria penalize complex models and favor parsimonious solutions that better generalize the data. As the adaptive GMM iterates through the model selection process, it incrementally prunes or adds components to the mixture, effectively adjusting the model complexity to fit the data.

Yang and Zwolinski[39] argue that investigating the mutual relationships between components offers a different and potentially more effective approach to achieving an optimal mixture model. The proposed method provides a means of selecting an appropriate set of components without relying on the information contained in the observed patterns directly. The paper presents an algorithm based on mutual information theory and outlines its potential benefits in pattern recognition and clustering tasks.

Another aspect of adaptivity in the GMM is the use of varying covariance structures for each component. In traditional GMM, all components share the same covariance matrix, which might not be suitable for datasets with non-spherical or anisotropic clusters. Adaptive GMM addresses this limitation by allowing each component to have its own covariance structure, leading to a more accurate and flexible representation of cluster shapes.

The adaptive GMM has proven to be a powerful tool in handling complex and high-dimensional datasets, offering advantages over fixed-parameter models like K-means or traditional GMM. By adapting its model complexity to the data, the adaptive GMM can identify meaningful clusters even in the presence of noisy or overlapping regions. This makes it well-suited for a wide range of applications, including image segmentation, anomaly detection, and pattern recognition, where data characteristics can vary significantly across different regions of the dataset. Overall, the adaptive GMM provides a robust and versatile clustering solution that can adapt to the intricacies of real-world data, making it a valuable technique in data analysis and machine-learning tasks.

- **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**

DBSCAN [61] is a popular density-based clustering algorithm used for grouping data points in a dataset based on their density. Unlike k-means, which require the number of clusters as a hyperparameter, DBSCAN does not need a predefined number of clusters and is capable of identifying clusters of various shapes and sizes.

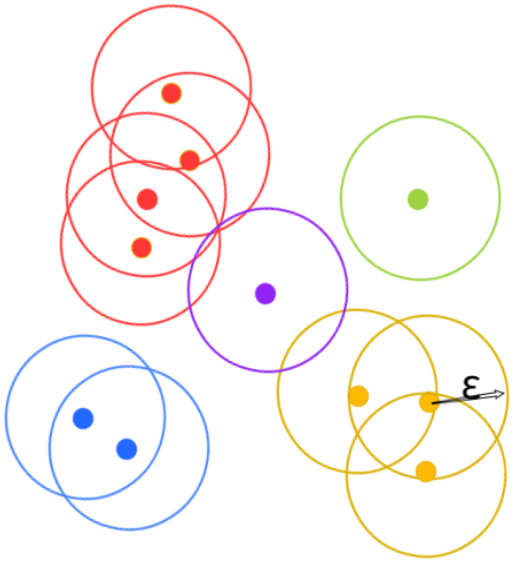


Figure 4.1: Example of DBSCAN clustering. For given ϵ and minPts=3, we get two clusters(red and yellow) and four noise points(2 blue, 1 purple and 1 green) using DBSCAN. There are two core points and two border points in the red cluster whereas yellow has only one core and two border points.

The DBSCAN algorithm works by defining two critical hyperparameters: "epsilon" (ϵ), representing the maximum distance between two points for them to be considered neighbors, and "minPts," which sets the minimum number of points required within (ϵ) distance to form a core point. The algorithm proceeds by examining each data point in the dataset, determining if it qualifies as a core point based on the number of neighboring points within ϵ distance. If a point has at least "minPts" neighbors, it forms a cluster. Additionally, points that are reachable from a core point but do not have enough neighbors to be core points themselves are classified as border points and are also included in the cluster. Points that have fewer neighbors than "minPts" and are not reachable from any core point are considered noise and are not assigned to any cluster. The value of "minPts" should be at least greater than the dimension of the data but generally, it is taken to be twice the dimension or as per the domain knowledge.

DBSCAN's ability to discover clusters of arbitrary shapes, robustness to noise, and not require a predefined number of clusters make it suitable for various applications, especially in cases where the number of clusters is not known beforehand or when dealing with datasets with irregularly shaped clusters and varying densities. However, its performance might degrade with high-dimensional datasets due to the "curse of dimensionality." As with any clustering algorithm, proper tuning of hyperparameters is essential to achieving optimal results in different scenarios.

Selection of Number of Clusters

Various distribution-based, centroid-based, and hierarchical clustering algorithms require a number of clusters as one of the major hyperparameters. Through domain knowledge or data information, if we can have this information beforehand; that is best. But, our data being from the sampled continuous system, we don't have that knowledge and so we need to figure out the number of clusters through analysis.

The idea behind the finding optimal number of clusters is that a good clustering solution should have clusters that are tightly packed together (small within-cluster variance) while being well-separated from each other (large between-cluster variance). Below are some of the techniques employed in our study to comprehend the data and identify the optimal number of clusters.

- **Elbow Method**: This technique involves plotting the within-cluster sum of squares (inertia) against the number of clusters. The plot typically shows a

decreasing curve. The optimal number of clusters is often identified at the "elbow" point, where the inertia starts to level off, indicating diminishing returns in variance reduction with additional clusters. The inertia here is calculated as:

$$Inertia = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij}(x_i - c_j)^2 \tag{7}$$

Where:

n is the total number of data points in the dataset.

k is the number of clusters being evaluated.

x_i is the ith data point in the dataset.

c_j is the centroid (mean) of the jth cluster.

w_{ij} is a binary indicator function that takes the value 1 if data point x_i is assigned to cluster j and 0 otherwise.

While plotting the inertia VS K (number of clusters) we get a graph like shown in Figure 4.2. By the Elbow method, the point where the inertia starts to level off corresponds to the optimum number of clusters. In the demo example at 4.2, k=4 is the optimum value.

◇ **Silhouette Analysis:** Silhouette analysis evaluates the compactness and separation of clusters. It calculates the average silhouette score for different cluster numbers. The silhouette score ranges from -1 to +1, with higher values indicating better-defined clusters.

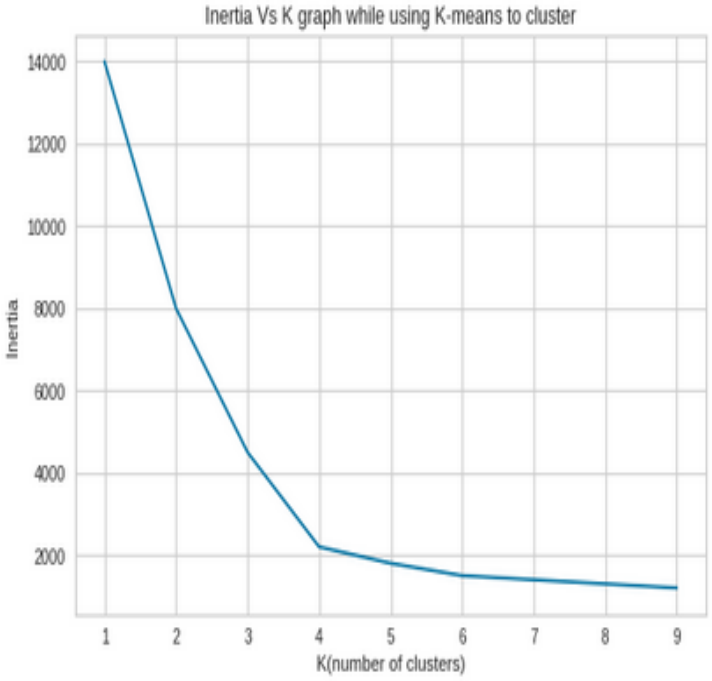


Figure 4.2: Inertia Vs K(number of cluster) plot. Here, using the elbow method, K=4 is the optimum number of clusters

Silhouette score for each data point i is calculated as:

$$s(i) = \frac{b(i) - a(i)}{max(a(i), b(i)} \tag{8}$$

where,

a(i) is the average distance between data point i and all other data points in the same cluster. This measures the cohesion of data i with its cluster.

b(i) is the average distance between data point i and all data points in the nearest neighboring cluster (i.e., the cluster to which data point i is not assigned). This measures the separation of data i from other clusters.

The overall Silhouette score for the clustering solution is the average of all

individual silhouette scores. To determine the optimal number of clusters using silhouette analysis, you calculate the silhouette score for different values of k(number of clusters) and choose the value that maximizes the average silhouette score. A higher average silhouette score suggests a better-defined clustering solution with well-separated and compact clusters.

A silhouette plot is a visual representation used to evaluate the quality of clustering results. It is a graphical tool that helps assess how well data points are assigned to their respective clusters. The silhouette plot displays individual data points represented as vertical bars within the plot. Each data point is assigned to a cluster, and the bars are color-coded based on their cluster assignment.

The silhouette value for each data point is calculated by considering its proximity to other data points within its own cluster (intra-cluster distance) and its proximity to data points in the nearest neighboring cluster (inter-cluster distance). A high silhouette value indicates that the data point is well-clustered and relatively closer to its own cluster compared to neighboring clusters. The silhouette plot allows us to visually inspect the distribution of silhouette values across all clusters. Ideally, we want to see predominantly positive silhouette values, which indicate well-separated and compact clusters.

The silhouette plot analysis is a valuable tool for identifying the optimal number of clusters (k) in a dataset. By considering specific criteria from the silhouette plot, we can determine the best k value for the clustering task. Firstly, it is important to select a value of k that ensures all clusters have

silhouette scores above the average silhouette score. This criterion helps in identifying consistent and well-separated clusters, as clusters with below-average silhouette scores indicate inconsistencies in the clustering process. Secondly, observing wide fluctuations in the size of silhouette plots suggests uneven cluster sizes, which is unexpected for balanced data. Therefore, we should aim for a value of k that leads to more uniform and evenly sized clusters. Additionally, the presence of a large number of data points with negative silhouette scores signifies that a substantial amount of data is being wrongly clustered, implying that these data points are closer to the next cluster than their assigned one. By taking these criteria into account, along with aiming for higher silhouette scores, we can pinpoint the optimal number of clusters that result in a more reliable and well-defined clustering solution for the given dataset.

Silhouette analysis provides a valuable and intuitive way to assess the quality of clustering results and helps in selecting the optimal number of clusters for a given dataset. However, it is essential to be cautious when interpreting the results, especially if the silhouette scores are close to 0, as it may indicate overlapping or poorly separated clusters. Therefore, it is recommended to combine silhouette analysis with other clustering evaluation methods to make a more informed decision about the optimal number of clusters.

- **Calinski-Harabasz Index**: Calinski-Harabasz Index evaluates the quality and compactness of the clustering solution based on the ratio of between-cluster variance to within-cluster variance. The index aims to identify the

number of clusters that results in well-separated and tightly packed clusters. The number of clusters corresponding to the highest Calinski-Harabasz Index is considered optimal for the problem.

The Calinski-Harabasz Index is calculated using the following formula:

$$CH = \frac{B(k)}{W(k)} X \frac{n-k}{k-1} \tag{9}$$

Where:

CH is the Calinski-Harabasz Index.

B(k) is the between-cluster variance, representing the variance between cluster centroids.

W(k) is the within-cluster variance, representing the variance within each cluster.

n is the total number of data points in the dataset.

k is the number of clusters being evaluated.

When using the Calinski-Harabasz Index, it is essential to be mindful of some limitations. The index can be sensitive to the dataset's size and dimensionality, and it may not perform well with irregularly shaped clusters or datasets with varying densities. Therefore, it is advisable to combine the Calinski-Harabasz Index with other clustering evaluation metrics to get a comprehensive understanding of the optimal number of clusters for a given dataset.

◇ **Davies-Bouldin Index(DB Index)**: This index evaluates the average similarity between each cluster and its most similar cluster. A lower Davies-Bouldin Index indicates better-defined clusters. The number of clusters with the lowest index is chosen as the optimal number.

$$DBI = \frac{1}{k}\sigma_{i=1}^{k} max_{j \neq i}(\frac{S_i + S_j}{M_{ij}}) \tag{10}$$

Where:

DBI is the Davies-Bouldin Index.

k is the number of clusters being evaluated.

S_i is the average distance (similarity) of each data point in cluster i to the centroid of cluster i.

S_j is the average distance (similarity) of each data point in cluster j to the centroid of cluster j.

M_{ij} is the distance between the centroids of clusters i and j.

To determine the optimal number of clusters using the Davies-Bouldin Index, you would calculate the index for different values of k and choose the value of k that minimizes the Davies-Bouldin Index. The k value that leads to the smallest Davies-Bouldin Index is considered the optimal number of clusters for the given dataset.

It is important to note that the Davies-Bouldin Index should be used in conjunction with other clustering evaluation metrics to gain a comprehensive

understanding of the clustering solution's quality and stability. Additionally, domain knowledge and practical interpretation of the clustering results should also be considered when determining the final number of clusters.

- ◇ **Hierarchical Clustering**: Hierarchical clustering can be visualized using a dendrogram, a tree-like diagram illustrating data points merging into clusters based on similarity or distance, displaying step-by-step relationships and grouping data points into clusters at various similarity levels. The optimal number of clusters can be determined by cutting the dendrogram at a height that best fits the problem's requirements.

- ◇ **Density-Based Clustering**: In density-based methods like DBSCAN, the number of clusters is not explicitly specified. Instead, the algorithm identifies clusters based on density-reachability relationships. The number of clusters can be considered to be the number of distinct clusters found by the algorithm. Various hyper-parameter and their relation to the features of data is a huge factor in determining the number of clusters in this algorithm.

- ◇ **Domain Knowledge**: Expert domain knowledge or specific problem requirements may also help determine the appropriate number of clusters based on the practical interpretation of the data. In our application, we tried various methods to find the optimum number of clusters but we always put an eye out that most of the data from the same APC should fall in the same or neighboring cluster.

Sorting of clusters

Clustering enables us to group similar data into different clusters, but it doesn't provide specific information about which cluster corresponds to a particular level of compaction density. To gain a better understanding of the relationship between data and compaction density, we devised a mechanism to organize the clusters based on certain features. Several options were considered for sorting, such as Total Power, Power in the fundamental frequency, Average Pass-ID, and Average Pass Count. Since Power in the fundamental frequency is already used in clustering, it was not the most suitable option. We also found that Average Pass-ID did not effectively represent the rolling pattern as Average Pass Count (APC) did. After careful consideration, we concluded that ordering the clusters using APC or Total Power would be the most effective approach.

