EVALUATION OF GASOLINE QUALITY IN AIZAWL AND CONTENT OF AROMATIC AND OXYGENATE COMPOUNDS

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

LALBIAKTLUANGA

MZU REGISTRATION NO.: 102 of 2015

Ph.D. REGISTRATION NO.: MZU/Ph.D./1141 of 24.04.2018

DEPARTMENT OF PHYSICS

SCHOOL OF PHYSICAL SCIENCES

FEBRUARY, 2025

EVALUATION OF GASOLINE QUALITY IN AIZAWL AND CONTENT OF AROMATIC AND OXYGENATE COMPOUNDS

By

Lalbiaktluanga

Department of Physics

Name of Supervisor: Dr. Hranghmingthanga

Submitted

In partial fulfillment of the requirement of the Degree of Doctor of Philosophy

in Physics of Mizoram University, Aizawl

Dr. Hranghmingthanga
Associate Professor
Department of Physics
Mizoram University
Tanhril - 796004

MIZORAM UNIVERSITY
DEPARTMENT OF PHYSICS
AIZAWL 796 004, MIZORAM
Phones: 0389-2330435, -2330522
FAX : 0389-2330522
Mobile : +919436779952
E-mail : hthanga@rediffmail.com

(A Central University Established by an Act of Parliament)

CERTIFICATE

This is to certify that the thesis entitled '*Evaluation of Gasoline Quality in Aizawl and Content of Aromatic and oxygenate compounds*' submitted by Lalbiaktluanga, for the degree of Doctor of Philosophy of the Mizoram University, Aizawl, embodies the record of original investigations carried out by him under my supervision. The thesis presented is worthy of being considered for the award of the Ph. D. degree. This work has not been submitted for any degree to any other University.

(Dr. HRANGHMINGTHANGA)

Supervisor

Declaration

Mizoram University

February, 2025

I, **Lalbiaktluanga**, hereby declare that the subject matter of this thesis is the record of work done by me, that the contents of this thesis did not form basis of the award of any previous degree to me or to the best of my knowledge to anybody else, and that the thesis has not been submitted by me for any research degree in any other University or Institute.

This is being submitted to the Mizoram University for the degree of Doctor of Philosophy in Physics.

(LALBIAKTLUANGA)

Candidate

(Prof. B. LALREMRUATA)

Head

(Dr. HRANGHMINGTHANGA)

Supervisor

Acknowledgment

First and foremost, I want to convey my sincere appreciation to my supervisor, Dr. Hranghmingthanga, Associate Professor in the Department of Physics at Mizoram University, for his respected guidance and oversight throughout my research. His steadfast support and insightful feedback have been crucial to the completion of this thesis.

I am deeply thankful to Prof. Diwakar Tiwari, Dean of the School of Physical Sciences, for his assistance in bringing my work to fruition. I also acknowledge the support received from Prof. B. Lalremruata, Head of the Department of Physics, and all the faculty members of the department at Mizoram University. My gratitude extends to the non-teaching staff, including Mr. Mala, Mrs. Luri, and Mrs. Hmai.

I am profoundly grateful to my seniors, Dr. J. Lalnunthari, Dr. Lalremruata Hlondo, and Dr. Rebecca Lalnuntluangi, for their inspiration and guidance. Their valuable advice, moral support, and encouraging words have been greatly appreciated.

I wish to express my heartfelt thanks to my friends Mrs. Vanzarmawii, Mr. Lalruatpuia Ralte, Mr. Lalmuanchhana, Mr. H. Lalthantluanga, Joseph, J. Lalramnghaka and Mr. Laldingngheta Ralte, for their companionship, support, and vibrant presence throughout this journey. Their support made the process significantly easier.

Additionally, I extend my thanks to my fellow research scholars in the Department of Physics at Mizoram University for their help and valuable advice, which was instrumental in completing my work.

I am eternally grateful to my parents and family for their encouragement and support throughout my studies. Their unwavering belief in me has kept my motivation and spirits high during this period. My thanks go to everyone who has been a part of this challenging yet rewarding journey.

Lastly, I thank the Almighty God for being with me through these years and granting me the health and strength needed to complete this work.

Dated: February, 2025

Department of Physics,

Mizoram University, Aizawl. (**LALBIAKTLUANGA**)

CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Diverse Products of Petroleum: From Refining to Fuel and Beyond

The word petroleum is derived from the Latin words "Petro" (stone) and "Oleum" (oil). It indicates stone oil (Haney., 1923). Petroleum, often referred to as crude oil, is a naturally occurring, flammable liquid found in rock formations in the Earth's crust. It is formed by the degradation of biological substances over millions of years (Ali., 2017). Petroleum is made up of hydrocarbons, which are organic substances that include hydrogen and carbon atoms. However, it is an extremely complex substance with no simple formula, containing primarily of carbon (93-97%) and hydrogen (10-14%) with small amounts of nitrogen (0.1-2%), oxygen (1-1.5%), and sulfur (0.5-6%) (Ali., 2017). Under typical conditions, it can exist as a gas, a liquid, or a solid. Hydrocarbons such as alkanes, cycloalkanes, and aromatic hydrocarbons are important components of petroleum. Crude oil's particular composition varies based on its source, and various areas produce crude oils with varying qualities. Crude oil is a mixture of Hydrocarbon molecules containing 1 to 60 carbon atoms (Gautam et al., 1998). So that the properties of hydrocarbons depend upon the number and arrangement of carbon and hydrogen atoms in their molecules. Unlike water, crude oil is not a single compound, it can be a mixture of large and small hydrocarbons. So that a wide variety of products can be formed from petroleum distillation. The composition and appearance of crude oil can be difference according to the region and location.

In the petroleum refining sector, crude oil distillation is an essential procedure that is the first stage in transforming crude oil into a variety of valuable products. Crude oil distillation is primarily used to divide the complex combination of hydrocarbons

found in crude oil into several fractions according to their boiling points (Behrenbruch et al., 2007). Each fraction obtained by distillation has distinct features and applications. However, crude oil distillation is essential for unlocking the value of crude oil by separating it into profitable fractions with a wide range of uses. This process is important because it meets energy demands, supports industrial activities, and provides necessary materials for many parts of modern life.

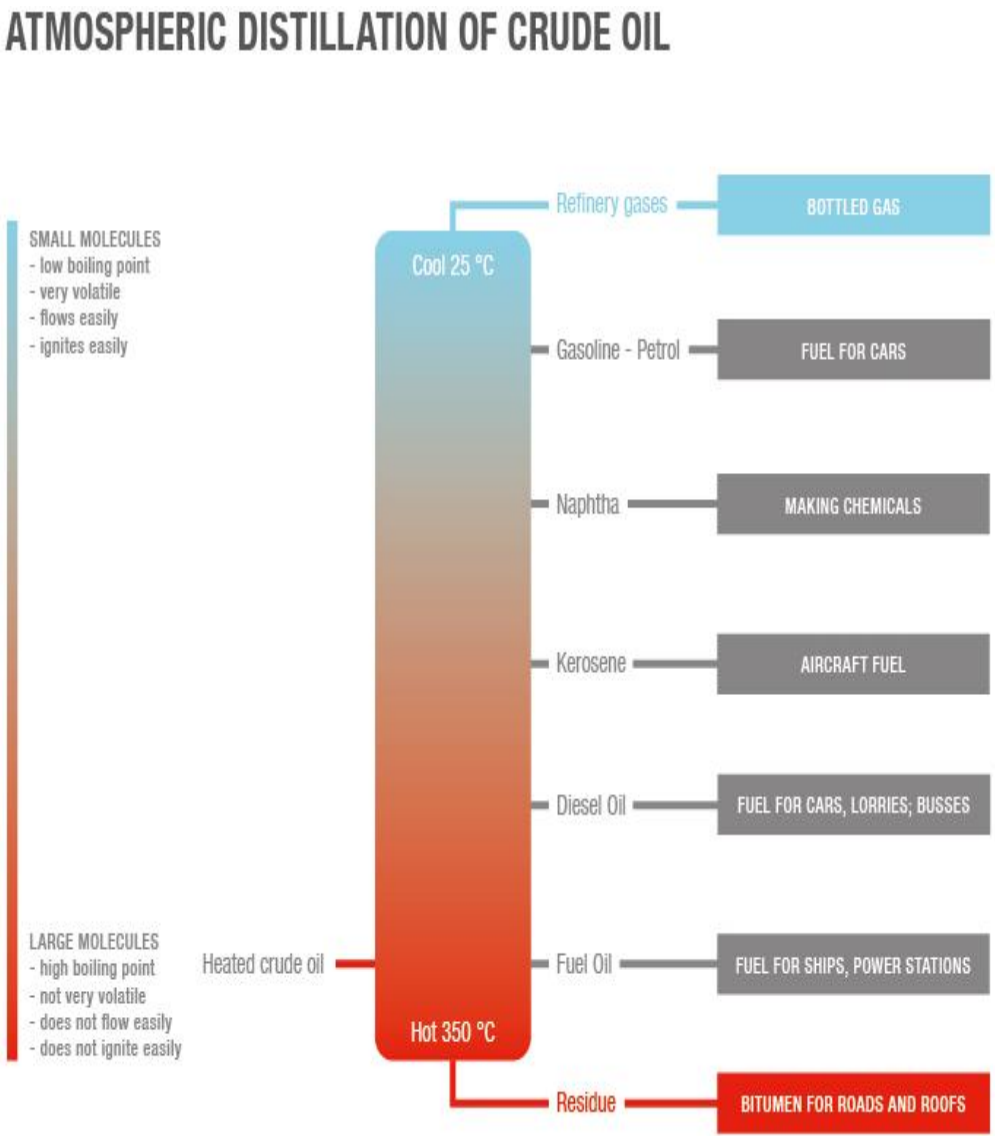

***Figure 1**: Atmospheric distillation of crude oil. (Anton-Paar. 2009)*

As can be seen from Fig. 1, a wide variety of product are formed from crude oil distillation process. Based on its distinct characteristics, each of these goods can be used for a different purpose. Petroleum products are essential components in many industries and provide basic human needs. Transportation fuels are the most widely used petroleum products in the world. These fuels are essential for powering automobiles and aircraft, which allows individuals and objects to move worldwide. The primary transportation fuels generated from crude oil are gasoline, diesel, and jet fuel. Whereas gasoline remained the primary fuel used in automobiles with internal combustion engines. Thus, the qualities of gasoline used in automobiles vary depending on engine structure, characteristics, and operation circumstances (Lu et al., 2011).

1.2 Gasoline and Its Properties: Composition, Performance, and Environmental Impact

Gasoline, often referred to as petrol, is a flammable liquid fuel that is commonly used in internal combustion engines, particularly in cars, motorcycles and light trucks. It is a complex blend of hydrocarbons with varying chemical compositions and properties. It is primarily composed of hydrocarbons, including alkanes, aromatics, olefins, and cycloalkanes, along with additives such as oxygenates (e.g., ethanol or MTBE) to enhance combustion and reduce emissions. Its chemical composition varies depending on the refining process, but it typically contains a mixture of these compounds, contributing to its energy content, octane rating, and environmental impact. Gasoline, a ubiquitous liquid that fuels modern civilization, embodies both the promise of progress and the perils of dependency. Derived from

fossil fuels, primarily crude oil, gasoline has become the lifeblood of transportation, industry, and commerce worldwide. However, its profound impact on the environment, geopolitics, and human health raises critical questions about its sustainability and the imperative for alternative energy sources.

At its core, gasoline represents unparalleled energy density, enabling the efficient movement of goods and people across vast distances. From the combustion engine to the jet turbine, gasoline has revolutionized transportation, connected distant lands and fostered economic interdependence. Its role in shaping the modern world cannot be overstated, underpinning globalization and the interconnectedness of nations.

Yet, this very dependence on gasoline comes with significant environmental consequences. The combustion of gasoline releases carbon dioxide and other greenhouse gases, contributing to climate change and global warming. The extraction, refining, and distribution of petroleum exact a heavy toll on ecosystems, polluting air, water, and soil. Moreover, the finite nature of fossil fuels necessitates a reckoning with their depletion, prompting urgent calls for sustainable alternatives.

Geopolitically, gasoline holds immense strategic importance, shaping the dynamics of international relations and conflicts. Nations capable with abundant oil reserves wield significant influence on the world stage, while others contend with geopolitical instability and resource dependency. The quest for energy security drives geopolitical manoeuvring, underpinning alliances, rivalries, and conflicts that reverberate across the globe.

On a human level, the widespread use of gasoline exacts a toll on public health, with air pollution from vehicle emissions linked to respiratory illnesses, cardiovascular diseases, and premature deaths. Vulnerable communities, disproportionately affected by pollution, bear the brunt of environmental injustices, exacerbating socioeconomic disparities and perpetuating inequities.

In light of these challenges, the imperative for innovation and transition to sustainable energy sources has never been clearer. From electric vehicles to renewable energy technologies, the quest for alternatives to gasoline is underway, driven by technological advancements and environmental imperatives. Electric vehicles, powered by clean energy sources, offer a promising path forward, reducing greenhouse gas emissions and mitigating environmental degradation.

Moreover, investments in renewable energy infrastructure and sustainable transportation systems herald a paradigm shift towards a greener, more resilient future. From solar and wind power to biofuels and hydrogen, diverse energy solutions promise to reduce dependence on fossil fuels and mitigate the impacts of climate change.

However, the transition away from gasoline requires concerted efforts at the individual, societal, and policy levels. Incentives for renewable energy adoption, investment in green technologies, and regulatory measures to curb emissions are essential steps towards a sustainable energy future. Education and awareness campaigns can empower individuals to make informed choices and embrace environmentally conscious lifestyles.

Therefore, gasoline stands as both a symbol of human ingenuity and a harbinger of environmental degradation. Its pivotal role in driving progress must be

tempered by a recognition of its limitations and environmental costs. The transition to sustainable energy sources represents a moral imperative and an opportunity to forge a more equitable, resilient future for generations to come. By embracing innovation, collaboration, and stewardship of the planet, we can chart a course towards a world less reliant on gasoline and more attuned to the imperatives of sustainability and justice.

Gasoline exerts profound effects on both the environment and human health due to its chemical properties and combustion characteristics.

Combustion and Energy Release: Gasoline is highly combustible due to its composition of hydrocarbons, which contain a high amount of chemical energy stored in their bonds. When ignited in the presence of oxygen, gasoline undergoes combustion, releasing heat energy that powers internal combustion engines in vehicles, generators, and other machinery. This combustion process produces carbon dioxide (CO2), water vapor (H2O), and various other combustion byproducts.

Emissions and Air Pollution: The combustion of gasoline emits not only CO2 and water vapor but also pollutants such as nitrogen oxides (NOx), carbon monoxide (CO), volatile organic compounds (VOCs), and particulate matter (PM). These pollutants contribute to air pollution, smog formation, and adverse health effects in humans, including respiratory issues, cardiovascular diseases, and exacerbation of asthma.

Vapor Pressure and Volatility: Gasoline's chemical properties include its vapor pressure and volatility, which determine its ability to vaporize and form combustible mixtures with air. Gasoline with higher vapor pressure and volatility evaporates more

readily, contributing to air pollution through vapor emissions, especially in warm weather or during refuelling operations.

Octane Rating: Gasoline's octane rating measures its resistance to knocking or detonation in internal combustion engines. Higher-octane gasoline is less prone to premature ignition, allowing for more efficient engine performance and reduced engine damage. This property is crucial for optimizing engine performance and fuel efficiency in vehicles.

Chemical Composition and Additives: Gasoline comprises a diverse mixture of hydrocarbons, including alkanes, alkenes, and aromatic compounds, each with distinct chemical properties and combustion characteristics. Additionally, gasoline often contains additives such as detergents, antioxidants, and anti-knock agents to enhance engine performance, reduce emissions, and prevent engine deposits and corrosion.

Environmental Impact: Gasoline's chemical properties and combustion emissions have significant environmental consequences, contributing to climate change, air pollution, and ecosystem degradation. The combustion of gasoline releases CO2, a major greenhouse gas implicated in global warming and climate disruption. Furthermore, emissions of NOx, VOCs, and other pollutants contribute to the formation of ground-level ozone, smog, and acid rain, with deleterious effects on human health, vegetation, and aquatic ecosystems.

In summary, gasoline's chemical properties influence its combustion behaviour, emissions profile, and environmental impact. Understanding these properties is essential for mitigating the adverse effects of gasoline use, promoting cleaner technologies, and transitioning towards more sustainable energy sources.

Efforts to improve fuel efficiency, reduce emissions, and develop alternative fuels and propulsion systems are critical for addressing the environmental and health challenges associated with gasoline consumption.

Gasoline, a quintessential fuel powering our modern world, represents a double-edged sword in its purity and adulteration. The adulteration of gasoline, a practice as old as its commercialization, poses multifaceted risks to public health, the environment, and the economy. This essay delves into the ominous implications of gasoline adulteration, exploring its causes, consequences, and potential remedies.

1.3 Blended Properties of Gasoline: Performance and Environmental Compliance

In gasoline formulation, various additives are blended with the base gasoline to enhance performance, improve fuel efficiency, and reduce engine knock. Two such additives that have been used historically in gasoline formulations are Benzene and MTBE (methyl tertiary-butyl ether). However, both of these substances have raised significant environmental and health concerns, leading to changes in regulations and their usage in gasoline.

Benzene is a colourless, volatile liquid with a sweet odour that occurs naturally in crude oil and is present in gasoline. As one of the simplest aromatic hydrocarbons, benzene has been used in gasoline since its early days, primarily due to its ability to increase the octane rating of fuel. It was historically added to gasoline to enhance engine performance. The octane number is crucial in preventing engine knocking, which can cause damage to engines. A higher-octane rating allows for more efficient

combustion. Despite its benefits in improving fuel performance, benzene is a known carcinogen and poses significant health risks. Chronic exposure to benzene, even in small amounts, can cause blood disorders such as leukaemia and other cancers (Bonfim et al., 2012). Benzene also contributes to air pollution, and its emissions from gasoline combustion contribute to smog formation. This has led to significant regulatory measures to reduce its levels in gasoline. The U.S. Environmental Protection Agency (EPA) and similar bodies around the world have enforced strict limits on benzene content in gasoline. The Clean Air Act of 1990 in the U.S. significantly reduced the amount of benzene allowed in gasoline (Swick *et al.,* 2014). As a result, refiners have taken steps to lower benzene concentrations by using alternative additives and adjusting the formulation of gasoline.