In the final step, the points within clusters possessing the smallest average ordering feature are assigned the value 1, the next smallest cluster is assigned 2, and so on. This establishes a relationship where the clustering order is determined by the powers. We use this clustering order to identify the cluster to which a vibration at a specific time step belongs. Additionally, we leverage the Average Pass Count (APC) to estimate the density level of the corresponding cluster. This process enables us to effectively determine the compaction density associated with each cluster.

This novel idea in our research allowed us to incorporate the rolling pattern into the Learning Phase itself, resulting in significant improvements to the performance

of our Intelligent Compaction mechanism. By sorting the clusters based on relevant features, we were able to gain valuable insights and enhance the overall efficiency of our approach.

4.2 Mapping Clusters to Density

Finally, our task was to convert the sorted clusters into density predictions. The main objective here was to establish a suitable function that could map the clusters to corresponding density values. To accomplish this, we obtained data on the minimum possible density at which the pavers lay the mix(d_{min}) and also the maximum density achievabled_{max}, determined from tests conducted in the HMA mixing laboratory. These two density values served as reference points for our function fitting process, enabling us to accurately convert the clusters into corresponding density predictions.

Exactly, in the clustering process, the cluster that represents the lowest Average Pass Count (APC) should be associated with the minimum possible density d_{min}, as it indicates the lowest level of compaction achieved. On the other hand, the cluster that represents the highest APC should correspond to the maximum achievable density d_{max}, as it signifies the highest level of compaction that can be attained. These are going to the limits of the prediction function.

We tried linear and sigmoid functions to convert data from cluster numbers to density prediction.

- **Linear fitting**: We tried a straight-line relationship between the cluster

number and density prediction. In a linear function, the output variable y (density) is directly proportional to the input variable (cluster number) with a constant rate of change, known as the slope or gradient.

The general form of a linear function can be expressed as:

$$y = mx + c \tag{11}$$

where:

y is the density prediction.

x is the cluster number predicted.

m represents the slope or gradient, which determines how steep or flat the line is. It signifies the rate of change of y with respect to x.

b is the y-intercept, which is the value of y when x is equal to zero. It indicates the point where the line intersects the y-axis.

We just need two points to fit the function and we used (1,d_min) and (k,d_max), where k is the optimum number of clusters selected, to create the linear fitting. After fading the m and c, we predict the density using the cluster number determined in the last step.

- **Sigmoid function**: We observe that low clusters are extensively present in low APC while high clusters are present in high APC in abundance. Whereas, there is not much of an average cluster present in either condition. From this, we can infer that there is an abrupt change in low to high density after some specific amount of compaction rather than linear. So, we tried sigmoid

fitting. The general form of a sigmoid function is:

$$f(x) = L + (\frac{1}{1 + e^{-m(x-x_0)}})X(H - L) \tag{12}$$

where: f(x) is the output value of the sigmoid function for a given input x.

L is the lower limit of the function i.e. d_min

H is the highest limit of the function i.e. d_max

e is the base of the natural logarithm (approximately 2.71828)

m is the slope of the curve, determining the steepness of the sigmoid

x0=$\frac{k+1}{2}$ is the x-coordinate of the midpoint, which indicates the point where the sigmoid function transitions from the lower value to the upper value

k=optimum number of cluster

Here, at any point, at most we want the slope to be 1, so we specify m=1. Then, as per the project, we specify L, H, and k and predict the density using the cluster number determined in the last step.

5. EVALUATION METRICS

5.1 Metrics for evaluating Power Clustering Algorithms

Evaluation metrics play a crucial role in assessing the effectiveness and quality of clustering algorithms. Clustering is an unsupervised learning technique that aims to group similar data points together based on their intrinsic characteristics. However, determining the optimal number of clusters and evaluating the clustering results can be challenging tasks. Evaluation metrics provide objective measures to quantify the performance of clustering algorithms and help in selecting the most appropriate clustering solution for a given dataset. These metrics enable us to compare different clustering methods, assess the consistency of cluster assignments, and make informed decisions about the quality of the clustering results. A range of evaluation metrics, including the Silhouette Score, Davies-Bouldin Index, and Calinski-Harabasz Index, have been devised to assess distinct facets of clustering evaluation, such as cluster compactness and separation. These metrics can be applied both for determining the optimal number of clusters for specific algorithms, as discussed in Section 4.1.2, and for identifying the most suitable clustering approach in this context. The careful selection and application of these metrics are crucial for gaining insights into the clustering performance and ensuring the reliability and validity of the clustered data for subsequent analyses and decision-making processes.

- Silhouette Coefficient:

  The silhouette coefficient is a measure used to evaluate the quality of clustering

results in unsupervised machine learning. It assesses how well each data point fits into its assigned cluster while also considering its similarity to other clusters. The Silhouette coefficient calculation is described in 8. The silhouette coefficient ranges from -1 to 1, where:

1. A coefficient close to +1 indicates that the data point is well-clustered and located far away from neighboring clusters.
2. A coefficient close to 0 indicates that the data point is near the decision boundary between two clusters.
3. A coefficient close to -1 indicates that the data point may have been assigned to the wrong cluster.

A higher silhouette coefficient indicates better-defined and well-separated clusters, while a lower value suggests that the clustering may not be appropriate or that data points are assigned to incorrect clusters.

- The Davies-Bouldin Index (DB index): DB Index, equation 10, is a clustering evaluation metric used to assess the quality of clustering results. It quantifies the compactness of individual clusters and the separation between different clusters in a clustering solution. Conversely, a higher DB index implies that the clusters are less distinct and may overlap, indicating suboptimal clustering. The lower the DB index, the better the clustering quality.
- Calinski-Harabasz Index (CH Index): The Calinski-Harabasz Index, also known as the Variance Ratio Criterion and calculated as in equation 9, is a clustering evaluation metric used to assess the quality of clustering results.

It quantifies the separation between clusters and the compactness of each cluster in a clustering solution. A higher Calinski-Harabasz Index value indicates better-defined and well-separated clusters.

All these evaluation metrics look at the similarity between the intra-cluster data and the difference between inter-cluster measures but we also need to satisfy a few application-based metrics. So, we created a few application-based metrics to measure output of clustering.

It is observed that first passes will have low compaction density and is highly affected by underlying substrate composition and thus is very uneven. But when the roller finishes its rolling pattern, the compaction density is high and smooth. So, two major concepts taken into account while creating these metrics are that in the final passes of well constructed pavements, value of the cluster mean should be high and there should be high amount of smoothness of cluster values in final passes.

- **Cluster mean in final pass(f_{mean}):** For any given stretch, the final pass is the last pass the roller vibrates in it. It is supposed to have compacted to the highest quality possible. Looking at the mean of the cluster predicted here to let us know if the clustering done can detect the high-density regions or not. A higher mean in the final pass in the cluster is desirable.

- **Standard deviation in final pass(f_{std}):** Standard deviation is a statistical measure that quantifies the amount of variability or dispersion in a set of data points. It indicates how much individual data points deviate from the

mean (average) of the dataset. It is desired to have a uniform high cluster prediction in the final pass, so a smaller standard deviation in the final pass is a measure of the smoothness of clustering.

- **One lag auto-correlation of cluster predicted in final pass (*one_lag_auto*):** Autocorrelation, also known as serial correlation, is a statistical concept that measures the degree of similarity between a time series data and a delayed version of itself. In other words, it quantifies the relationship between an observation at a particular time and previous observations in the same time series. One lag autocorrelation is autocorrelation between time series data delayed by one time_step. Autocorrelation can be used as the measure of the smoothness of the data.

  The formula for autocorrelation at lag k in a time series data is given by:

$$Autocorrelation\ at\ lag\ k = \frac{\sum_{t=k+1}^{N}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2} \tag{13}$$

  where:

  N is the total number of observations in the time series data

  x_t is the value of the time series at time t

  x_{t-k} is the value of the time series at time $t - k$ (lagged version)

  $\bar{x}$ is the mean of the time series data

  The larger the value of $final\_autocorr$, the smoother the prediction in the final pass. Thus, a higher value of *one_lag_auto* is desirable.

After determining the proper clustering mechanism, we used those clusters

to convert to density values. Results of these values are directly compared to the ground truth values and different results were compared using evaluation metrics designed to asses density prediction.

5.2 Metrics for evaluating Cluster to Density Algorithms

Evaluating the final density prediction to the ground_truth is the final measure of the IC system we are developing. Even though we are using clustering as pattern recognition too, we are to predict the compression density as if we were doing regression. So, to evaluate the prediction of density, we are using the RMSE(Root Mean Squared Error) and R^2 (R-squared) metric.

RMSE(Root Mean Squared Error)

The Root Mean Squared Error (RMSE) is a commonly used evaluation metric in regression analysis and machine learning to measure the accuracy of a predictive model. It quantifies the difference between the predicted values of the model and the actual (observed) values of the target variable. The RMSE provides a measure of how well the model's predictions match the true values, and it is particularly useful when dealing with continuous numeric data.

Mathematically, the RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{n}} \tag{14}$$

where:

n is the total number of data points in the dataset.

y_i is the actual (observed) value of the target variable for data point i.

$\bar{y_i}$ is the predicted value of the target variable for data point i generated by the regression model.

The RMSE computes the average of the squared differences between the predicted and actual values, takes the square root of the result, and provides a single scalar value representing the average prediction error. The lower the RMSE, the better the model's predictions align with the actual data, indicating a more accurate and precise predictive model.

The RMSE is sensitive to outliers, as large errors in predictions can lead to higher values of RMSE. It is commonly used as part of model evaluation and selection to compare different regression models or to tune hyperparameters in the model to achieve the best performance.

R2 (R-squared)

The R2 (R-squared) evaluation metric, also known as the coefficient of determination, is a statistical measure used to assess the goodness of fit of a regression model. It indicates the proportion of the variance in the dependent variable (predicted density) that is explained by the independent variables (predictor variables) in the model.

The R2 value ranges from 0 to 1, where:

- R2=0 indicates that the model does not explain any of the variance in the

dependent variable. In other words, the model does not fit the data well.

- R2=1 indicates that the model explains all the variance in the dependent variable. The model perfectly fits the data.

Mathematically, the R2 value is calculated as follows:

$$R2 = 1 - \frac{Sum\ of\ Squared\ Residuals}{Sum\ of\ Squared\ Residuals} \tag{15}$$

where:

The Sum of Squared Residuals is the sum of the squared differences between the actual values of the dependent variable and the predicted values from the regression model

The Total Sum of Squares is the sum of the squared differences between the actual values of the dependent variable and the mean of the dependent variable

The R2 score varies from 0 to 1, with larger values signifying a stronger alignment between the regression model and the data. Nevertheless, it's vital to recognize that a high R2 score doesn't automatically imply the model's suitability or the existence of a cause-and-effect link between the predictor and target variables. Instead, it reflects how much the model explains variability.

Therefore, when appraising the effectiveness and credibility of the predictive model, it's prudent to jointly consider both the root mean squared error (RMSE) and the R2 score. This holistic evaluation approach ensures a thorough and well-rounded assessment.

6. RESULTS AND DISCUSSION

6.1 Results of clustering step

Detailed examination is conducted on data sourced from two significant HMA (Hot Mix Asphalt) pavement construction projects: one located in Shawnee, Oklahoma, and the other in Lathrop, California. For each of these projects, two distinct sets of data segments are scrutinized at each phase of analysis. This approach facilitates the generalization of findings across diverse construction methodologies employed by different companies at varying geographical sites.

In each case, the initial data segment serves as training data for clustering, while subsequent segments are utilized for validation purposes. This methodology enhances our comprehension of the data and affords us comprehensive insight into the behavior across entire stretches, enabling location-based analyses. However, during the testing phase, the entirety of remaining project data is employed. This strategy ensures that evaluations encompass a wide array of real-world construction site scenarios.

The training and validation processes involve the deliberate selection of distinct stretches from the time-series data. These chosen stretches exhibit clear and discernible multiple passes, which is essential for accurate Asphalt Pavement Compactor (APC) calculations during the training phase and for result interpretation within the validation dataset. Each of the selected stretches comprises five passes, contributing to the robustness of the analysis.

The amount of data in number of time_steps in the constrained stretches is listed below:

- Lathrop Project:
  - Test Stretch: 2558

    Pass1:517 Pass2:508 Pass3:533 Pass4:492 Pass5:508
  - Test Stretch: 2234

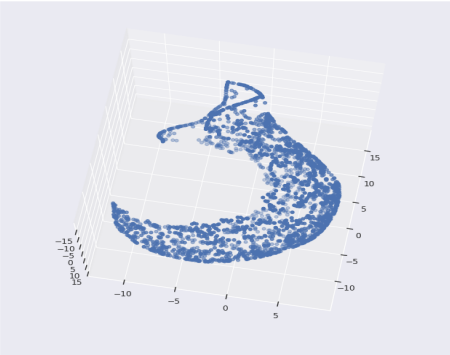
    Pass1:438 Pass2:443 Pass3:451 Pass4:451 Pass5: 451
- Shawnee Project:
  - Test Stretch: 1805

    Pass1:357 Pass2:359 Pass3:365 Pass4:365 Pass5:359
  - Test Stretch: 1054

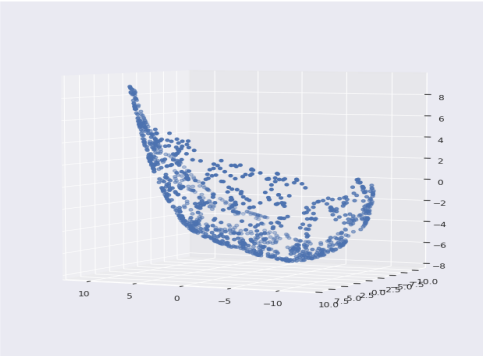
    Pass1:225 Pass2:225 Pass3:185 Pass4:186 Pass5:233

Visualization

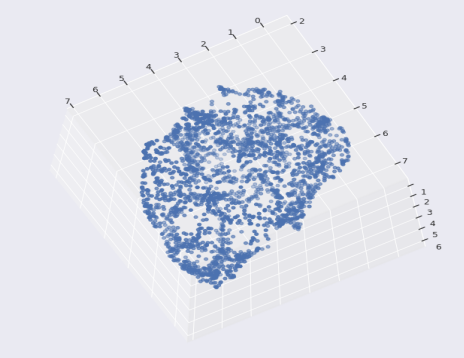
We employed t-SNE and UMAP as feature extraction mechanisms and used visualization of those resultant features to gain insights into the underlying data patterns. These visualization techniques allowed us to observe the data in lower-dimensional spaces, making it easier to identify clusters and potential relationships among the data points. By employing t-SNE and UMAP, we aimed to gain a better understanding of the structure and grouping of the power features, which could potentially reveal valuable insights regarding the compaction process and its variations during pavement construction.