In addition to Benzene, MTBE, a synthetic organic compound, has been utilized as an oxygenate additive in gasoline. It was introduced in the 1970s as a substitute for lead and other harmful compounds to enhance combustion efficiency and reduce air pollution (Yu et al., 2006). MTBE is an oxygenate, meaning it increases the oxygen content in gasoline, promoting more complete combustion. This helps reduce the emissions of carbon monoxide and other pollutants, particularly in urban areas with high traffic volumes. MTBE also improves octane levels in gasoline, making it an attractive additive for refineries.

While MTBE can reduce harmful emissions from vehicles, it has raised serious environmental concerns. MTBE is highly soluble in water and can contaminate groundwater if it leaks or spills. Even at low concentrations, MTBE can make water sources undrinkable due to its strong, unpleasant taste and odour. Furthermore, MTBE

is not easily removed from water, making it a persistent environmental contaminant. Due to concerns about MTBE contamination of drinking water and its potential health impacts, several states in the U.S. have banned or phased out its use in gasoline. The EPA, in the early 2000s, began recommending the phase-out of MTBE in favour of safer alternatives (Kümmerer., 2011). As a result, ethanol, another oxygenate, has largely replaced MTBE in gasoline.

However, the blending of benzene and MTBE in gasoline can compound the risks associated with fuel. Benzene's toxicity as a carcinogen and MTBE's potential to contaminate water supplies can create a dangerous synergy. As a result, refineries have been under increasing pressure to reduce or eliminate the use of both substances (Keenan *et al.,* 2010).

1.4 Adulteration in Gasoline: Impact on Quality and Engine Performance

Gasoline adulteration, the deliberate mixing of gasoline with cheaper or harmful substances, stems primarily from profit motives and lax regulatory oversight. Unscrupulous actors within the fuel industry resort to this practice to increase profit margins by diluting high-cost gasoline with cheaper substances like kerosene, diesel, or even water. The absence of strong regulatory frameworks, combined with ineffective monitoring systems creates wrongdoers feel empowered to engage in unethical activities without fear of consequences.

The consequences of gasoline adulteration reverberate across multiple domains. From an environmental standpoint, the combustion of adulterated gasoline releases a cocktail of toxic pollutants into the atmosphere, exacerbating air quality

issues and contributing to global warming. Moreover, the discharge of unburned hydrocarbons and particulate matter from adulterated gasoline engines poses a significant threat to ecosystems and public health, manifesting in respiratory ailments and ecological degradation.

The adverse economic ramifications of gasoline adulteration are equally profound. The compromised performance of vehicles running on adulterated gasoline leads to increased maintenance costs and reduced fuel efficiency, burdening consumers and undermining the automotive industry's productivity. Furthermore, the erosion of consumer trust in the integrity of gasoline suppliers' precipitates market distortions and fosters an atmosphere of uncertainty detrimental to economic stability.

Addressing the scourge of gasoline adulteration demands a multipronged approach encompassing legislative reforms, technological innovations, and public awareness campaigns. Governments must enact stringent regulations mandating periodic quality assessments and imposing severe penalties on violators to deter illicit practices. Leveraging advanced analytical techniques such as chromatography and spectroscopy can enhance the detection capabilities of regulatory agencies, facilitating the swift identification of adulterated gasoline.

Additionally, fostering a culture of consumer vigilance through educational initiatives and outreach programs can empower individuals to make informed choices and demand accountability from gasoline suppliers. Collaborative efforts between governmental agencies, industry stakeholders, and civil society organizations are indispensable in combating the pervasive threat of gasoline adulteration and safeguarding the collective well-being of society. So that, the adulteration of gasoline

represents a pernicious challenge with far-reaching implications for human health, environmental sustainability, and economic prosperity. As custodians of our planet's future, it behoves us to confront this menace with resolve and resilience, marshalling the collective wisdom and resources necessary to preserve the integrity of our most vital resources.

By confronting the adulteration of gasoline head-on, we not only mitigate immediate hazards but also pave the way for a cleaner, safer, and more equitable world for generations to come.

In India, the most commonly used transport fuel of vehicles is gasoline or petrol. Among the largest petrol pump companies is Indian Oil Corporation Ltd. (IO), with over 34,000 retail outlets across the country. The other major companies are Hindustan Petroleum Corporation Ltd (HP) and Bharat Petroleum Corporation Ltd. (BP), also called reformulated gasoline (RFG). RFG is gasoline blended to burn more cleanly than conventional gasoline and to reduce smog-forming and toxic pollutants in air. One of the popular chemical substances blended with gasoline is oxygenated compounds, namely, methyl tert-butyl ethers (MTBE) to reduce pollution carbon monoxide (CO) as well as to increase the octane number of engine (Li et al., 2011). In the state of Mizoram, North-Eastern of India, the largest fuel pump station is also IO, which have over 80% outlets, and the rest of them are 20% of HP and BP. However, the other main problems of gasoline transport fuel are adulteration. In the past few years, the Supreme Court (SC) itself state that auto adulteration is rampant India and largest cases of such adulterated petrol pump outlets are IO and next followed by HP and BP. Some of the different substances that blended gasoline adulteration are

naphtha, kerosene, rubber solvents, aromatics, ethanol, light and heavy aromatics, thinners, white spirits, turpentine, and lubricants. However, in spite of these, adulteration is still remaining difficult to monitor, control and detection (Vempatapu et al., 2017).

1.5 Condition of Gasoline in Mizoram: Quality, Composition, and Challenges

In Mizoram (Northeast India), gasoline is the most commonly used transportation fuel for vehicles. Among the 54 (2019) gas stations are Indian Oil Corporation Ltd. (IOC), which is the largest petrol supplier with 46 petrol stations across the state of Mizoram. The other major companies are Hindustan Petroleum Corporation Ltd (HPC) and Bharat Petroleum Corporation Ltd. (BPC), also called reformulated gasoline (RFG). RFG is a gasoline blend with oxygenator that burns cleaner than regular gasoline and reduces the formation of smog and toxic pollutants in the air. However, the quality of gasoline products is generally controlled by blending different refinery streams which in turns is reflected by engine performance. So that, in order to be marketed, Benzene, the most widely added aromatic compound to gasoline, is limited in India due to its carcinogenic chemical substance (Bonfim et al., 2012). To create a different grade of gasoline, oxygenator is also to fuel added chemical compounds to enhance octane rating and reduce air pollution. Methyl tert-butyl ether (MTBE) is the most often used oxygenator. Additionally, it is strictly limited inside the nation due to environmental concerns, mostly pertaining to ground water pollution (Prasad *et al.,* 2008). The permissible levels of aromatics and oxygenates in gasoline may vary by region and are subject to regulation. Different

regions or countries may have different standard and methods for testing gasoline quality, further complicating the identification and quantification of these components.

On the other hand, the other main problems of gasoline transport fuel are adulteration. Adulteration is the unlawful addition of any foreign material to a product that causes it to fail beyond its original specifications. It is common in the motor vehicle fuels (gasoline and diesel) sectors in developing countries, causing fuel quality to become poorer. Adulteration of petroleum-based automotive fuels is indeed a serious with several negative consequences. Adulterated fuels with impurities or lower quality substances can harm the engine's performance, can lead to long term damage to various engine components. Adulterated fuels often burn less cleanly, leading to higher emissions of harmful pollutants. This contributes to air pollution and possess to both the environment and living beings. Some of the different substances that blended gasoline adulteration are PDS Kerosene, naphtha, rubber solvents, aromatics, thinners, white spirits, turpentine, and lubricants (Vempatapu et al., 2017). To combat gasoline adulteration, government and regulatory bodies must enforced stringent regulations, conducted regular testing, and invested in advanced detection method such as Chromatography and Spectroscopy. Nevertheless, adulteration continues to be challenging to detect, monitor, and control.

1.6 Purpose of the Research Work

Several factors influence the quality and performance of gasoline produced at refineries, including boiling range, flash point, octane number, and oxygenates. Aromatic compounds like benzene, toluene, and xylene, as well as oxygenators, are used to control the octane rating. Oxygenators, due to their oxygen content, play a

crucial role in enhancing the fuel's combustion efficiency. However, benzene is a known carcinogen, and MTBE is also suspected of being environmental pollutant. Due to environmental and health concerns, the use of MTBE has decreased in the United States because it is highly soluble in water and can easily contaminate groundwater if leaked or spilled. Consequently, regulations now limit the concentration of these compounds in gasoline to between 1% and 15%. This study focuses on analysing the levels of these compounds in gasoline samples from Aizawl, using FTIR and multivariate techniques to investigate local variations. Additionally, it aims to detect gasoline adulteration in Mizoram (Northeast India) using FTIR combined with chemometric tools, and to compare various multivariate data analysis methods for classification and quantification. Particularly, there is a lack of reports on vibrational analysis of gasoline quality check with chemometric techniques in India, particularly in the Northeastern region.