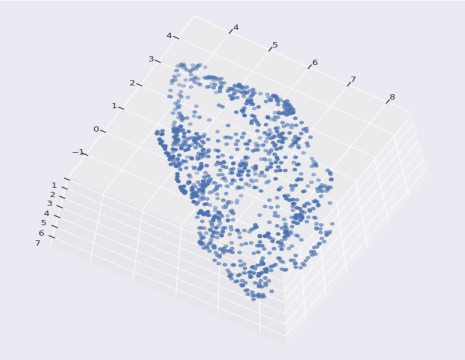
(a) 3D t-SNE to see if there are some clear clusters in Lathrop train data



(b) 3D t-SNE to see if there are some clear clusters in Shawnee train data



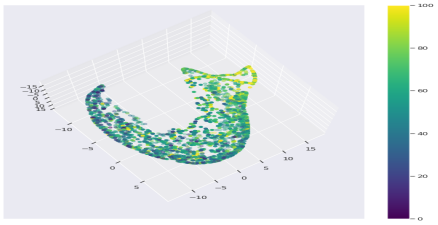
(c) 3D UMAP to see if there are some clear clusters in Lathrop train data



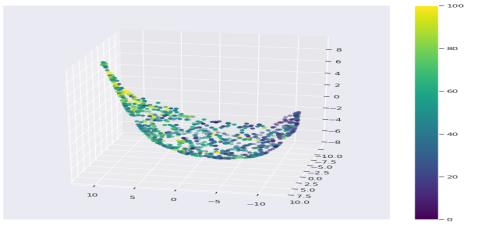
(d) 3D UMAP to see if there are some clear clusters in Shawnee train data

Figure 6.1: Visualization of the data in t-SNE and UMAP space to observe if there is existence of clusters

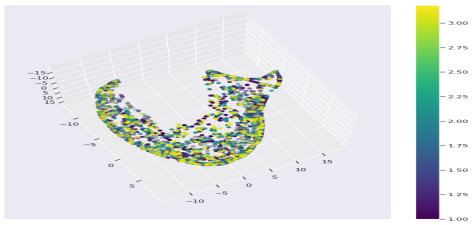
Upon examining the t-SNE and UMAP visualizations of the training data in Figure 6.1, it became apparent that distinct clusters of the data points were not present. This observation aligns with our expectations since our system deals with continuous data, making the expectation of discrete clusters infeasible. Despite the absence of clear clusters, we decided to further investigate the distribution of the t-SNE and UMAP visualizations concerning the average pass count and total power characteristics. By analyzing the visualizations in relation to these attributes, we aimed to gain a better understanding of any potential patterns or trends that might emerge in the data.



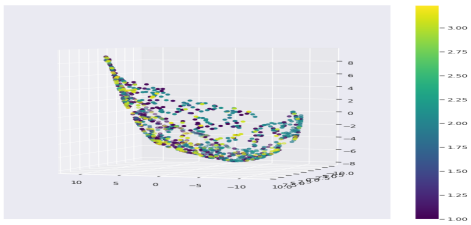
(a) Visualize using t-SNE to see the existence of a trend of power change with the features in Lathrop train data



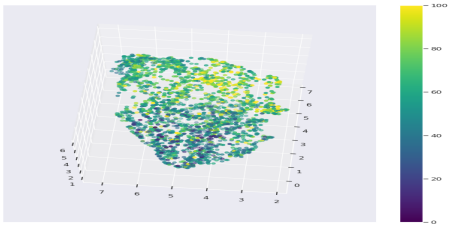
(b) Visualize using t-SNE to see the existence of a trend of power change with the features in Shawnee train data



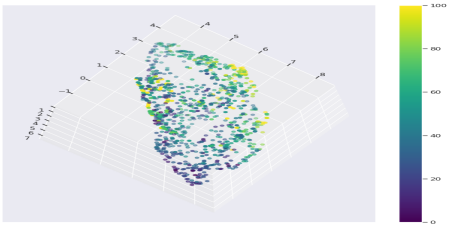
(c) Visualize using t-SNE to see the existence of a trend of APC change in Lathrop train data



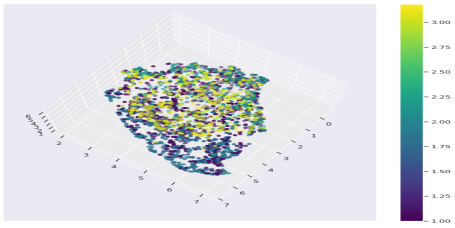
(d) Visualize using t-SNE to see the existence of a trend of APC change in Shawnee train data



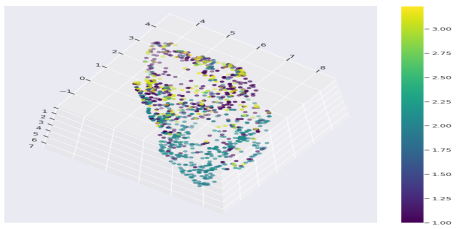
(e) Visualize using UMAP to see the existence of a trend of power change in Lathrop train data



(f) Visualize using UMAP to see the existence of a trend of power change in Shawnee train data

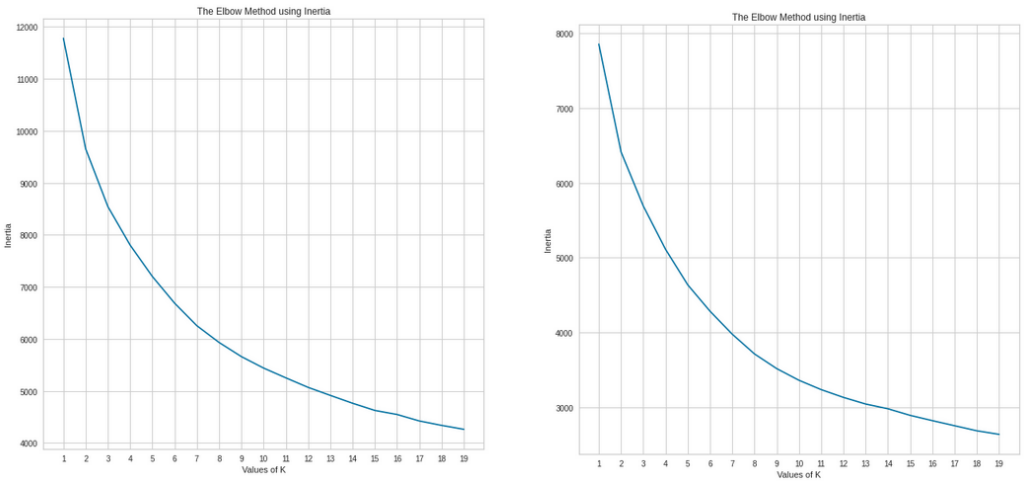


(g) Visualize using UMAP to see the existence of a trend of APC change in Lathrop train data



(h) Visualize using UMAP to see the existence of a trend of APC change in Shawnee train data

Figure 6.2: Visualization of power and APC trend on the training data



(a) Change in inertia with K at Lathrop_train(b) Change in inertia with K at Shawnee_train

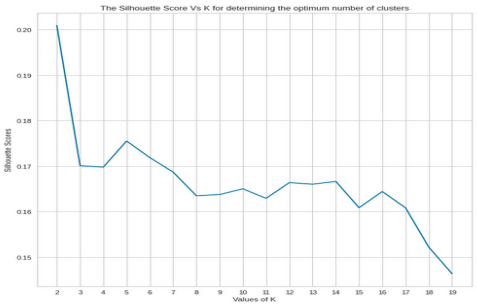
Figure 6.3: Change in Inertia Value with Increasing Number of Clusters (k) in K-Means. Clustering at Figure 6.3b Lathrop_location_1 and Figure 6.3b Shawnee_location_1

Based on the Figure 6.2, we noticed the absence of clear definitive clusters in the data. However, a discernible trend in power emerged, with a concentration of data points representing high powers in specific regions. The same pattern holds for Average Pass Count (APC) as well. This observation provides crucial insights, suggesting that we can utilize total_power or APC as attributes to guide the ordering of the clusters we create in subsequent steps.

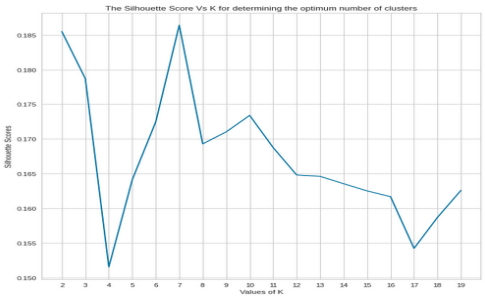
Optimum clustering mechanism

The initial step in our analysis involved determining the optimal number of clusters. To accomplish this, we conducted an exhaustive search by employing various clustering algorithms and implemented the cluster selection techniques mentioned in Section 4.1.2. Initially, we applied the K-Means algorithm with multiple clusters and evaluated different clustering indices. Subsequently, we determined the appropriate base cluster size for our further analysis.

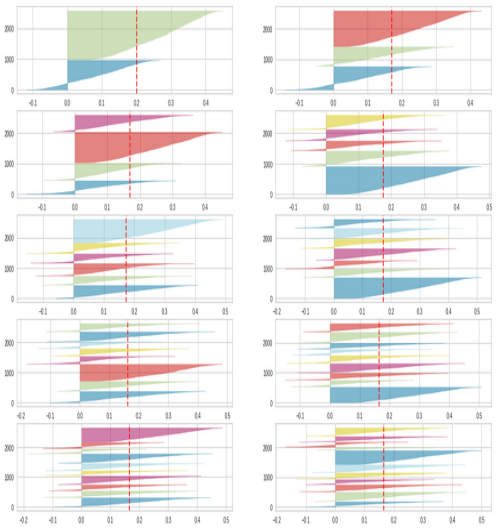
First, we implemented the elbow method based on cluster inertia to find the optimum number of clusters. We observed the decrease in inertia values for different K-values in two projects, Shawnee and Lathrop as seen in Figure 4.2. The curves obtained are pretty continuous and there is no clear elbow position in any of the K values. So, using this we cannot have any conclusive value of K.



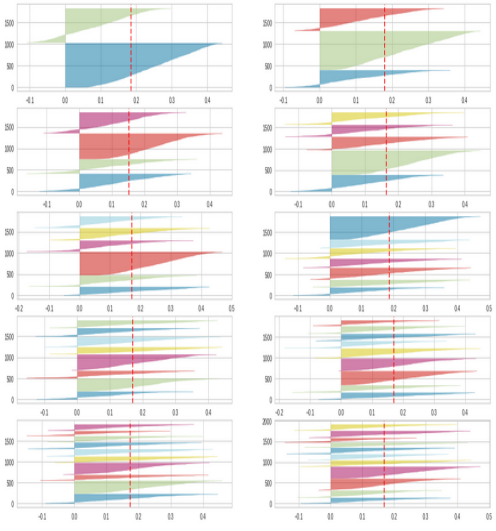
(a) Change in average silhouette coefficient with respect to K at Lathrop_location_1 when using K_Means clustering



(b) Change in average silhouette coefficient with respect to K at Shawnee_location_1 when using K_Means clustering



(c) Shiloutte plot for K_Means clustering using K from 2 to 11 at Lathrop_location_1



(d) Shiloutte plot for K_Means clustering using K from 2 to 11 at Lathrop_location_1

Figure 6.4: Silhouette graphs for various cluster sizes while using K-Means clustering to determine the optimum number of clusters

The overall trend of the silhouette score is in decreasing order with respect to the increase in K, as seen in Figure 6.4(a) and Figure 6.4(b) at Lathrop and

Shawnee projects respectively. Looking into the silhouette score VS K plots only, we see that the silhouette score value, for K greater than two, is maximum K=5 and so is the optimum cluster number for the Lathrop project but for Shawnee, it is at K=7. But when we dive deep into the silhouette plot, where we also look at the number of negative coefficients in each cluster, we observe that using K=7 at Shawnee is not always the best option. Since the amount of data points with negative values is similar in the range of k=5,6,7,8, and 9; using any of these are equally near-optimal solutions when using K-Means Clustering.

Likewise, when we utilized DB and CH indices to assess the clustering process, we obtained near-optimal solutions with either 4, 5, or 7 clusters. These are the k values where CH is maximum or DB is minimum.

Table 6.1: Optimum number of the cluster for K-Means and Gaussian Mixture Model(GMM) clustering using feature with no transformation

Clustering	Methods	Optimum number of cluster(K)	
		Shawnee	Lathrop
K-Means	Elbow(Inertia)	Inconclusive	Inconclusive
	Silhouette	5	5/7
	DB	7	4
	CH	5	4
GMM	Silhouette	7	7
	DB	7	5
	CH	3	5

Based on Table 6.1, when utilizing K-Means as a clustering technique, we find that the optimal value of K can be 4, 5, or 7, depending on the specific project and the evaluation parameter used. Similarly, in our exploration with Gaussian Mixture Model (GMM), as shown in Table 6.1, we discovered that the optimum

number of clusters also falls in the range of 3, 5, or 7, depending on the evaluation parameter and project.

This suggests that when employing distance-based clustering mechanisms, it is preferable to embrace a dynamic number of clusters tailored to the specific characteristics of the project. Nevertheless, if there is a need to set a fixed value for K, opting for 5 appears reasonable, considering it represents a central near-optimal choice for K and also the most frequently occurring one in the above result at Table 6.1.

We also experimented with Adaptive Gaussian Mixture Model (AGMM) based on the information content. In this context, we observed that the number of clusters selected varies significantly depending on the hyperparameter's definition for the start of the cluster size, from which the model attempts to reduce the cluster number to the optimum value called max_cluster. From Table 6.2, it is

Table 6.2: Numbers of clusters selected by Adaptive Gaussian Mixture Model for clustering

Max_Cluster	Number of Clusters selected	
	Shawnee	Lathrop
5	5	5
10	9	9
15	14	13
20	19	17
25	23	24

evident that as the value of max_cluster increases, the number of clusters opted by AGMM also increases. This indicates that AGMM is not a suitable fit for clustering the data because it fails to effectively reduce the number of clusters when starting from a higher value of max_cluster.

Table 6.3: Numbers of clusters selected by DBSCAN (Here, Sh signify Shawnee project and La signify Lathrop project)

	Number of cluster selected					
min_sample	<=2	=4		=6		>=8
	Sh/La	Sh	La	Sh	La	Sh/La
eps=0.6	>15	10	10	0	0	0
eps=0.8	>15	7	7	1	1	1
eps=1	>15	7	5	5	5	1
eps=1.2	>15	5	5	2	2	1
eps=1.4	>15	2	2	1	1	1

Finally, we also tried DBSCAN for the sole purpose to observe what optimum number of cluster it provide for our data. The two major hyperparameters for DBSCAN are eps and min_sample as described in Section 4.1.1. A small change in any of these will cause a change in the number of clusters.

Based on Table 6.3, it is evident that varying hyperparameter values result in a significant difference in the number of clusters obtained. Moreover, there is no single value for the number of clusters that appears across multiple hyperparameter combinations. To use the DBSCAN, one needs to tune the hyperparameter eps and min_sample and make the decision. Among the different values observed, K=5 occurs most frequently, indicating that using a density-based selection of the number of clusters, K=5 is the preferable choice compared to others. As the result of using density-based and distance-based algorithms is similar, we move ahead to do experiments in distance-based clusterings.

Then, various feature extraction mechanisms were tested to see if better transformations can be obtained to have better clustering results. For this, PCA, Isomap, and Kernel-PCA were tried. In this experiment, we first transformed the features into PCA, KPCA, and Isomap space and clustered using K-Means or

Gaussian mixture Model for different numbers of transformed features. Finally, the Silhouette Score, CH Index, and DB Index of each of the permutations is calculated and analyzed.

Upon utilizing the transformed features in K-Means clustering and evaluating the results, the outcomes are displayed in Table 6.4. For the Shawnee project, we have the best silhouette score to be on the original feature, using 6 or 7 features in PCA transformation. It is the same case when evaluating using the CH Index or DB Index in the Shawnee project. In the case of the Lathrop project, we see that no transformed feature can match the original feature in the case of the Silhouette Coefficient, CB Index or DB Index. It is evident that the best Silhouette Coefficient, CH Index, and DB Index are obtained when using the original features without any transformation and with K=5. As a result, in the subsequent steps involving K-Means clustering, we choose to employ the original features without any transformation and set K=5.

We also performed a similar exhaustive experiment using Gaussian Mixture Model and the result can be soon in Table 6.5. Here, in the Shawnee project, we observe that the best silhouette score is when using one or two principle components with maximum information in PCA space. While evaluating DB or CH score, the best is when using only the first principle component. The case in Lathrop project data is also exactly the same, to use one or two principle components in PCA space.

So, for experiments in the next sages, the three best clustering methods we have selected are:

Table 6.4: Result of use of various Feature Extraction methods in K-Means Clustering of power data. Here, Sh signifies Shawnee project and La signifies Lathrop project

No. of features	Evaluation Metrics	Original Features		PCA		KPCA		Isomap	
		Sh	La	Sh	La	Sh	La	Sh	La
1	Silh	–	–	0.02	0.02	-0.15	-0.1	0.02	0.02
	CH	–	–	154.64	171.55	48.67	66.61	150.54	157.6
	DB	–	–	4.31	4.7	9.47	8.98	4.43	4.88
2	Silh	–	–	0.08	0.08	-0.11	-0.11	0.07	0.08
	CH	–	–	223.72	262.19	78.98	81.53	199.98	250.5
	DB	–	–	2.49	2.57	4.04	3.59	2.58	2.66
3	Silh	–	–	0.1	0.11	-0.11	-0.1	0.1	0.1
	CH	–	–	254.62	303.32	70.08	90.63	242.15	283.03
	DB	–	–	2.15	2.19	2.49	2.68	2.15	2.31
4	Silh	–	–	0.13	0.13	-0.1	-0.09	0.13	0.12
	CH	–	–	271.38	326.55	69.14	88.47	257.11	306.53
	DB	–	–	1.92	1.91	2.38	2.12	1.98	1.98
5	Silh	–	–	0.12	0.13	-0.1	-0.09	0.13	0.12
	CH	–	–	274.61	330	70.75	89.16	260.29	307.57
	DB	–	–	1.92	1.9	2.34	2.12	1.93	1.98
6	Silh	–	–	0.14	0.13	-0.11	-0.09	0.13	0.12
	CH	–	–	282.44	330.38	66.42	89.76	266.11	304.83
	DB	–	–	1.82	1.9	2.58	2.13	1.89	1.98
7	Silh	0.14	0.13	0.14	0.13	-0.11	-0.08	0.13	0.12
	CH	283.46	336.28	283.46	336.28	71.23	92.35	265.48	308.81
	DB	1.81	1.88	1.81	1.88	2.59	2.16	1.9	1.92

i) K-Means with k=5 in original power feature space with all seven features (KMC7), Algorithm1

ii) GMM with k=5 using first principle component in PCA space as feature (GMM1), Algorithm2

iii) GMM with k=5 using first two principle components in PCA space as features (GMM2): Algorithm3

Algorithm 1 KMC7: K-Means Clustering with K=5 taking all seven powers from original space as features

Input: Dataset X
Fix_hyperparameter: Number of clusters $K \leftarrow 5$
Output: Cluster centers C, Cluster assignments S
Initialize: Randomly choose K data points as initial cluster centers C
repeat
Assign: For each data point $x_i \in X$, find the nearest cluster center c_k and assign x_i to cluster k: $s_i = \arg\min_k \|x_i - c_k\|$
Update: For each cluster k, update the cluster center c_k as the mean of all data points assigned to the cluster: $c_k = \frac{1}{|s_i=k|} \sum_{x_i \in X} x_i \mathbb{1}(s_i = k)$
until Convergence
Return: Final cluster centers C and cluster assignments S

Algorithm 2 GMM1: PCA Feature Extraction and GMM Clustering with First Principal Component

Input: Dataset X
Fix_hyperparameter: Number of clusters $K \leftarrow 5$
Output: Cluster centers C, Cluster assignments S
PCA Feature Extraction:
Compute the mean: Calculate the mean vector $\bar{x}$ of the data points in X
Center the data: Subtract the mean vector $\bar{x}$ from each data point to get the centered data matrix $\hat{X}$
Compute the covariance matrix: Compute the covariance matrix C of the centered data matrix $\hat{X}$ as $C = \frac{1}{n}\hat{X}^T\hat{X}$, where n is the number of data points
Compute the eigenvectors and eigenvalues: Compute the eigenvectors and eigenvalues of the covariance matrix C. Let v_1 be the eigenvector corresponding to the largest eigenvalue.
GMM Clustering with First Principal Component:
Project data onto the first principal component: Project the data points in X onto the first principal component v_1 to obtain the 1-dimensional feature vectors.
Initialize GMM: Randomly initialize K cluster centers C.
repeat
Expectation Step: For each data point, compute the probability of belonging to each cluster using the current cluster centers.
Maximization Step: Update the cluster centers as the weighted mean of the data points, where the weights are the probabilities obtained in the expectation step.
until Convergence
Assign data points to clusters: Assign each data point to the cluster with the highest probability obtained in the expectation step.
Return: Final cluster centers C and cluster assignments S

Algorithm 3 GMM2: PCA Feature Extraction and GMM Clustering with First Two Principal Components

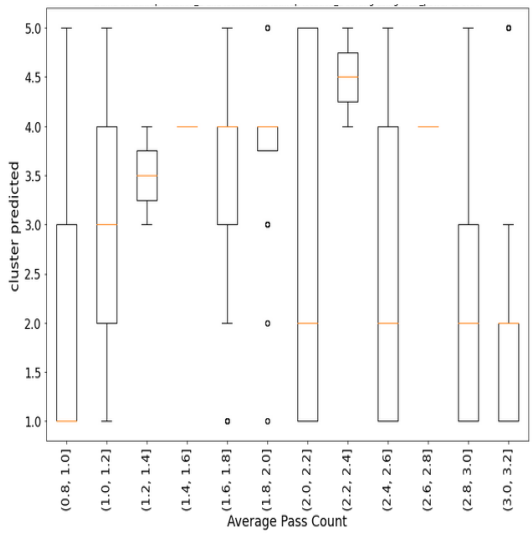
Input: Dataset X
Fix_hyperparameter: Number of clusters $K \leftarrow 5$
Output: Cluster centers C, Cluster assignments S
PCA Feature Extraction:
Compute the mean: Calculate the mean vector $\bar{x}$ of the data points in X
Center the data: Subtract the mean vector $\bar{x}$ from each data point to get the centered data matrix $\hat{X}$
Compute the covariance matrix: Compute the covariance matrix C of the centered data matrix $\hat{X}$ as $C = \frac{1}{n}\hat{X}^T\hat{X}$, where n is the number of data points
Compute the eigenvectors and eigenvalues: Compute the eigenvectors and eigenvalues of the covariance matrix C. Sort them in decreasing order and select the first two eigenvectors, denoted by v_1 and v_2.
GMM Clustering with First Two Principal Components:
Project data onto the first two principal components: Project the data points in X onto the plane spanned by the first two principal components v_1 and v_2 to obtain the 2-dimensional feature vectors.
Initialize GMM: Randomly initialize K cluster centers C.
repeat
Expectation Step: For each data point, compute the probability of belonging to each cluster using the current cluster centers.
Maximization Step: Update the cluster centers as the weighted mean of the data points, where the weights are the probabilities obtained in the expectation step.
until Convergence
Assign data points to clusters: Assign each data point to the cluster with the highest probability obtained in the expectation step.
Return: Final cluster centers C and cluster assignments S

Table 6.5: Evaluation Metrics of use of various Feature Extraction methods in GMM Clustering of power data

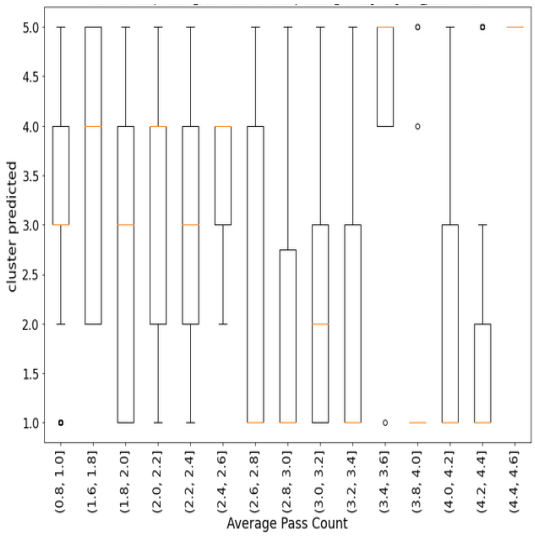
No. of features	Evaluation Metrics	Original Features		PCA		KPCA		Isomap	
		Sh	La	Sh	La	Sh	La	Sh	La
1	Silh	–	–	0.14	0.18	-0.14	-0.1	0.03	0.02
	CH	–	–	395.42	470.12	62.49	78.55	151.12	158.16
	DB	–	–	2.27	2.23	9.05	8.57	4.28	4.86
2	Silh	–	–	0.14	0.18	-0.05	-0.04	0.06	0.08
	CH	–	–	383.16	466.78	96.63	122.69	196.35	248.29
	DB	–	–	2.3	2.25	4.34	4.46	2.69	2.73
3	Silh	–	–	0.13	0.18	-0.04	-0.03	0.11	0.1
	CH	–	–	366.06	463.69	82.64	131.86	222.87	268.85
	DB	–	–	2.34	2.25	7.08	4.55	2.31	2.23
4	Silh	–	–	0.11	0.17	-0.04	0.01	0.1	0.12
	CH	–	–	269.34	458.04	83.84	146	205	296.87
	DB	–	–	2.73	2.28	7.01	3.79	2.17	2
5	Silh	–	–	0.09	0.12	-0.01	0.03	0.07	0.1
	CH	–	–	218.54	378	94.45	145.09	185.67	277.83
	DB	–	–	3	2.43	4.69	4.08	2.43	2.11
6	Silh	–	–	0.1	0.11	-0.03	0.03	0.05	0.02
	CH	–	–	232.86	334.35	90.93	142.4	164.6	175.78
	DB	–	–	2.93	2.54	5.4	4.18	2.45	2.63
7	Silh	0.09	0.11	0.1	0.11	0	0.01	0.07	0.07
	CH	247.85	340.57	247.85	340.57	97.81	131.27	181.72	243.24
	DB	2.79	2.57	2.79	2.57	4.57	4.82	2.43	2.41

Selection of cluster ordering criterion

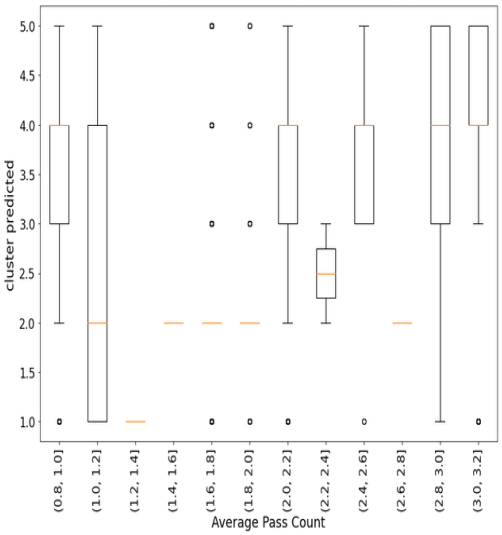
We have determined the optimal clustering mechanisms for our application in the previous step. However, the next major step is to identify which clusters represent highly compacted regions and which ones represent low compacted regions. To achieve this, we need to arrange these clusters based on certain characteristics. In addition to the seven power features, each time_step data also includes total power, passID, and APC characteristics. While PassID simply indicates the increase in pass count without considering any overlap between consecutive passes, APC takes into account both the progressive pass count over time and the overlap



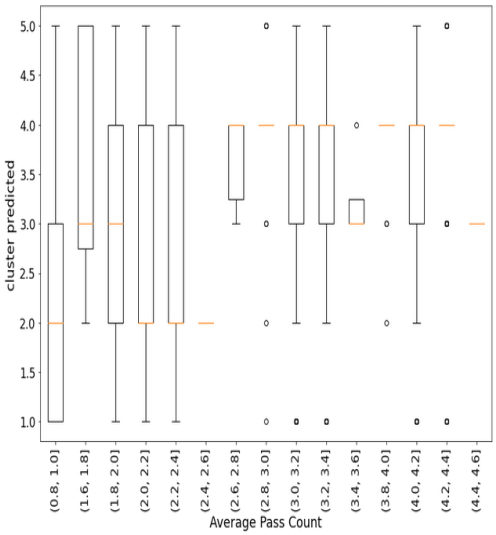
(a) Train data-set of Lathrop using total_power as ordering attribute