Detection and identifying adulterated gasoline, as well as determining the contents of aromatics and oxygenators in gasoline, can indeed be challenging task due to the complex nature of the solvents used for this purpose and the presence of similar molecules in unadulterated gasoline. In order to identify and quantification of gasoline adulteration, gas chromatography (GC)/mass spectrometry offered excellent sensitivity and specificity, making it a valuable tool in the analysis of fuels and other substances (Skrobot *et al.,* 2005). It is also considered to be one of the most accurate and reliable analysis methods. However, there are some limitations, including the high cost and time needed to make a single chromatogram, which can take more than 30 minutes, and the analysis sometimes requires skilled experts in the field. In some cases,

a combination of different analytical techniques may be used to achieve a balance between accuracy and efficiency. Apart from these, the analytical methods that combined vibrational spectroscopy Near-Infrared (NIR), Mid-infrared (MIR), and Raman spectroscopy with chemometric tools are commonly used to analyse gasoline fuels (Bezerra *et al.*, 2019). These techniques offer a powerful and non-destructive way to obtained detail information about the composition and properties of gasoline samples. It also allows for rapid and accurate analysis, helping to ensure product quality, compliance and regulations, and process optimization in petroleum industry. They are particularly valuable for accessing chemical composition, quality, and properties of gasoline without the need of for extensive sample preparation or time-consuming laboratory techniques.

Several research have been performed regarding the uses of spectroscopic methods combined with multivariate data analysis. In Yousefinejad et al, 264 fuel oil samples from three different types of oil—furnace oil, gas oil, and mazut oil—were classified using ATR-FTIR spectroscopy. The outcome demonstrates that while Extended Canonical Variates Analysis (ECVA) performed well for classification purposes, PCA was unable to distinguish between different fuel oils. However, internal Extended Canonical Variates Analysis (iECVA) proved successful in highlighting the FTIR spectra's fingerprint sub-regions and reducing the complexity of the model. It concludes that the combination of chemometrics with ATR-FTIR spectroscopy provides a quick, effective, and trustworthy method for determining the grade of fuel oils, which is crucial for quality assurance and environmental monitoring(Yousefinejad et al., 2016). In Pereira *et al.* investigation, the results show

that FTIR combined with PCA-LDA is a successful technique for automotive gasoline quality control. The technique's sensitivity was 8% v/v, and its classification accuracy was 96% for adulterated and unadulterated gasoline and 93% for the type of solvent used (Pereira *et al.* 2006). Mohammadreza *et al.* showed that using the support vector machine regression (SVMR) model, the paraffin, olefin, naphthene, and aromatic contents of gasoline samples were also determined using the ATR-FTIR spectrum data. For comparison, partial least square regression (PLSR) and partial least squares discriminant analysis (PLS-DA) were also used for prediction and classification. Comparing the genetic algorithm data driven soft independent modeling of class analogy (GA-DD-SIMCA) and GA-SVMR methods to the GA-PLS-DA and GA-PLSR methods, respectively, showed that the first could provide superior classification and prediction capabilities (Khanmohammadi et al., 2022). Mahmodi *et al.* was carried out employing a variety of techniques for classification of biodiesel, including as support vector machines (SVM) and linear and quadratic discriminant analysis (LDA and QDA). The findings showed that the classification precisions of SVM, QDA, and LDA were, respectively, 94.8%, 94.1%, and 87.1%. Furthermore, the SVM's discriminating and classification precision for the two groups of pure and impure fuels was higher (about 95.4%). Applying the overall desirability function demonstrated that the QDA technique performed better than LDA and SVM in terms of discrimination and classification capacity (Mahmodi et al., 2019). Xu *et al.* investigated the findings of several chemometric techniques, including PLS-DA, LDA, PCA-SVM, and LDA-SVM in MIR spectroscopy applications for lubricating oil classification. The LDA-SVM model produced 100% favourable results in both the calibration and verification sets (Xu *et al.,* 2023). However, it is important to note that the success of FTIR in

combination with multivariate data analysis in classifying and determining gasoline samples relies on the choice of appropriate chemometric methods, and a well-established reference database for model development. Proper calibration and validation of the models are also crucial to ensure reliable results.

CHAPTER 2

METHODOLOGY

This chapter details the systematic approach to collecting and preparing gasoline samples from various sources in Mizoram, ensuring that the samples are representative and suitable for analysis. The combination of FTIR spectroscopy and chemometric tools provides a powerful methodology for assessing gasoline quality, identifying adulterants, and ensuring that fuel meets required standards for performance and environmental compliance. This robust analytical framework contributes to understanding fuel composition and detecting any irregularities that could affect both engine performance and public health.

2.1 Sample Collection and preparation

Indian Oil Corporation Limited (IOCL) supplied the majority of the fuel used by motor vehicles in Mizoram. Another petroleum company operating a network of filling stations throughout the state is Bharat Petroleum Corporation Limited (BPCL). In addition, Hindustan Petroleum Corporation Limited (HPCL) also maintains fuel distribution networks within Mizoram. During 2017-2020, the samples collection was done thrice a year on different filling stations inside Aizawl city. The 147 unknown test samples from Mizoram were also provided by IOCL, which is collected from different filling stations. Since there is no depot within the state, all fuel is imported from neighbouring states. As comparison samples, 12 additional unknown samples from the nearby state of Assam were collected. The samples collected from nearby

states are all around oil depots. It is also thought to be as pure and of higher grade than Mizoram. Additionally, a sample of 2.5 litres of gasoline is taken from the oil depot. This sample is used to make standard solutions with kerosene. The public distribution system (PDS) Kerosene fuel, which is used as a solute in the standard solution is also acquired from one retail outlet in Mizoram.

The samples are collected directly from the distribution sight and stored in polyethylene terephthalate in quantities of 500 ml apiece. All samples were stored in a cool and dry place before analysis. Thus, all the samples are divided into three categories: Mizoram samples, samples from nearby states and kerosene blended samples, indicated as MZ, NS and KS respectively.

2.2 Fourier Transform Infrared (FTIR) spectroscopy

Fourier Transform Infrared (FTIR) spectroscopy stands as a cornerstone technique in modern analytical chemistry, providing invaluable insights into molecular structures and compositions across diverse scientific domains. This essay delves into the principles, applications, and significance of FTIR spectroscopy in elucidating the mysteries of matter and enhancing our understanding of natural and synthetic compounds.

2.2.1 Principles of FTIR Spectroscopy:

At its core, FTIR spectroscopy connects the fundamental principles of molecular vibrational spectroscopy. When subjected to infrared radiation, molecules absorb energy at characteristic frequencies corresponding to the vibrational modes of their constituent bonds. The FTIR instrument, equipped with an interferometer, splits

incident infrared radiation into two beams, facilitating the measurement of absorption spectra and enabling the identification of functional groups and chemical compositions within samples.