(b) Test data-set of Lathrop using total_power as ordering attribute



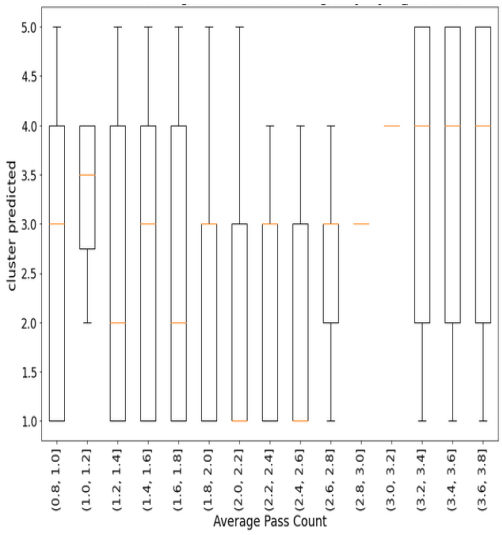
(c) Train data-set of Lathrop using APC as ordering attribute



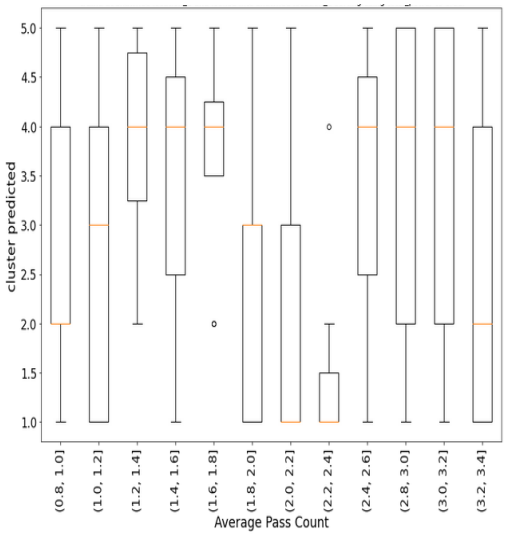
(d) Test data-set of Lathrop using APC as ordering attribute

Figure 6.5: Box and whisker plot illustrating the results of clustering using K-Means clustering, utilizing original power feature, on Lathrop project data with different cluster ordering attributes.

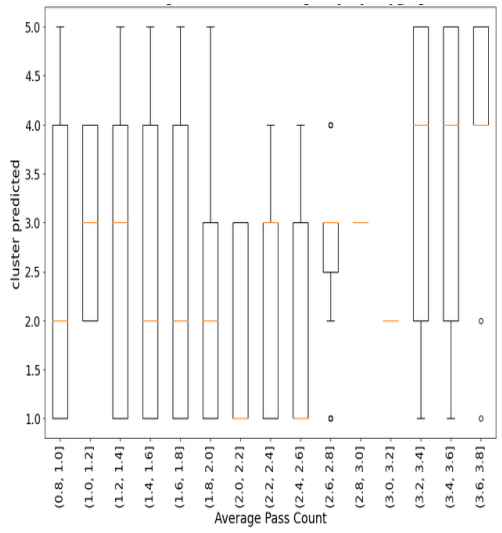
of the previous passes in the position of the current time_step. Total_power represents the weighted sum of powers in the fundamental frequency and the first six harmonics. Hence, it is expected to hold important implications for the amount of compaction. Consequently, we proceed to examine which characteristic, total_power or APC, is optimal to use as ordering criteria.



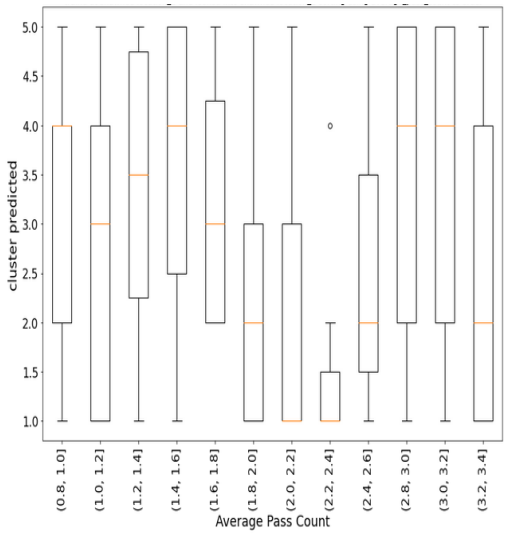
(a) Train data-set of Shawnee using total_power as ordering attribute



(b) Test data-set of Shawnee using total_power as ordering attribute

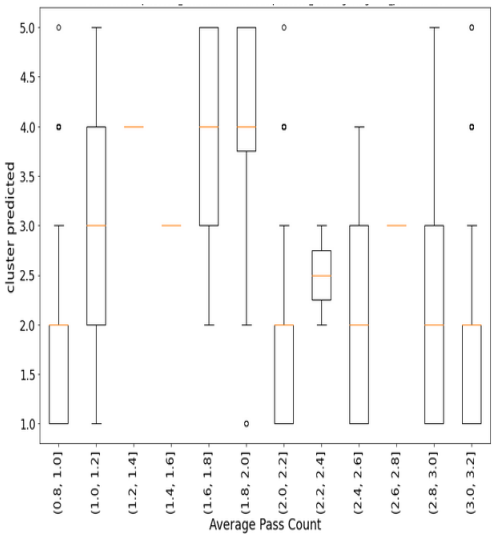


(c) Train data-set of Shawnee using APC as ordering attribute

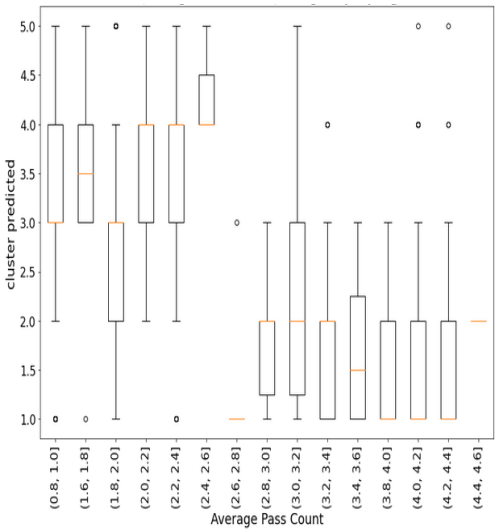


(d) Test data-set of Shawnee using APC as ordering attribute

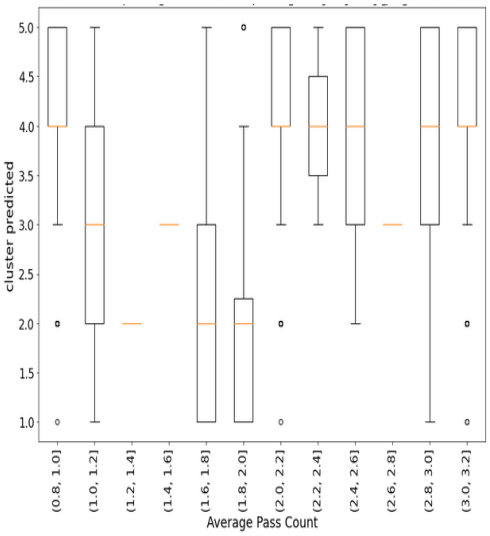
Figure 6.6: Box and whisker plot illustrating the results of clustering using K-Means clustering, utilizing original power feature, on Shawnee project data with different cluster ordering attributes.



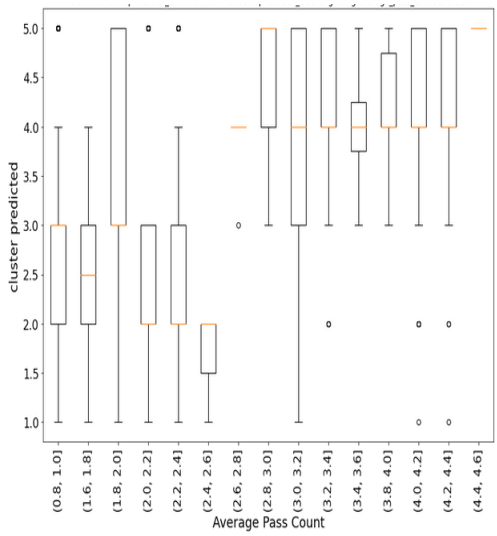
(a) Train data-set of Lathrop using total_power as ordering attribute



(b) Test data-set of Lathrop using total_power as ordering attribute

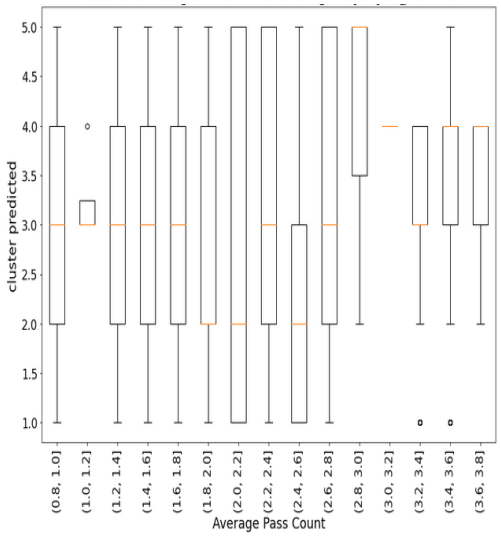


(c) Train data-set of Lathrop using APC as ordering attribute

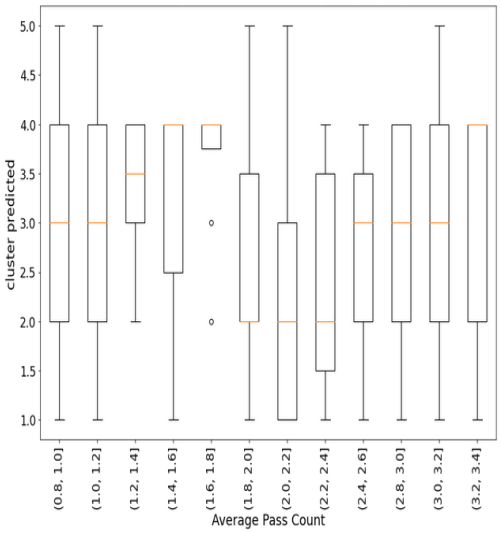


(d) Test data-set of Lathrop using APC as ordering attribute

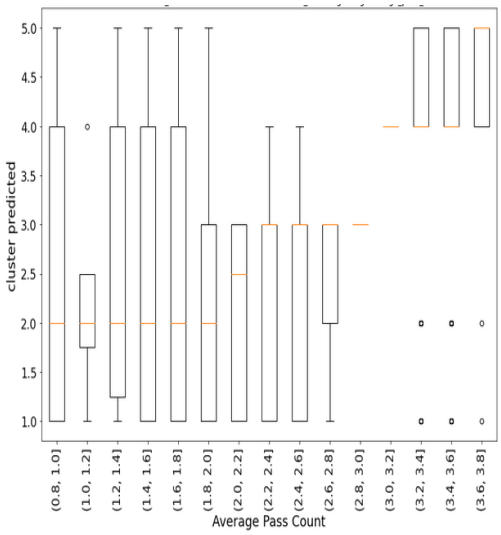
Figure 6.7: Box and whisker plot illustrating the results of clustering using Gaussian Mixture Model (GMM), utilizing the first principle component from PCA space as a feature, on Lathrop project data with different cluster ordering attributes.



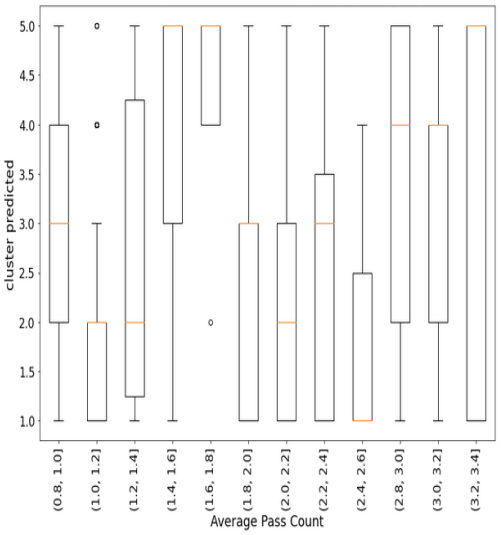
(a) Train data-set of Shawnee using total_power as ordering attribute



(b) Test data-set of Shawnee using total_power as ordering attribute

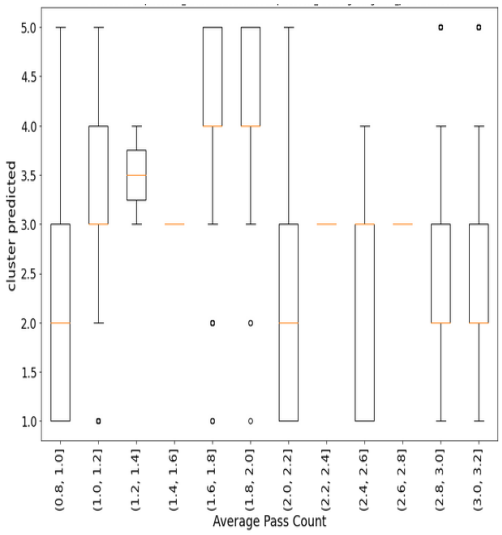


(c) Train data-set of Shawnee using APC as ordering attribute

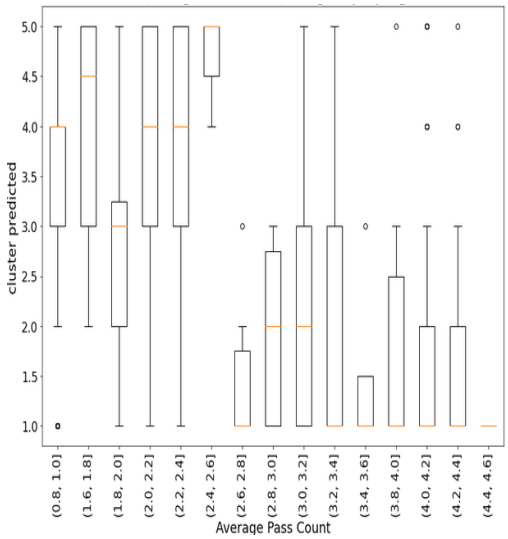


(d) Test data-set of Shawnee using APC as ordering attribute

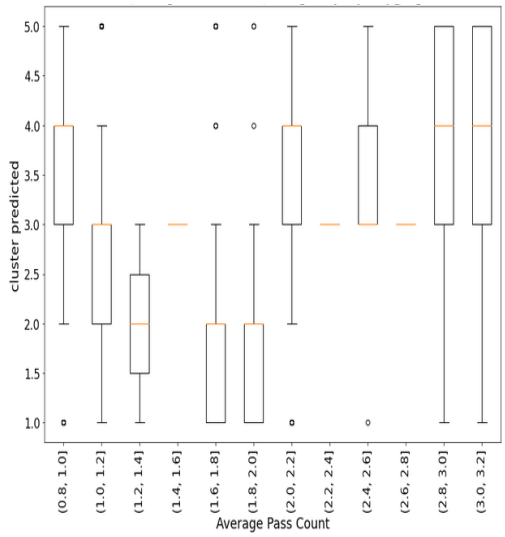
Figure 6.8: Box and whisker plot illustrating the results of clustering using Gaussian Mixture Model (GMM), utilizing the first principle component from PCA space as a feature, on Shawnee project data with different cluster ordering attributes.



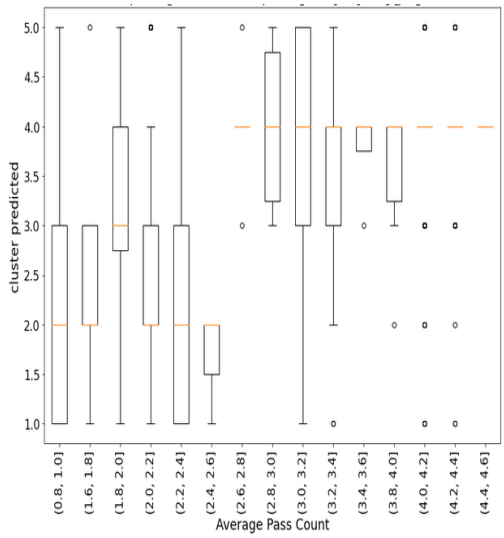
(a) Train data-set of Lathrop using total_power as ordering attribute



(b) Test data-set of Lathrop using total_power as ordering attribute

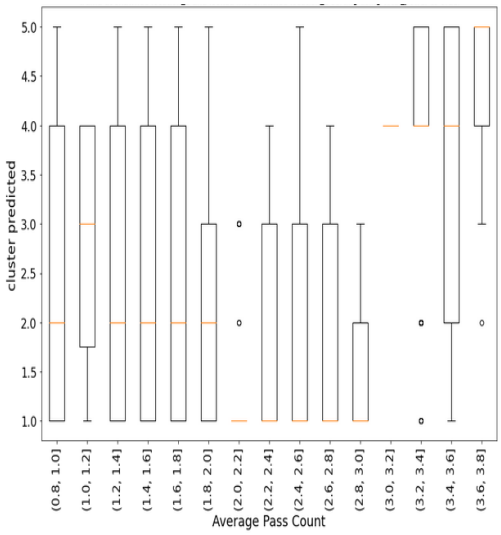


(c) Train data-set of Lathrop using APC as ordering attribute

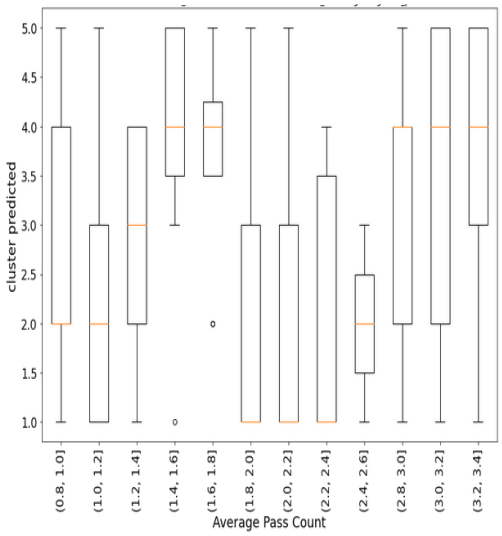


(d) Test data-set of Lathrop using APC as ordering attribute

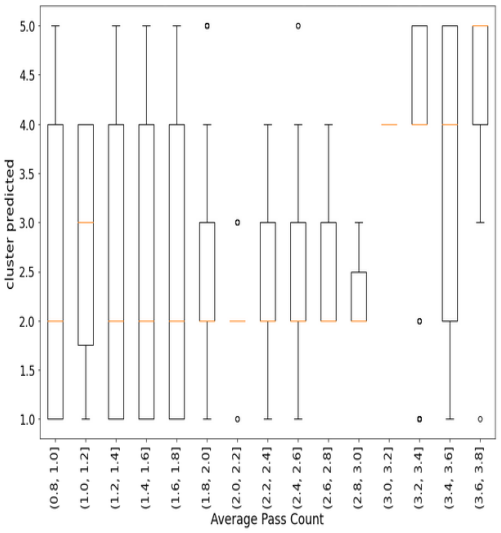
Figure 6.9: Box and whisker plot illustrating the results of clustering using Gaussian Mixture Model (GMM), utilizing the first two principle component from PCA space as feature, on Lathrop project data with different cluster ordering attributes.



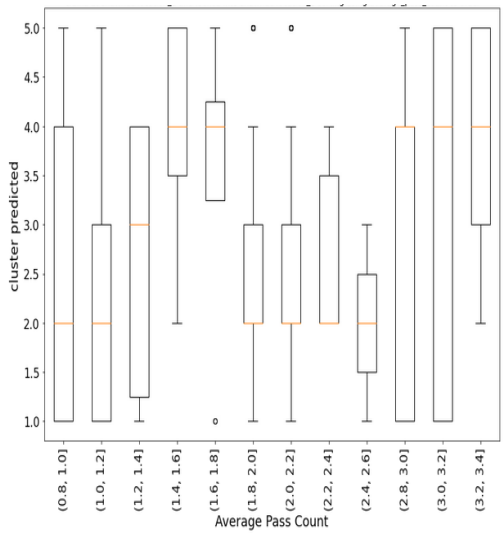
(a) Train data-set of Shawnee using total_power as ordering attribute



(b) Test data-set of Shawnee using total_power as ordering attribute



(c) Train data-set of Shawnee using APC as ordering attribute



(d) Test data-set of Shawnee using APC as ordering attribute

Figure 6.10: Box and whisker plot illustrating the results of clustering using Gaussian Mixture Model (GMM), utilizing the first two principle component from PCA space as feature, on Shawnee project data with different cluster ordering attributes.

From Figure 6.5 and Figure 6.6, we can see that while using total_power as the cluster ordering attribute, there are huge amounts of low cluster predictions in high APC even in the control training stretch. If this is true, then the system had not obtained compaction enough at these locations but, it being a controlled stretch where the final level of compaction was good enough to proceed to conduct the construction, that is not the case. Then when we tried using APC as our cluster ordering attribute, we see better distribution of high clusters in the high APC and low clusters in the low APC. This trend is not applicable to the training data but is the same in the test data in both Lathrop and Shawnee projects. This shows us that using APC as the cluster ordering parameter is the optimum option. We tried using both total_power and APC as cluster ordering attributes while clustering with GMM with a single principle component and two principle components in Figure 6.7, Figure 6.8, Figure 6.9 and Figure 6.10 respectively. Here too, we see a high average cluster value in high APC and also a smaller standard deviation in the final passes. So, from the visualization itself, we can say that using APC as the ordering attribute of the cluster is better than using total_power.

For the quantitative evaluation, we calculated the f_mean, f_std, and one_lag_auto evaluations metrics described in Section 5.1. These parameters give us the measure of how well is the algorithm working in the final pass. For our data, the mechanism that gives higher f_mean and one_lag_auto along with low f_std is the optimum mechanism.

Examining the data presented in Table 6.6, we observe that when "total_power" is used as the cluster ordering attribute for the Shawnee project and applying the

Table 6.6: Comparison of evaluation metrics when using total_power and APC as cluster ordering attributes

Clustering mechanism		Shawnee				Lathrop			
		total_power as ordering attribute		APC as ordering attribute		total_power as ordering attribute		APC as ordering attribute	
		train	valid	train	valid	train	valid	train	valid
KMC7	f_mean	2.66	2.37	2.66	2.44	2.94	1	2.91	2.63
	f_std	1.41	1.5	1.41	1.5	1.24	1.62	1.23	0.87
	one_lag_auto	0.34	0.37	0.34	0.3	0.25	0.19	0.23	0.22
GMM1	f_mean	2.87	2.12	2.87	2.47	3.07	1.55	3.17	3.19
	f_std	1.26	1.05	1.26	1.4	0.99	0.75	0.99	0.68
	one_lag_auto	0.35	0.23	0.35	0.48	0.23	0.49	0.21	0.12
GMM2	f_mean	2.58	2.37	2.58	2.16	2.91	1.57	2.89	2.83
	f_std	1.53	1.39	1.53	1.63	1.1	0.93	1.1	0.65
	one_lag_auto	0.27	0.4	0.27	0.31	0.42	0.39	0.44	0.23

KMC7 clustering mechanism, the average cluster values in the final pass are 2.66 for the training set and 2.37 for the test set. When substituting "APC" as the ordering attribute, the corresponding values become 2.66 for training and 2.44 for testing, yielding similar outcomes. The f_mean results remain consistent in this project across GMM1 and GMM2 implementations, and a similar trend is evident for f_std and one_lag_auto metrics. Hence, for the Shawnee project, the choice between "APC" and "total_power" as the cluster ordering attribute appears to yield comparable results.

However, a different pattern emerges when considering the Lathrop project. For instance, when using KMC7, GMM1, and GMM2 with "total_power" as the ordering attribute, the f_mean values in the train set are 2.94, 3.07, and 2.91 respectively, and in the test set, they are 1, 1.55, and 1.57. Notably, the f_mean value in the test set is significantly lower than that in the train set, indicating inconsistent prediction characteristics. Conversely, when "APC" is employed for cluster ordering, the f_mean values in the train set are 2.91, 3.17, and 2.89, and

Table 6.7: Metrics obtained when using features from 1 time_step and 2 time_step for KMC7, GMM1 and GMM2. Cluster ordering is done here using APC

Clustering mechanism		Shawnee				Lathrop			
		Using 128ms time_step		Averaginig 2X 128ms time_step		Using 128ms time_step		Averaginig 2X 128ms time_step	
		train	valid	train	valid	train	valid	train	valid
KMC7	f_mean	2.66	2.44	2.95	2.49	2.94	2.63	2.77	2.78
	f_std	1.41	1.5	1.18	1.39	1.24	0.87	1.27	0.97
	one_lag_auto	0.34	0.3	0.48	0.62	0.25	0.22	0.47	0.33
GMM1	f_mean	2.87	2.47	2.91	2.35	3.07	3.19	3.1	3.16
	f_std	1.26	1.4	1.32	1.54	0.99	0.68	1.09	0.65
	one_lag_auto	0.35	0.48	0.62	0.74	0.23	0.12	0.59	0.49
GMM2	f_mean	2.58	2.16	2.99	2.5	2.91	2.83	3.23	2.95
	f_std	1.53	1.63	1.36	1.6	1.1	0.65	0.95	0.46
	one_lag_auto	0.27	0.31	0.43	0.64	0.42	0.23	0.67	0.5

in the test set, they are 2.63, 3.19, and 2.83 for KMC7, GMM1, and GMM2 respectively. In this case, the f_std is also lower when using "APC," and the prediction consistency between the test and train sets improves. These findings underscore the superiority of "APC" as a cluster ordering attribute, resulting in the same or better performance in terms of f_mean and f_std.

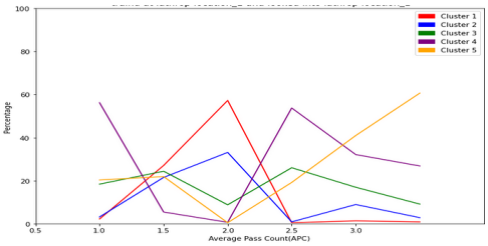
However, it's worth noting that when "APC" is utilized as the ordering parameter, the one_lag_auto metric, which gauges the smoothness of cluster predictions, experiences a slight decline. To investigate whether this change is influenced by noise within the system, a comparison was made between using a single 128ms time_step and taking the average of the present and last time_step (256ms). The resulting metrics are detailed in Table 6.7.

When we were using features from the given time step to cluster, we obtained one_lag_auto of 0.34,0.35 in the train set and 0.3,0.48 and 0.31 in the test set

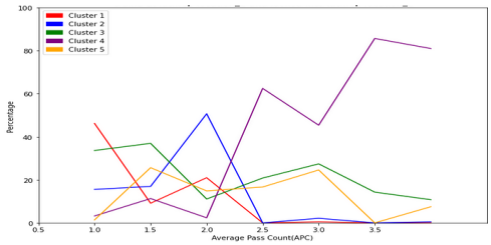
at Shawnee for KMC7, GMM1, and GMM2 respectively. But, when we averaged the features from current time_step to the previous time_step, the values of one_lag_auto were 0.48, 0.62, and 0.43 in test data and 0.62, 0.74, and 0.64 in test data at Shawnee for KMC7, GMM1, and GMM2. This shows that we can increase the smoothness of the final pass prediction using information from consecutive time_step. But the best part is that the f_mean and f_std are similar to using features of given time_step only. This result indicates that we can use two time_step and get similar predictions with better smoothness in comparison to using only one. Also, in the future, we may experiment with using 256ms as time_step and observe the results.

The best f_mean at Lathrop for the train is 3.1 and the test is 3.16, both when using GMM1 with APC as cluster ordering and average of two time_step features to cluster. But the f_std and one_lag_auto are best when using GMM2 for the same cases. And. looking into the Shawnee, it is observed that the best values of the f_mean and f_std are best while using GMM2 but one_lag_auto is best for GMM1. So, we can clearly state that using KMC7 is not the best option, and it is not much difference between GMM1 and GMM2.