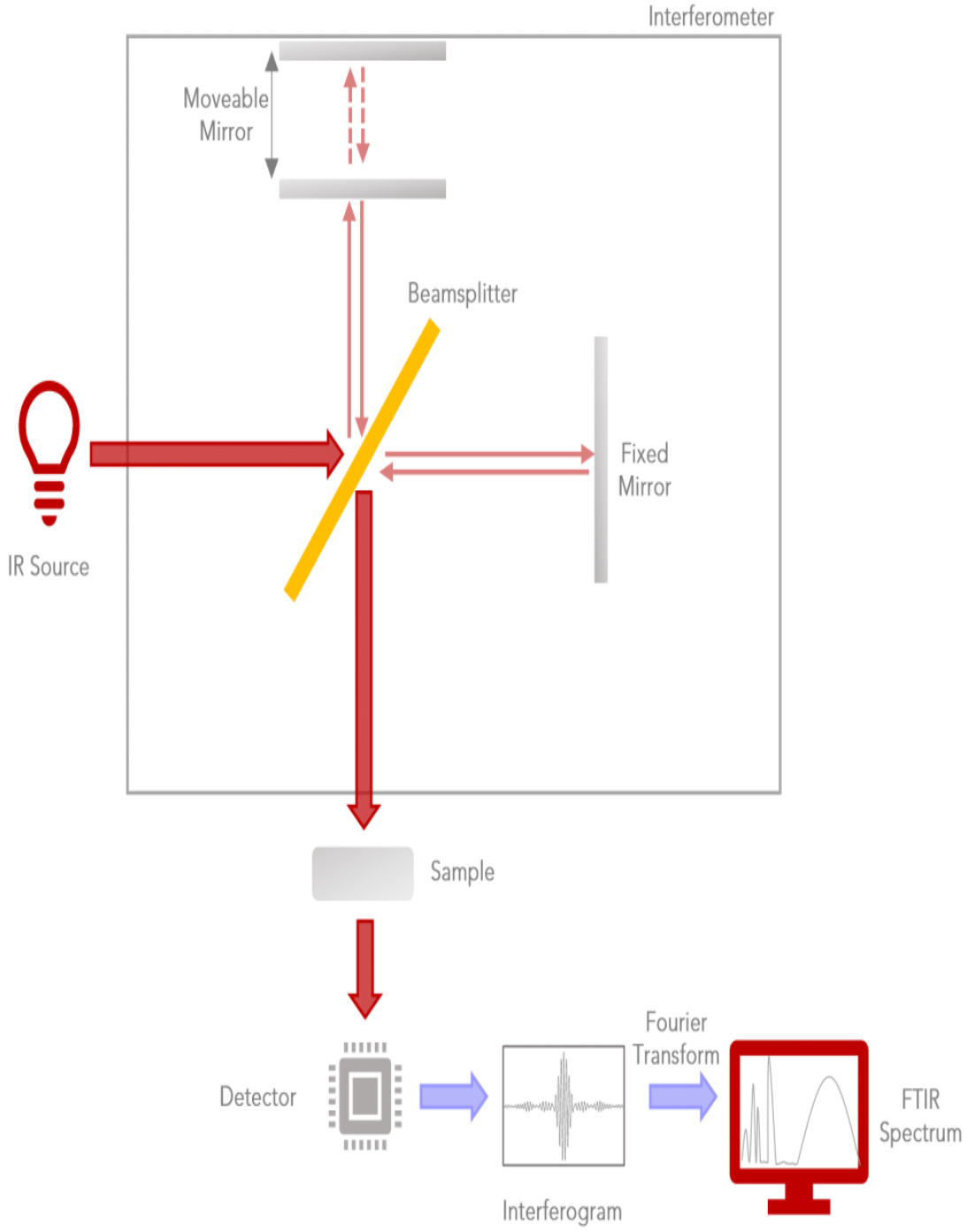

Figure 2.1: *Schematic diagram of FTIR (Edinburgh Instrument)*

2.2.2 Applications of FTIR Spectroscopy:

Chemical Analysis: FTIR spectroscopy serves as a powerful tool for qualitative and quantitative analysis of organic and inorganic compounds. By elucidating absorption patterns in the infrared spectrum, chemists can identify specific functional groups and assess molecular structures with unparalleled precision.

Material Science: In the realm of material science, FTIR spectroscopy plays a pivotal role in characterizing polymers, composites, and other materials. Researchers leverage its capabilities to investigate bonding interactions, assess material purity, and monitor structural changes under various conditions.

Pharmaceutical Industry: Within the pharmaceutical industry, FTIR spectroscopy emerges as a critical technique for drug development, formulation, and quality control. By analyzing infrared spectra, scientists can validate drug compositions, detect impurities, and ensure the efficacy and safety of pharmaceutical products.

Environmental Monitoring: FTIR spectroscopy finds widespread applications in environmental monitoring and analysis. From detecting pollutants in air and water to assessing soil contamination, this technique facilitates the identification and quantification of harmful substances, aiding in environmental preservation and remediation efforts.

Forensic Science: Forensic scientists rely on FTIR spectroscopy to analyse trace evidence, such as Fibers, paints, and illicit drugs. By examining infrared spectra, forensic analysts can establish crucial links between crime scenes and suspects, contributing to the administration of justice.

2.2.3 Significance of FTIR Spectroscopy:

The significance of FTIR spectroscopy transcends disciplinary boundaries, offering multidimensional insights into the molecular world. Its non-destructive nature, high sensitivity, and versatility render it indispensable in scientific research, industrial applications, and forensic investigations. As technology continues to advance, the integration of FTIR spectroscopy with complementary techniques promises to unlock new frontiers in analytical chemistry, materials science, and biomedical research, paving the way for groundbreaking discoveries and innovations.

Therefore, Fourier Transform Infrared spectroscopy stands as a beacon of scientific inquiry, illuminating the intricate tapestry of molecular structures and compositions with unparalleled clarity. From elucidating the secrets of ancient artifacts to advancing modern drug discovery, its applications are as diverse as the molecules it scrutinizes. As we journey deeper into the realms of science and technology, the enduring legacy of FTIR spectroscopy persists as a testament to human ingenuity and the relentless pursuit of knowledge.

2.3 FTIR-ATR Spectral Processing

The spectra of collected fuel samples and laboratory kerosene blended were acquired using ABB MB3000 FTIR spectrometer equipped with Zinc Selenide (ZnSe) Attenuated Total Reflectance (ATR) crystal purchased from ABB Inc. Quebec, Canada. Which is one of the most reliable FT-IR laboratory analyzer with the lowest cost of ownership. Each spectrum was collected in the mid-IR band of 3200 - 650 cm^{-1} with a resolution of 4 cm^{-1} and 30 scans. Every specimen was conducted in a closed

ATR with no further evaporation. The sampling area was then cleaned thrice with the following sample in between runs.

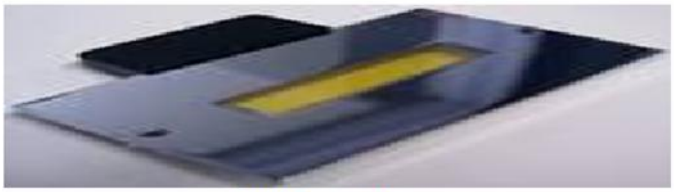
ATR Crystal

Figure 2.2: *Horizon MB 3000 FTIR ATR Crystal Spectrometer.*

2.4 ASTM D6277

ASTM D6277 is a standard test method developed by ASTM International (formerly known as the American Society for Testing and Materials) that is used to determine the aromatic content of gasoline and other petroleum products. This test is important because the aromatic content of fuel can affect engine performance, emissions, and overall fuel quality.

This test method determines the concentration of benzene in spark-ignition engine fuels. It applies to benzene levels ranging from 0.1 to 5 percent by volume. The specified units of measurement are the preferred units and are used throughout this standard. However, it is designed for quality control during the production and distribution of spark-ignition engine fuels. The Mid-IR Test Apparatus, which is suitable for this method, uses specific wavelengths from the spectrum that are strongly correlated with benzene or potential interferences. These wavelengths are selected using selective bandpass filters or through mathematical selection of specific areas of the entire spectrum. A multivariate mathematical analysis then converts the detector's response for these selected areas into a benzene concentration for the unknown sample. This test method is quick, easy to perform, and cost-effective (D02 Committee, n.d.).

2.5 Chemometric Analysis

Chemometrics analysis is a data driven method that analyses chemical data to produce the most pertinent chemical information possible. It does this by using methods such as statistical analysis, formal logic, and mathematics to design or choose the best measurement protocols and experiments (Héberger., 2008).

Chemometric data preprocessing was obtained by using The Unscrambler X (version 10.4.1) CAMO chemometric tool. The primary objective of data preprocessing is to reduce unwanted variability or effects from the signal, giving useful data about the property of interest to be used for efficient modelling. The specific objectives of the preprocessing approaches depend on the type of artefacts to be dealt with (Mishra *et al.,* 2020). So that, the preprocessing method of the savitzky-golay smoothing and standard normal variate were used before analysis.

2.5.1 Savitzky-Golay smoothing and Standard Normal Variate (SNV)

The Savitzky-Golay algorithm fits a polynomial to each succeeding curve segment, resulting in more regular variations of the original values. It is a very successful method for removing spectral noise spikes while keeping chemical information. All mathematical equations used in the research methodology were directly generated by the Unscrambler® 10.4.1 software (CAMO software AS, Oslo, Norway).

By applying SNV, each spectrum is centered and then scaled by dividing by its standard deviation, i.e.,

$$x_{ij}^{SNV} = \frac{(x_{ij} - \bar{x}_i)}{\sqrt{\frac{\sum_{j=1}^{p}(x_{ij} - \bar{x}_i)^2}{p-1}}} \qquad (1)$$

where x_{ij}^{SNV} is the element of the transformed spectrum and x_{ij} is the corresponding original element of the spectrum i at variable j, $\bar{x}_i$ is the mean of spectrum i, and p is the number of variables or wavelengths in the spectrum

2.5.2 Principal Component Analysis (PCA)

The principal component analysis (PCA) can uncover the hidden structure within significant datasets. It offers a graphical depiction of the correlations between variables and samples, along with specifics regarding the method by which measured variables figure out samples are similarities or dissimilarities. PCA is a bilinear modelling method that provides an interpretable overview of the main information contained in a multidimensional table. It is also known as a projection method, because it takes information carried by the original variables and projects them onto a smaller

number of latent variables called Principal Components (PC). Each PC explains a certain amount of the total information contained in the original data, and the first PC contains the greatest source of information.

As for the first analysis, an unsupervised chemometric method of Principal Component Analysis (PCA) is used for finding the variation of gasoline fuel quality. The method involves decomposing a data matrix X into a structure part and a noise part, i.e.

$$X = TP^T + E \qquad (2)$$

where *T* and *P,* are the structure part of PCA and they represent the score vector and loading vector of the PCA, and *E* is the noise part that represent the residual matrix. The structured element of the data is made up of scores and loadings together. By multiplying the scores and loadings, the whole structure of the original data set may be recreated, with the intention of leaving just a tiny residual of random fluctuations that cannot be meaningfully modelled. When evaluating the findings of a PCA, one focusses on the structure component while discarding the residual part.

2.5.3 Soft Independent Modelling of Class Analogy (SIMCA)

Soft Independent Modelling of Class Analogy (SIMCA) means developing a PCA model for each class in the training set. Unknown samples are then compared to class models and classified based on their similarity to the training samples. The classification method is such a method that each class of samples is described by its own principal component model.

The term "*soft*" denotes that the approach is entirely data-driven, with no previous assumptions made about the distribution of the information collected.

"*Independent*" means that each class of objects under study is treated individually and separately, contrary to how a standard discrimination strategy would operate; "*modelling of class analogy*" implies that SIMCA focusses on the similarities among the samples belonging to the individually investigated category rather than the differences that would allow it to be distinguished from the others, again in contrast to how discriminant techniques operate (Vitale *et al.,* 2023).

SIMCA modelling requires the building of a single PCA model for each class that as accurately describes its structure as possible. The ideal number of PCs should be determined for each model independently, using a proper validation process. Each model should be examined for any outliers and adjusted if necessary (as with any PCA model). SIMCA analysis produces specific results in addition to the regular PCA results such as scores, loadings, and residuals. After each class has been modelled, and assuming that the classes do not considerably overlap, additional samples are projected onto each model used in the SIMCA classification process. The classification determination rule is based on traditional statistical methods. If a sample belongs to a class, it should have a close relationship to the class model.

In the present work, unknown gasoline samples are then compared to the class models and assigned to classes according to their proximity to the training samples. Classification by SIMCA is preferable as it obtain accurate results as compared to other type of classification.

2.5.4 Partial Least Square – Discriminant Analysis (PLS-DA)

Partial Least Square – Discriminant Analysis (PLS-DA), which is the use of Partial Least Square - Regression (PLS-R) for discrimination or classification

purposes, it is employed for classification of different gasoline samples. Partial Least Squares Regression (PLSR), also known as Projection to Latent Structures or simply PLS, simultaneously models the X- and Y-matrices to discover the latent (or hidden) variables in X which most accurately predict the latent variables in Y. These PLS components are analogous to primary components, but are referred to as factors. The PLS-R decomposition takes the form as

$$X = TP^T + E \qquad (3)$$

$$y = Uq^T + f \qquad (4)$$

where T and $P$are the scores and loadings of spectral data matrix X, E the residual matrix of X, U is the scores matrix of concentration matrix y, q and f are the loading and residual vector of y. PLS Discriminant Analysis (PLS-DA) is a classification approach that uses PLS to represent differences between several classes. If there are only two classes to differentiate, the PLS model employs a single response variable that codes for class membership as follows: -1 for members of one class, +1 for members of the other. If there are three or more classes, the model assigns one response variable (-1/+1 or 0/1, which is equivalent) to each class. The model also includes various Y-variables.

2.5.5 Support Vector Machine Classification (SVMC)

The Support Vector Machine (SVM) classification method is based on statistical learning. When a linear function is unable to describe complicated separations, SVM uses kernel functions to translate the original space to the feature space. The function can have many different forms, making it possible to handle

nonlinear classification cases. The kernels may be thought of as a mapping of nonlinear data to a higher dimensional feature space, while providing a computational shortcut by allowing linear algorithms to operate with higher dimensional feature spaces.

SVM classification is a supervised technique of classification. The data used for SVM must be a data matrix with a single category variable describing which classes the model will distinguish. For SVM classification, the X and Y matrices must contain equal numbers of rows (samples) and no missing data. The Y matrix must include a single column of category variables. The X data must be numerical and have no missing values. Support vector machines build a maximum separation hyperplane from input vectors in a higher-dimensional space. Two parallel hyperplanes are built on either side of the hyperplane to split the data and maximise the distance between the two parallel hyperplanes. The concept is that the greater the margin or distance between these parallel hyperplanes, the better the generalisation error of the classifier will be (Balabin et al., 2010). The main result of the SVM is the confusion matrix, which represents how many samples have been classified in each class, and the prediction matrix, which gives the classification computed for each sample in the training set.

2.5.6 Linear Discriminant Analysis (LDA)

The discriminant analysis is also a supervised classification and dimensionality reduction method, which is used to build models that can classify or categorize samples into predefined classes. It projects the data onto a lower-dimensional subspace while preserving as much class discrimination as possible. Unknown samples are categorised into the most likely class using the models. The application of discriminant analysis is

significant in aiding in the interpretation of variations amongst sample groupings. Since the objects need to be categorised using the created known model, LDA's discriminant methodology qualifies as a classification technique. Finding the best fitting parameters for classifying data from a created model is the primary objective of LDA. Following that, unknown samples are categorised using the model

The regression models were evaluated according to the following parameters of Root Mean Square Error of Prediction (RMSEP) and Coefficients of Determinants (R^2).

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n}} \quad (5)$$

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (6)$$

According to Equation 5 and 6, Specificity and Sensitivity were used to assess the SIMCA and PLS-DA models, respectively, in order to obtain a dependable calibration model.

$$Specificity = \frac{TP}{FP+TN} \quad (7)$$

$$Sensitivity = \frac{TN}{FN+TP} \quad (8)$$

While TP and TN represent the number of true positives and true negatives, respectively, FP and FN represent the number of false positives and false negatives.

2.5.7 Principal Component Regression (PCR)

PCR is a method for relating the variance in a response variable (Y-variable) to the variance of several predictors (X-variables), with explanatory or predictive purposes. PCR is a two-step procedure which first decomposes an X-matrix by PCA,

then fits an MLR model, using the PC scores instead of the original X-variables as predictors. This method performs particularly well when the various X-variables express common information, i.e., when there is a large amount of correlation, or even collinearity. In PCR, the X matrix is approximated by the first few principal components, obtained by one of the methods described in Algorithms Used for Calculating Principal Components in The Unscrambler.

$$X = TP^T + E = (UD)V^T + E \qquad (9)$$

Where,

- T = Scores (usually the first 'a' scores),
- P = Loadings,
- E = Residual,
- U = First 'a' left singular vectors,
- D = Singular Values,
- V = First 'a' right singular value.

The next step is to regress Y on the first few scores, using MLR and then calculating regression coefficients as follows,

$$B = P(T^T)^{-1}T^TY = VD^{-1}U^TY \qquad (10)$$

When the outcomes of multiple linear regression analysis differ from realities, it's often caused by multicollinearities between independent variables. At that point, one can do an analysis using PCR. Principal component regression analysis is an efficient technique. It is able to not only detect collinearity in each independent variable, and it also has the ability to correct the problem (Liu *et al.,* 2003).

2.5.8 Partial Least square Regression (PLSR)

Partial Least Squares (PLS), also known as Projection to Latent Structures, models both the X- and Y-matrices at the same time to determine which latent variables in X best predict latent variables in Y. These PLS components are similar to primary components, but are referred to as Factors. PLS maximises the covariance of X and Y.

Unlike PCR, which only takes into account the variance structure of the X matrix, the NIPALS PLS regression algorithm starts with extracting the largest eigenvector of $S = X^T Y$, which includes information on *X* and *Y* and their covariance (or correlation if the data are scaled to unit variance). This eigenvector is named *w*, or the first loading weight vector. This leads to the expression,

$$t = X_w = E_w \qquad (11)$$

w is then normalized to length 1.

X and *Y* loadings are then calculated by regressing against *t*, calculated above.

$$p = \frac{E^T t}{t^T} \qquad (12)$$

$$q = \frac{F^T t}{t^T} \qquad (13)$$

E and F are initially X and Y and are "deflated" during the calculation of PLS factors. The so-called Y-scores, *u*, are calculated from

$$u = \frac{F_q}{q^T q} \qquad (14)$$

Thus, the models for *X* and *Y* can be written as:

$$X = TP^T \qquad (15)$$

$$Y = UQ^T \qquad (16)$$

The inner relation in PLS regression is the relation between T and U for the individual factors:

$$h = u^T t / t^T t \qquad (17)$$

As u is a constant multiplied with t, it is conceptually simpler to have the same scores expressing both X and Y:

$$X = TP^T \qquad (18)$$

$$Y = TQ^T \qquad (19)$$

The process continues by deflating. The information of the PLS factors (i.e. the outer products, tp^T and tq^T) is subtracted from E and F to obtain

$$E_{n+1} = E_n - tp^T \qquad (20)$$

$$F_{n+1} = F_n - tq^T \qquad (21)$$

The process is now repeated to find the next PLS factors by finding the eigenvector $E_{n+1}^T F_{n+1}$. The estimation of the PLS loadings, loading weights and scores may also be achieved by extracting eigenvectors of the smallest size of products of X, X^T, Y and Y^T, which are the basis for other PLSR algorithms like the kernel and wide kernel methods (see below). The matrices, W, T, P and Q are then stored in The Unscrambler® Project Navigator with the PLSR results, for further diagnostic purposes. To ensure that the columns of the matrix W relate to the original matrix X, the weights may be expressed as,

$$R = W \left(P^T W\right)^{-1} \qquad (22)$$

The scores T are now used to calculate the regression coefficients, using the following expression,

$$B = R(T^T T)^{-1} T^T Y = W(P^T W)^{-1} Q^T \quad (23)$$

Since the normalization step can be introduced at various points in the calculation, using other variants of the PLSR algorithm, this can make it difficult to compare scores and loadings calculated by these variants.

However, PLSR aims to predict dependent variables based on a collection of independent factors. This prediction is made by extracting from the predictors a collection of orthogonal components known as latent variables that have the highest predictive ability. PLS regression is especially effective for predicting a collection of dependent variables from a (very) large set of independent variables (i.e., predictors) (Abdi., 2010).

2.5.9 Support Vector Machine Regression (SVMR)

SVMR is a regression technique based on statistical learning. When a linear function is unable to model complicated systems, SVMR uses kernel functions to convert the original space to the feature space. The function can have many different forms, making it possible to handle nonlinear regression cases. The kernels may be described as a mapping of nonlinear data to a higher dimensional feature space, while providing a computational shortcut by allowing linear algorithms to operate with higher dimensional feature spaces. SVM principles may readily be applied to problems related to regression, so they are becoming increasingly popular for quantitative spectrum analysis (Leal *et al.,* 2020).