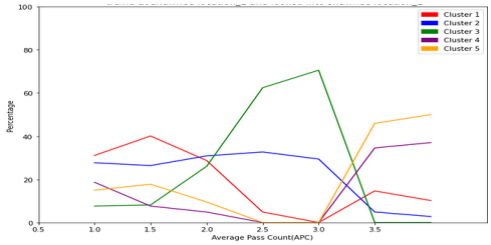
The primary goal behind clustering and ordering of clusters is to establish a mechanism that accurately reflects the compaction level trend. Our foundational understanding is that regions with high APC values should correspond to high compaction, moderate APC values to medium or low compaction, and low APC values to low compaction. However, a closer inspection of the cluster distribution depicted in Figure 6.8 reveals that this trend is not consistently adhered to,



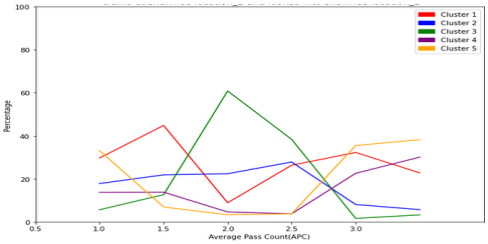
(a) Distribution of percentage of clusters in training data from Lathrop project



(b) Distribution of percentage of clusters in test data from Lathrop project



(c) Distribution of percentage of clusters in training data from Shawnee project

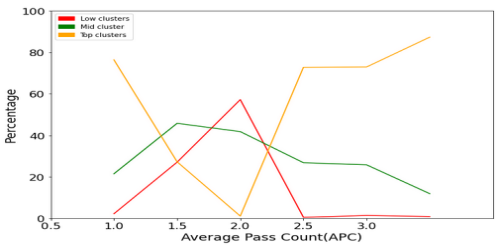


(d) Distribution of percentage of clusters in test data from Shawnee project

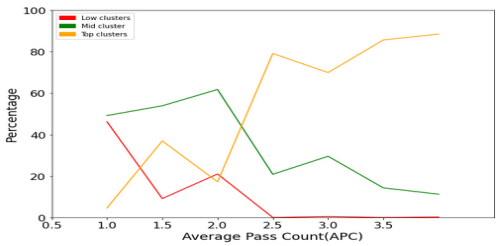
Figure 6.11: Plots showing the distribution of clusters in various locations

particularly in areas with lower pass counts. Notably, there is a substantial presence of high clusters even in regions with low APC values. Another intriguing observation is that the percentage of high clusters is minimal when APC values are low, a trend consistent across both the Shawnee and Lathrop project data.

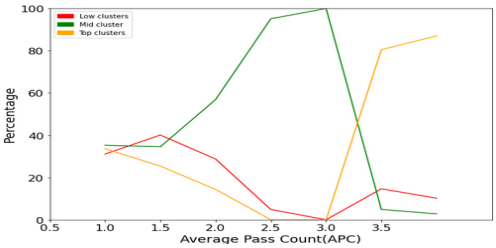
At first glance, these results might seem to indicate that the developed system struggles to align with the known information about HMA construction that with increase in APC, there should be increase in the level of compaction density. Yet, a deeper analysis is warranted to comprehend the underlying dynamics. When APC is approximately 1, signifying substantial uncompressed HMA, the roller's movement leads to material shifting rather than compaction. Consequently, the vibration data obtained is significantly influenced by the underlying pavement substrate. In such regions, the data appears scattered, making it challenging to accurately predict asphalt compaction levels. Additionally, considering that the



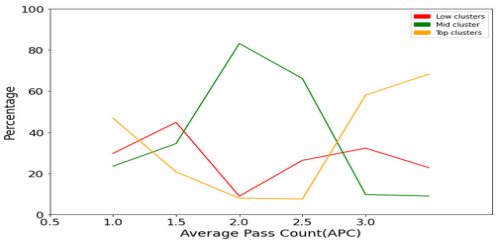
(a) Distribution of percentage of clusters in training data from Lathrop project



(b) Distribution of percentage of clusters in test data from Lathrop project



(c) Distribution of percentage of clusters in training data from Shawnee project



(d) Distribution of percentage of clusters in test data from Shawnee project

Figure 6.12: Plots of levels of cluster distribution in train and validation data

roller's width is around 5-6 feet and it typically compacts within a lane width of about 10 feet, initial passes focus on compressing the lane's edges – whether they are shoulders, curves, or adjacent lanes. As a result, areas with APC values near 1 are inherently susceptible to substantial noise.

Given these insights, it becomes reasonable to expect a notable presence of high clusters in such regions. However, as the roller progresses through its final passes, consistent high compaction is achieved, resulting in a predominance of high clusters, as depicted in Figure 6.11. This visual representation reinforces the trend that high compaction is indeed achieved in the later stages. If we make the cluster prediction into three level, high clusters(4 and 5), mid_clusters(2 and 3) and low clusters(cluster 1) as in Figure 6.12, the trend is more prominent to the expectation. We can observe that with increase of the APC, there is higher percentage of high clusters and low percentage of low clusters. The percentage of

high and low clusters is reversed in low APC, other than in the error region of APC<1.5.

In essence, while the observed cluster distribution may appear initially counterintuitive, a comprehensive understanding of the underlying factors, such as the roller's initial lane-edge compression and the impact of the underlying substrate, reveals that the system's behavior aligns with the complexity of the compaction process. This deeper analysis underscores the system's capability to capture the nuanced behavior of compaction and provides context for the observed cluster patterns.

In Figure 6.11(a), we observe the distribution of clusters in the training data from Lathrop, and in Figure 6.11(b), we see the test data from the same project. In the region with APC=1-1.5, all clusters are present, with cluster 4, a high cluster, being the most prominent one in the training set. However, as we move to the region of APC=1.5-2, we notice all the high clusters vanish, and the low clusters become prominent in this region. Then, in the region of APC=2-2.5, there is a slight increase in the percentage of high clusters, and the number of low clusters decreases considerably. Finally, for the region with APC>2.5, high clusters (4 and 5) take over.

Figure6.11(c) and Figure6.11(d) show the distribution of clusters in various APC values in the train and test data from the Shawnee project. The trend is similar here as well. All these trends are generated from the results of implementing Algorithm 4. Therefore, we can conclude that by implementing Algorithm 4, we can effectively capture the trend of changes in the compactness of HMA pavement

Algorithm 4 Optimum Clustering Mechanism: PCA Feature Extraction, GMM Clustering, and Cluster Ordering

Input: Dataset X
Fix_hyperparameter: Number of clusters $K \leftarrow 5$
Output: Cluster centers C, Cluster assignments S
PCA Feature Extraction:
Compute the mean: Calculate the mean vector $\bar{x}$ of the data points in X
Center the data: Subtract the mean vector $\bar{x}$ from each data point to get the centered data matrix $\hat{X}$
Compute the covariance matrix: Compute the covariance matrix C of the centered data matrix $\hat{X}$ as $C = \frac{1}{n}\hat{X}^T\hat{X}$, where n is the number of data points
Compute the eigenvectors and eigenvalues: Compute the eigenvectors and eigenvalues of the covariance matrix C. Sort them in decreasing order and select the first two eigenvectors, denoted by v_1 and v_2.
GMM Clustering with First Two Principal Components:
Project data onto the first two principal components: Project the data points in X onto the plane spanned by the first two principal components v_1 and v_2 to obtain the 2-dimensional feature vectors.
Initialize GMM: Randomly initialize K cluster centers C.
repeat
Expectation Step: For each data point, compute the probability of belonging to each cluster using the current cluster centers.
Maximization Step: Update the cluster centers as the weighted mean of the data points, where the weights are the probabilities obtained in the expectation step.
until Convergence
Assign data points to clusters: Assign each data point to the cluster with the highest probability obtained in the expectation step.
Cluster Ordering based on Average APC Attribute:
Compute the Average APC for each cluster: Calculate the average APC (Average Pass Count) attribute for each cluster using the data points assigned to that cluster.
Order the clusters: Order the clusters based on their average APC values, with the cluster having the highest average APC considered as the most compacted with cluster number K and the cluster with the lowest average APC considered as the least compacted with cluster number 1.
Return: Final cluster centers C and cluster assignments S in ordered fashion

during construction through the cluster numbers.

6.2 Results in density prediction

Table 6.8: Prediction results in the various core data collected

Core_number	Latitude	Longitude	Density	Predicted_cluster	Density(linear)	Density(sigmoid)
Shawnee_1	35.68284924	-96.66262486	95.2	5	96.5	95.72
Shawnee_2	35.67784402	-96.66244896	94.6	4	94.875	94.75
Shawnee_3	35.67193842	-96.66235509	96.2	5	96.5	95.72
Shawnee_4	35.67062983	-96.6623161	94.6	4	94.875	94.75
Shawnee_5	35.67062625	-96.66230751	94.8	4	94.875	94.75
Lathrop_1	37.81907004	-121.2891194	95.2	5	95.5	94.84
Lathrop_2	37.81907203	-121.2885476	95	5	95.5	94.84
Lathrop_3	37.81905966	-121.2879115	94	4	94.125	94.02
Lathrop_4	37.81903288	-121.2904883	93.4	4	94.125	94.02
Lathrop_5	37.81903881	-121.2911619	93.8	4	94.125	94.02
Lathrop_6	37.81902592	-121.2915524	92.7	3	92.75	92.75

The subsequent crucial phase involves predicting the density values based on the cluster numbers that represent varying levels of compactness. Initially, we explored employing linear regression to establish a relationship between the cluster numbers and density values. Subsequently, we also experimented with utilizing sigmoidal fittings to capture the underlying patterns more accurately.

In the Shawnee project, the construction involved distinct Breakdown and Intermediate rollers, with our system connected to the Intermediate roller. Consequently, our data collection commenced after a certain degree of pavement compression, subsequent to the Breakdown phase. The minimum density observed through our system aligned with measurements from an NDG density

measurement device, registering at 91%. The project's maximum achievable density was recorded as 96.5%. Similarly, in the Lathrop project, the determined min_density was 90%, while the max_density reached 95.5%.

To predict the final cluster number at a core location, data from the nearest final pass time_step was selected, as it was most relevant to the given core location for cluster prediction. Subsequently, both linear and sigmoidal fitting methodologies were applied. The outcomes, as detailed in Table 6.8, were analyzed.

The predicted density values, as showcased in Table 6.8, led to calculated Root Mean Squared Error (RMSE) values of 0.266 for linear fitting and 0.103 for sigmoidal fitting. Additionally, the computed R_squared values were 0.69 for linear fitting and 0.88 for sigmoidal fitting. These metrics provide insight into the predictive accuracy of each approach.

Based on these results, we can deduce that sigmoidal fitting outperforms linear fitting for this specific application. The lower RMSE and higher R_squared values associated with sigmoidal fitting indicate its superior suitability in capturing the complex relationship between cluster numbers and density values in this context.

7. CONCLUSION AND FUTURE WORK

After conducting various experiments and analyzing the results of this application-based research, we have identified an algorithm capable of monitoring the compaction level in HMA pavement during construction. The algorithm is designed specifically for vibratory roller compactors, as the vibration patterns contain crucial information about the pavement's compactness during construction.

Initially, a retrofit system was devised to attach to impulse vibratory roller compactors, enabling the collection of vibration, temperature, and location data. Vibration and temperature information were collected at 1000Hz, while GPS location data was collected at 1Hz. Utilizing FFT, we converted the time-domain vibration data into the frequency domain, observing significant changes around the fundamental frequency and its harmonics. Focusing on the readings in those regions, we termed them as the power of the given frequency band.

Later, we discovered that the power content in the fundamental frequency and its harmonics can serve as essential characteristics to determine the level of compaction density. Throughout our research, we primarily focused on a time_step of 128ms and formulated various attributes for that specific time interval using vibration and location data. Additionally, we attempted to reduce noise by averaging the last two 128ms time_steps, yet still predicting the clusters every 128ms.

To further enhance our understanding, we need to conduct a thorough

investigation to determine the optimum duration of the time_step. This investigation should strike a balance between having a small enough time_step to detect changes in density levels and having a long enough time_step to mitigate the impact of noise effectively. Such an analysis will aid in optimizing the system's performance and accuracy.

Determining the compaction density ideally should be approached as a regression problem. However, due to the limited availability of ground truth data, we need to proceed using unsupervised learning methods. Among various unsupervised techniques, we have considered clustering and ordering the resulting clusters to determine the different levels of compaction densities.

To find the optimum combination, we explored various feature extractions and different clustering algorithms with varying cluster numbers. Based on the Silhouette Score, DB Index, and CH Index analysis, we identified KMC7, GMM1, and GMM2 as the optimal clustering algorithms. Moving forward, we aimed to find the best attribute for ordering the clusters based on the level of compaction. It was determined that APC (Average Pass Count) is the most suitable attribute for executing the ordering.

The main objective of clustering and ordering the clusters is to develop a mechanism that accurately follows the trend of compaction levels. The most optimal resulting algorithm from all the experiments conducted is identified as Algorithm 4. Figure 6.11 demonstrates proper trend detection of compaction levels using this algorithm. As an unsupervised learning problem, we employ domain knowledge to assess the results of the clustering mechanism. Based on the

results, we can confidently assert that this system is effective in learning the various levels of compaction density. However, obtaining exact density values remains unattainable due to the limitations in ground truth data and the inherent nature of the unsupervised learning approach.

Furthermore, we attempted to determine the fitting mechanism for converting the cluster numbers into density predictions. Sigmoidal fitting was observed to be the best option, although further research is necessary to gain a better understanding. Due to limited ground truth data (11 cores from two projects), statistical significance is challenging to establish. More core data needs to be collected from real construction sites to facilitate comprehensive analysis. Additionally, a major limitation is that all the available ground truth data is from the final fully compacted regions, leaving us without data to evaluate the system's performance in low and intermediate compaction during non-final passes. Data from those situations is essential to make robust claims about the system.

8. APPENDIX

Table 8.1: Specifications of the two vibratory rollers used (referenced from the Caterpillar Tandem Vibratory Roller specifications)

Property	CB-7	CB-10
Operating Frequency Range (Hz)	42-63.3	40-63.3
Vibratory Amplitude(mm)	0.25-0.65	0.18-0.84
Drum width(mm)	1500	1700
Drum diameter(mm)	1108	1198
Operating weight(kg)	7990	9500

Table 8.2: Specifications of the 2240-010 Titanium Hermetic Accelerometer used in RICA (ref: *https* : *//www.silicondesigns.com*/2240 − 2480)

Parameter	**Min**	**Typ**	**Max**	**Units**
Bias Calibration Error (%)		0.25	0.6	± % of span
Bias Calibration Error (mV)		25	60	± mV
Scale Factor Calibration Error (1)		0.5	1.25	± %
Non-Linearity (-90 to +90% of span) (1)		0.15	0.5	± % of span
Bias Temperature Shift (Coefficient)	-100	0	+100	(PPM of span)/°C
Scale Factor Temperature Shift (Coefficient)	-150	0	+50	PPM/°C
Cross Axis Sensitivity		2	3	± %
Power Supply Rejection Ratio	50	>65		dB
Output Impedance		1		Ω
Output Common-Mode Voltage		2.5		VDC
Operating Voltage	8		32	VDC
Operating Current 2240 / 2480	6 / 19	7 / 23	12 / 27	mA DC
Operating Temperature	-55		+125	°C
Mass 2240 / 2480		15 / 25		grams

Table 8.3: Specifications of the CI Series Compact IR Temperature Sensor from Raytek used in RICA (ref: *https : //www.raytek − direct.com/pdfs/cache/www.raytek − direct.com/cm − series/datasheet/cm − series − datasheet.pdf*)

Parameter	**Range**
Temperature Range	-20ºC to 500ºC (-4ºF to 932ºF)
Accuracy (mV)	± 1.5% of reading or ± 2ºC, whichever is greater 1,2
Accuracy (t/c)	± 1.5% of reading or ± 4ºC, whichever is greater1,2,3
Spectral Response	8 to 14 microns
System Repeatability	± 0.5% of reading or ± 2ºC, whichever is greater
Temperature Resolution	0.1°C
Response Time (95%)	150 mSec
Emissivity	0.10 to 1.10 (adjustable)
Transmissivity	0.10 to 1.00 (adjustable)
Signal Processing	Peak hold, valley hold, variable averaging
Bore-Sight tolerance	3º @ focal point

Table 8.4: Specifications of the GNSS Receiver for GPS location (ref:*https : //www.navtechgps.com/wp − content/uploads/Vector − VS*1000_D*S.pdf*)

Parameter	**GNSS Receiver Specifications**
Receiver Type:	Vector GNSS RTK Receiver
Signals Received:	GPS, GLONASS, BeiDou, Galileo, & Atlas 3
Channels:	1059
GPS Sensitivity:	-142 dBm
SBAS Tracking:	2-channel, parallel tracking
Update Rate:	10 Hz standard, 20 Hz optional
Timing (1PPS) Accuracy:	20 ns
Rate of Turn:	100°/s maximum
Cold Start:	60 s (no almanac or RTC)
Warm Start:	30 s typical (almanac and RTC)
Hot Start:	10 s typical (almanac, RTC, and position)
Heading Fix:	10 s typical (valid position)
Antenna Input Impedance:	50 Ω
Maximum Speed:	1,850 kph (999 kts)
Maximum Altitude:	18,000 m (59,055 ft)
Differential Options:	SBAS, Atlas (L-band), RTK

REFERENCES

1. Gedik, A. A review on the evaluation of the potential utilization of construction and demolition waste in hot mix asphalt pavements. *Resources, Conservation and Recycling* **161,** 104956 (2020).

2. Yoon, S., Hastak, M. & Lee, J. Suitability of intelligent compaction for asphalt pavement quality control and quality assurance. *Journal of Construction Engineering and Management* **144,** 04018006 (2018).

3. Roberts, F. L., Kandhal, P. S., Brown, E. R., Lee, D.-Y. & Kennedy, T. W. Hot mix asphalt materials, mixture design and construction (1991).

4. Corps, U. A. *Hot-Mix Asphalt Paving Handbook* https://rosap.ntl.bts.gov/view/dot/42051/dot_42051_DS1.pdf (USA: US Army Corps of Engineers, 2000).

5. Scherocman, J. A. & Martenson, E. D. *Placement of asphalt concrete mixtures* in *Placement and Compaction of Asphalt Mixtures: A Symposium* (1984), 12–17.

6. Scherocman, J. Guidelines for compacting asphalt concrete pavement. *Better Roads* **54,** 12–17 (1984).

7. Brown, E. R. *Experiences of Corps of Engineers in compaction of hot asphalt mixtures* (ASTM International, 1984).

8. Hughes, C. S. *Compaction of asphalt pavement* **152** (1989).

9. Kassem, E., Scullion, T., Masad, E. & Chowdhury, A. Comprehensive evaluation of compaction of asphalt pavements and a practical approach for density predictions. *Transportation research record* **2268,** 98–107 (2012).

10. Kennedy, T. W., McGennis, R. & Roberts, F. *Effects of compaction temperature and effort on the engineering properties of asphalt concrete mixtures* 48–66 (American Society for TEsting and MAterials, 1984).

11. Pell, P. & Taylor, I. *Asphaltic road materials in fatigue* in *Association of Asphalt Paving Technologists Proc* **38** (1969), 371–464.

12. Epps, J. A. *et al. Influence of mixture variables on the flexural fatigue properties of asphalt concrete* in *Association of Asphalt Paving Technologists Proc* **38** (1969), 423–464.

13. Linden, R. N., Mahoney, J. P. & Jackson, N. C. Effect of compaction on asphalt concrete performance. *Transportation research record.* https://onlinepubs.trb.org/Onlinepubs/trr/1989/1217/1217-003.pdf (1989).

14. Finn, F. N., Nair, K. & Hilliard, J. *Minimizing premature cracking of asphalt concrete pavements* https://onlinepubs.trb.org/Onlinepubs/nchrp/nchrp_rpt_195.pdf (National Cooperative Highway Research Program Project 9-4., 1973).

15. McLeod, N. W. Influence of viscosity of asphalt-cements on compaction of paving mixtures in the field and discussion. *Highway Research Record* (1967).

16. Kandhal, P. S. & Koehler, W. C. Pennsylvania's experience in the compaction of asphalt pavements. *ASTM Special Technical Publication* **829,** 93–106 (1984).

17. Cooley, L. A., Prowell, B. D. & Brown, E. Issues pertaining to the permeability characteristics of coarsegraded superpave mixes. *NCAT Report,* 2–6. https://rosap.ntl.bts.gov/view/dot/34128/dot_34128_DS1.pdf (2002).

18. Butts, N. & Ksaibati, K. *Asphalt pavement quality control/quality assurance programs in the united states* in *Annual Meeting of the Transportation Research Board (Washington DC, United States)* (Jan. 2003).

19. Scherocman, J. A. & Walker, D. Factors affecting asphalt compaction. *Asphalt* **23.** https://onlinepubs.trb.org/onlinepubs/circulars/ec105.pdf (2008).

20. Mitchell, T. Density monitoring on asphalt pavement. *Better Roads* **54** (1984).

21. Kumar, V., Coleri, E., Obaid, I. A., *et al. Constructing High Performance Asphalt Pavements by Improving in Place Pavement Density* tech. rep. (Oregon. Dept. of Transportation. Research Section, 2021).

22. Yiqiu, T., Haipeng, W., Shaojun, M. & Huining, X. Quality control of asphalt pavement compaction using fibre Bragg grating sensing technology. *Construction and Building Materials* **54,** 53–59 (2014).

23. Steyn, W. & Sadzik, E. Application of the portable pavement seismic analyser (PSPA) for pavement analysis (2007).

24. Diamanti, N., Redman, J. D. & Annan, A. P. *A GPR-based sensor to measure asphalt pavement density* in *2018 17th International Conference on Ground Penetrating Radar (GPR)* (2018), 1–6.

25. Timm, A. *Comparison of nuclear and non-nuclear density gauges for determining in-place density of hot mix asphalt* PhD thesis (Washington State University, 2012).

26. Liao, J. R. & Powell, J. D. Intelligent Compaction: A New Approach to Asphalt Pavement Construction and Evaluation. *Transportation Research Board's 73rd Annual Meeting* (1994).

27. Mooney, M. A. & Rinehart, R. V. Field monitoring of roller vibration during compaction of subgrade soil. *Journal of Geotechnical and Geoenvironmental Engineering* **133,** 257–265 (2007).

28. Torres, A. & Arasteh, M. *Intelligent compaction measurement values (ICMV)—A road map. US Department of Transportation, Federal Highway Administration, Technical Brief* tech. rep. (FHWA-HIF-17-046, Washington, DC, 2017).

29. Chang, G. K., Xu, Q., Rutledge, J. L., Garber, S. I., *et al. A study on intelligent compacting and in-place asphalt density* tech. rep. (United States. Federal Highway Administration. Office of Pavement Technology, 2014). https://www.fhwa.dot.gov/construction/ictssc/pubs/hif14017.pdf.

30. Sivagnanasuntharam, S., Sounthararajah, A., Ghorbani, J., Bodin, D. & Kodikara, J. A state-of-the-art review of compaction control test methods and intelligent compaction technology for asphalt pavements. *Road Materials and Pavement Design* **24,** 1–30 (2023).

31. Sivagnanasuntharam, S., Sounthararajah, A. & Kodikara, J. In-situ spot test measurements and ICMVs for asphalt pavement: lack of correlations and the

effect of underlying support. *International Journal of Pavement Engineering* **24,** 2198770 (2023).

32. Xu, Q., Chang, G. K. & Gallivan, V. L. Development of a systematic method for intelligent compaction data analysis and management. *Construction and Building Materials* **37,** 470–480 (2012).

33. Zhang, W. *et al.* Investigation of the correlations between the field pavement in-place density and the intelligent compaction measure value (ICMV) of asphalt layers. *Construction and Building Materials* **292,** 123439 (2021).

34. Commuri, S., Mai, A. T. & Zaman, M. Neural network-based intelligent compaction analyzer for estimating compaction quality of hot asphalt mixes. *IFAC Proceedings Volumes* **41,** 2224–2229 (2008).

35. Commuri, S. & Zaman, M. A novel neural network-based asphalt compaction analyzer. *International Journal of Pavement Engineering* **9,** 177–188 (2008).

36. Commuri, S., Mai, A. T. & Zaman, M. Neural network–based intelligent compaction analyzer for estimating compaction quality of hot asphalt mixes. *Journal of Construction Engineering and Management* **137,** 634–644 (2011).

37. Commuri, S., Zaman, M., Barman, M., Nazari, M., Imran, S., *et al. Evaluation of performance of asphalt pavements constructed using intelligent compaction techniques.* tech. rep. (Oklahoma. Dept. of Transportation. Materials and Research Division, 2014).

38. Plati, C., Georgiou, P. & Loizos, A. Influence of different roller compaction modes on asphalt mix performance. *International journal of pavement engineering* **17,** 64–70 (2016).

39. Zheng, S., Dai, Q. & Lin, S. Response characteristics analysis of vibrating drum in vibratory roller considering the asymmetrical hysteresis of materials. *Journal of Machine Design* **33,** 87–91 (2016).

40. Wan, Y. & Jia, J. Nonlinear dynamics of asphalt–screed interaction during compaction: Application to improving paving density. *Construction and Building Materials* **202,** 363–373 (2019).

41. Ingrassia, L. P., Lu, X., Ferrotti, G. & Canestrari, F. Chemical, morphological and rheological characterization of bitumen partially replaced with wood bio-oil: Towards more sustainable materials in road pavements. *Journal of Traffic and Transportation Engineering (English Edition)* **7,** 192–204 (2020).

42. Gao, J. *et al.* High-temperature rheological behavior and fatigue performance of lignin modified asphalt binder. *Construction and Building Materials* **230,** 117063 (2020).

43. Nivitha, M. & Murali Krishnan, J. Rheological characterisation of unmodified and modified bitumen in the 90–200 C temperature regime. *Road Materials and Pavement Design* **21,** 1341–1358 (2020).

44. Halle, M., Rukavina, T. & Domitrovic, J. *Influence of temperature on asphalt stiffness modulus* in *Proceedings of the 5th Eurasphalt Eurobitume Congress, Istanbul, Turkey* (2012), 13–15.

45. Masad, E., Scarpas, A., Alipour, A., Rajagopal, K. R. & Kasbergen, C. Finite element modelling of field compaction of hot mix asphalt. Part I: Theory. *International Journal of Pavement Engineering* **17,** 13–23 (2016).

46. Masad, E. *et al.* Finite element modelling of field compaction of hot mix asphalt. Part II: Applications. *International Journal of Pavement Engineering* **17,** 24–38 (2016).

47. Beainy, F., Commuri, S., Zaman, M. & Syed, I. Viscoelastic-plastic model of asphalt-roller interaction. *International Journal of Geomechanics* **13,** 581–594 (2013).

48. Burgers, J. Mechanical considerations-model systems-phenomenological theories of relaxation and viscosity, First report on viscosity and plasticity. *Nordemann, New York* (1939).

49. Beainy, F., Commuri, S. & Zaman, M. Dynamical response of vibratory rollers during the compaction of asphalt pavements. *Journal of Engineering Mechanics* **140,** 04014039 (2014).

50. Rinehart, R. V. & Mooney, M. A. Instrumentation of a roller compactor to monitor vibration behavior during earthwork compaction. *Automation in Construction* **17,** 144–150 (2008).

51. Ingle, V. K. & Proakis, J. G. *Digital signal processing using matlab: a problem solving companion* (Cengage Learning, 2016).

52. Downey, A. *Think DSP: digital signal processing in Python* (" O'Reilly Media, Inc.", 2016).

53. Hotelling, H. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology* **24,** 417 (1933).

54. Mika, S. *et al.* Kernel PCA and de-noising in feature spaces. *Advances in neural information processing systems* **11** (1998).

55. Balasubramanian, M. & Schwartz, E. L. The isomap algorithm and topological stability. *Science* **295,** 7–7 (2002).

56. Schölkopf, B., Smola, A. & Müller, K.-R. *Kernel principal component analysis* in *International conference on artificial neural networks* (1997), 583–588.

57. Roweis, S. T. & Saul, L. K. Nonlinear dimensionality reduction by locally linear embedding. *science* **290,** 2323–2326 (2000).

58. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9** (2008).

59. McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426* (2018).

60. Yang, M.-S., Lai, C.-Y. & Lin, C.-Y. A robust EM clustering algorithm for Gaussian mixture models. *Pattern Recognition* **45,** 3950–3961 (2012).

61. Ester, M., Kriegel, H.-P., Sander, J., Xu, X., *et al. A density-based algorithm for discovering clusters in large spatial databases with noise* in *kdd* **96** (1996), 226–231.