CHAPTER 3

EVALUATION OF GASOLINE FUEL VARIATION IN AIZAWL AND CLASSIFICATION

In this chapter, gasoline fuels are evaluated and classified through two distinct approaches. First, the analysis focuses on fuels sourced from Mizoram, a neighbouring state, and laboratory kerosene blends. This step is aimed at assessing quality and making comparisons, using fuel exclusively from a single supplier. The second approach involves a broader classification of all samples, comparing conventional, reformulated, and kerosene-blended fuels. This comparison is intended to evaluate the quality of fuels available within Mizoram, considering that fuels are transported by road into the state at various times.

To improve the accuracy and predictive capability of FTIR for gasoline quality assessment, multivariate data analysis techniques such as PCA, SIMCA, PLSDA, and LDA are utilized. The FTIR spectroscopy in combination with these multivariate techniques provide a robust, efficient, and non-destructive approach to gasoline quality monitoring. By offering precise classification and predictive capabilities, this integrated approach enhances the quality control process, ensuring that gasoline meets the required standards for performance, safety, and environmental compliance.

In addition to quality monitoring, a comparison of these chemometric methods is also conducted for future applications. This is particularly relevant in regions where the preferred GC analysis is not accessible, highlighting the need for faster and more cost-effective techniques for future research.

3.1 Introduction

Gasoline adulteration, by mixing unwanted substances into the petroleum base fuels, is one of the acute problems in developing countries. Adulteration is described as the illegal or criminal insertion of any foreign substance into a product, causing it fails to comply with its specified specifications. Adulterants are foreign compounds that, when introduced, alter and reduce the quality of the base transport fuels. The basic purpose of adulteration is to deceive consumers into purchasing poor fuel in order to increase their earnings while avoiding the hazardous consequences of this action (Vempatapu et al., 2017). Some of the adverse effects of adulteration included engine malfunction, higher tailpipe pollution emissions, and health issues.

In India, transportation fuels are frequently adulterated with cheaper products, by-products, or waste hydrocarbon streams for monetary advantages. According to the Supreme Courts (SC) of the Central Government in the year 2016, one of the most common types of adulteration in the country is combining kerosene with gasoline and diesel. Not only that, but the amount of kerosene used into diesel can range from 20-30% while only a little amount is mixed with automobile gasoline (n.d.). Kerosene, also referred to as public distribution system (PDS), is a subsidised product provided by the government to those who are economically deprived. As a result, PDS kerosene is mostly used throughout the country for fuel adulteration.

Gasoline adulteration detection and identification are challenging tasks since the solvents used for this purpose are complex combinations of hundreds of distinct molecules, some of them might be found in unadulterated gasoline (Wiedemann et al., 2005). Detection of adulterants in automotive fuel using Gas Chromatography (GC)/

Mass Spectrometry is considered to be among the most accurate and reliable analytic methods (Doble *et al.,* 2003). However, there are several disadvantages, such as the high cost and required more than 30 minutes to obtain a single chromatogram, as well as sometimes requires trained technicians for analysis (Lee et al., 2018b).

The vibrational spectroscopy (NIR, MIR and RAMAN) in combination with chemometric methods has proven to a powerful tool for analysing gasoline fuels (Sales *et al.,* 2023). Detection and estimation of gasoline adulteration have been done by multivariate methods like PCA, PLS-DA and PLS regression, which were applied for statistical analysis of NIR spectral data (Mabood *et al.,* 2017). Raman spectroscopy can also be used to evaluate the quality of fuels in a remote, rapid, and non-destructive manner without the need for reagents (Bezerra *et al.,* 2019). However, Raman spectrometer is much more expensive than FTIR. Then, the FTIR spectroscopy combined with multivariate calibration analysis is a versatile, efficient and accurate technique for the simultaneous estimation of key gasoline properties (Q. Xia *et al.* 2019). Studies have shown that qualitative analysis result of Partial Least Squares Discriminant Analysis (PLS-DA) obtained 100% and 98.66% accuracy respectively in the calibration set and the prediction set (Q. Xia *et al.,* 2019). The Soft Independent Modeling Class Analogy (SIMCA) classification model had also been developed that predict the type of adulterant present at an error rate of 6.25% for Kerosene and 12.5% for premix (Dadson et al., 2018). FTIR associated with Principal Component Analysis – Linear Discriminant Analysis (PCA-LDA) had been used to determine gasoline adulteration with 96% efficiency and to identify the solvent added with 93% efficiency (Pereira *et al.,* 2006). However, in India vibrational analysis of gasoline adulterants with chemometric methods has not yet been reported in literature. Therefore, FTIR in

combination with multivariate data analysis of PCA, SIMCA, and PLS-DA has been selected to determine the regression and classification of adulteration of gasoline.

All petrochemical fuels available in Mizoram (North-east India) are imported from nearby states within India. As there have been several complaints in the past about adulteration of fuels and related concerns in vehicle engine performance, regular monitoring of gasoline quality against adulteration is critical for improving ambient air quality. During the past three years of 2019, the government received 4145 complaints from all other states except Mizoram, and 32 retail outlets were shut down nationwide. Since the state of Mizoram has so far not made any initiative steps in comparison to other states, it is essential to ensure and maintain fuel quality across the state. The present work is expected to be useful in fuel testing program to monitor the quality of fuels at the retail outlets and to ensure the compliant with the fuel quality standards. If the fuel quality is inspected at the distribution points with portable inexpensive tools and rapid measurement methods, adulteration of fuel might be gradually reduced.

3.2 Materials and Methods

Indian Oil Corporation Limited (IOCL) supplies most of the fuel consumed by automobiles in Mizoram. The gasoline obtained in Mizoram has not yet been monitored, and 142 samples were collected from various filling stations at IOCL. Since there is no depot within the state, all fuel is imported from neighbouring states. As comparison samples, 12 additional samples from the nearby state of Assam were collected. The samples obtained from nearby states are all close to oil depots. It is also believed to be as pure and of better quality than Mizoram. Additionally, a sample of 2.5 litres of gasoline is taken from the oil depot. This sample is used to make standard

solutions with kerosene contents of 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 22%, and 24% added to 10 mL of 21 gasoline samples, respectively. The public distribution system (PDS) Kerosene fuel, which is used as an adulterant in the standard solution, is also acquired from one retail outlet in Mizoram.

For the second analysis, the Mizoram samples, samples from nearby states, and kerosene blended samples, indicated as MZ, NS, and KS, respectively. A total of 78 gasoline samples were collected from filling stations of three companies, namely, Indian Oil (IO), Hindustan Petroleum (HP) and Bharat Petroleum (BP) in different parts of Aizawl, the capital of Mizoram, Northeastern India. Apart from this, one sample of conventional gasoline was collected from the Indian Oil Corporation Ltd of Marketing Division for standard sample, as well as 11 samples were also collected from nearby retail gasoline pumps of Silchar, Assam and taken as training set for comparison. The standard samples were then mixed with kerosene at the percentage level of 5-20 % v/v. In addition, the oxygenate was also chosen as standard model in order to span slight variations because the RFG samples that differ from conventional gasoline are included in the analysis. To identify the amount of oxygenate in gasoline, 15 standard solutions were prepared by blending conventional fuel with MTBE in the range of 1-15% v/v. Therefore, the standard samples were grouped into two sets: a calibration set representing 80% of the total samples, to create a model for analysis and a validation set representing 20% of the samples, to evaluate the performance of the calibration model.

3.3 Result and Discussion

3.3.1 First Analysis

3.3.1.1 Spectral Characterization

Figure 3.1 shows the actual FTIR-ATR spectra of gasoline sample and kerosene blended sample at the region of 3200-650 cm^{-1}. The primary distinction between gasoline and kerosene is that gasoline is a light mixture of hydrocarbons ranging from C_4 to C_{12} carbon atom per molecule, whereas kerosene is a medium weight mixture ranging from C_{10} to C_{16} carbon atom per molecule. The bands observed between 3200 and 2700 cm^{-1} correspond to alkyl C-H stretching vibrational modes, which are followed by in-plane C-H bending and C-C modes between 1600 and 1000 cm^{-1}, and the bands below 1000 cm^{-1} correspond to aromatic compounds' C-H out-of-plane (oop) modes. The primary distinction between gasoline and kerosene blended samples is visible at the peak of 2926 cm^{-1}, indicating that kerosene contaminated fuel exhibits a slightly higher peak at the asymmetric stretching of the methyl group of CH_2. Kerosene, on the other hand, has a lengthier carbon chain than gasoline fuel. Aside from that, another distinction between petrol and kerosene mixed samples can be found in the aromatic region of 800-650 cm^{-1}. It shows that the aromatic region peak is reduced by the kerosene concentration as compared to the gasoline sample. In addition to having a negative impact on motor engines and being less burnable than pure petrol, the adulterated fuel is hazardous for the ecosystem and people's health.

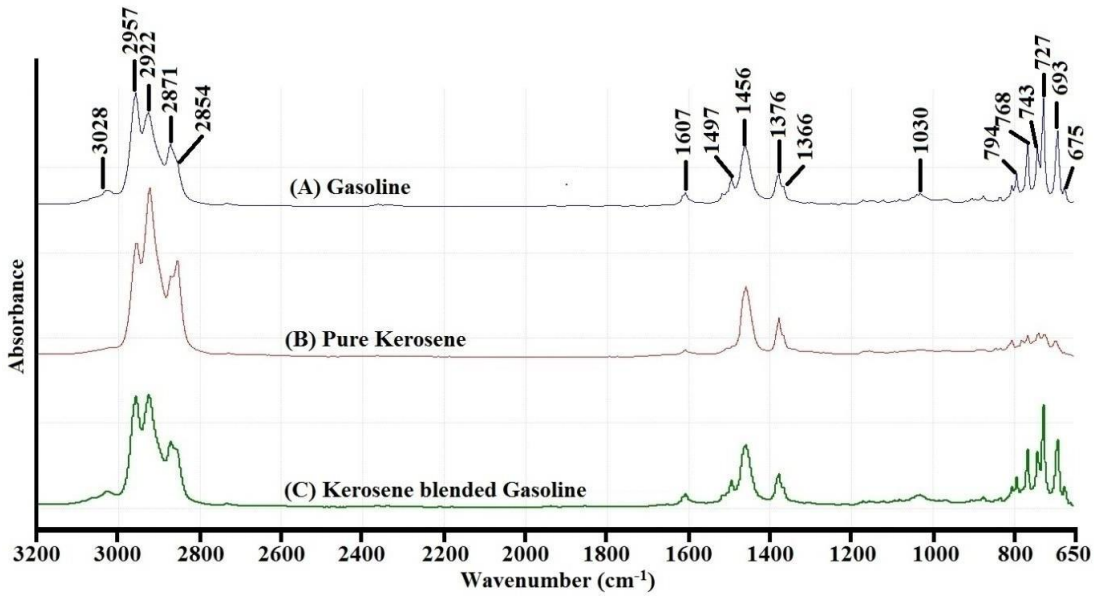

Figure 3.1: *IR Spectra of (A)Gasoline, (B)Pure Kerosene and (C)Kerosene blended in the region of 3200-650 cm^{-1}.*

3.3.1.2 Chemometric Data Processing

Chemometric analyses of unsupervised PCA, supervised SIMCA and PLS-DA techniques were performed on all of the samples at the mid-IR spectra ranging from 3200 to 650 cm^{-1}. Regarding the purpose of running supervised SIMCA and PLS-DA, the location of sample collection and the type of samples have been chosen as parameters. In order to enhance data analysis, all of the spectral data were initially processed using the Savistky-Golay smoothing and SNV processes as shown in Figure 3.2. However, it has produced better outcomes than without performing pre-processing method in the data analysis [24].

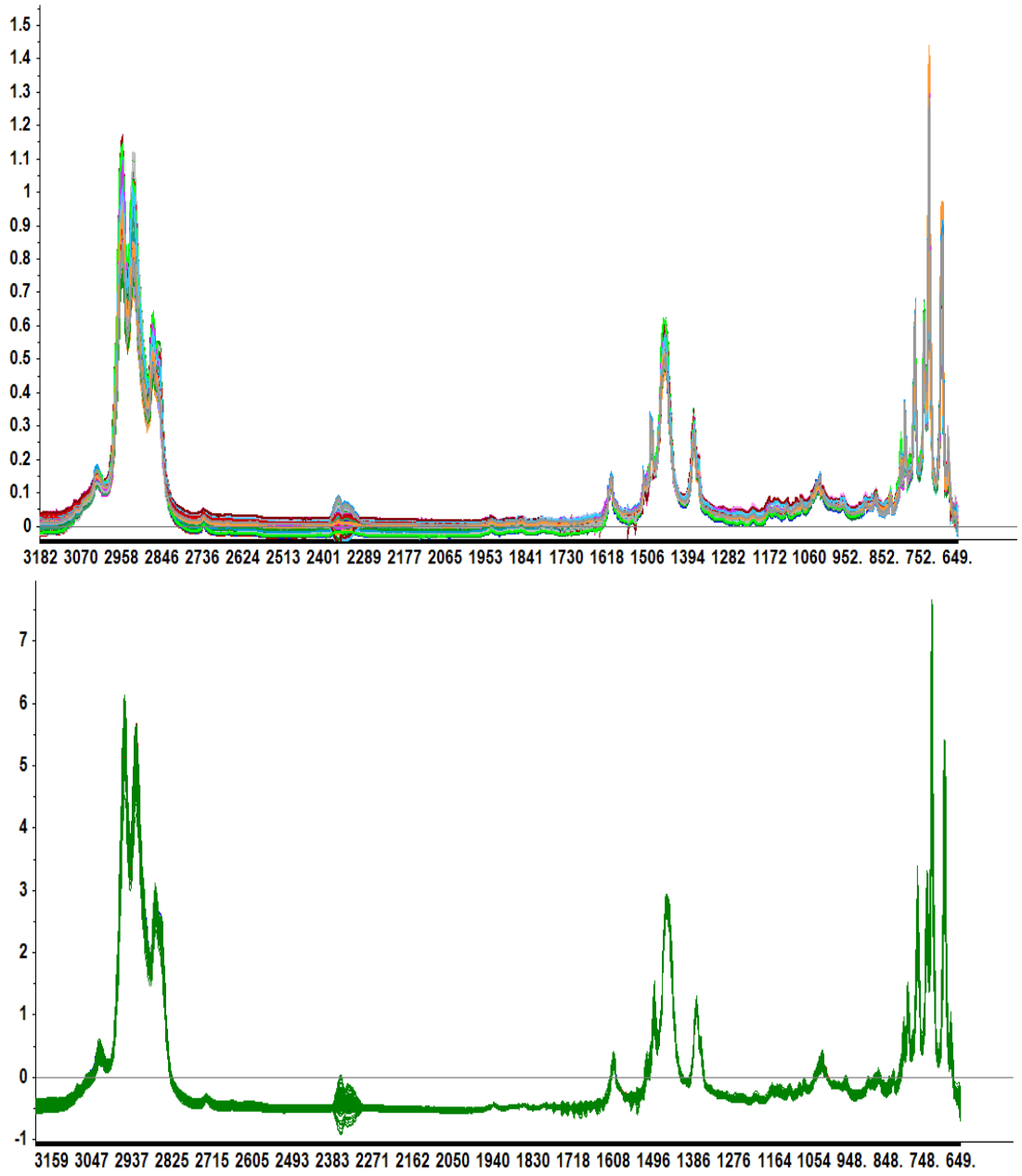

Figure 3.2: *IR Spectra of all samples (A) before and (B) after pre-processing transformation.*

3.3.1.3 Principal Component Analysis

Figure 3.3 illustrates the PCA 2D score and correlation loading plot of all type's gasoline samples in the entire region of 3200-650 cm^{-1}. In the score plot Figure 3.3 (a), the first two Principal Component explained 83% of the total variability on seven (7) Principal components i.e., PC1 68% and PC2 15%. If the sum of the explained variances for the first two components exceeds 70% of the total variability in PCA, the relationships can be interpreted with a high degree of certainty using the first two components. The Score plot shows the distribution of the samples depending on their similarities and difference of their IR spectra. Examining the score plot with the Hotelling T2 ellipse of 95% confidence level of detecting outlier. Figure 3.3(b) shows the 1D correlation loading Toolbar of the first two components. The peak observed in between the upper bound and lower bound are modelled by the PC's, while those that lies between the lower bounds are not. It is especially useful for evaluating important wavelengths in spectroscopic analysis or as contributing variables in time series data. The scores plot shows that all the kerosene blended samples are found inside the ellipse without any outlier. They all are cluster near the axes without giving separation from a large group of gasoline samples. The Major group of Mizoram samples including nearby state samples are mostly found inside the 1^{st} and 4^{th} quadrant. The variation of these two groups of samples is caused by the symmetric and asymmetric stretching vibrations of CH_3 at 2943, 2887 and 2864 cm^{-1} and the C-CH_3 symmetric bending vibration of the umbrella modes at 1371-1269 cm^{-1}. Which are the characteristics of normal gasoline sample and it is believed that this group of samples are supposed to be good quality in the analysis. Apart from that, three of the Mizoram

samples step out as outliers and several samples collected from the Mizoram area are obviously distributed away from large groups of samples on the negative side of PC1. The samples in this variation are also not only separated from major group of samples; it is also separated from kerosene blended samples. Thus, this variation is due to the out of plane C-H bending vibration of aromatic compounds at 796-692 cm^{-1}. It is also believed that the S=O stretching vibrations of sulphur oxide at 1030 cm^{-1} and the ring mode vibrations of Toluene at 1064-1496 cm^{-1} also contributed in this variation. The substantial variation of the Mizoram samples towards the positive side of PC1 is clearly related to the presence of many aromatic components in the sample. In addition, Fig.4 shows that a few samples are pulled to the positive side of PC2 and one sample from a nearby state being closer to the outlier. The separation towards positive side of PC2 is mostly driven by the Methylene group of C-CH_2 scissor bending vibration at 1446 cm^{-1} and the symmetric stretching vibration of CH_2 at 2848 cm^{-1}. The Methylene group is primarily responsible for samples variation in PC2.

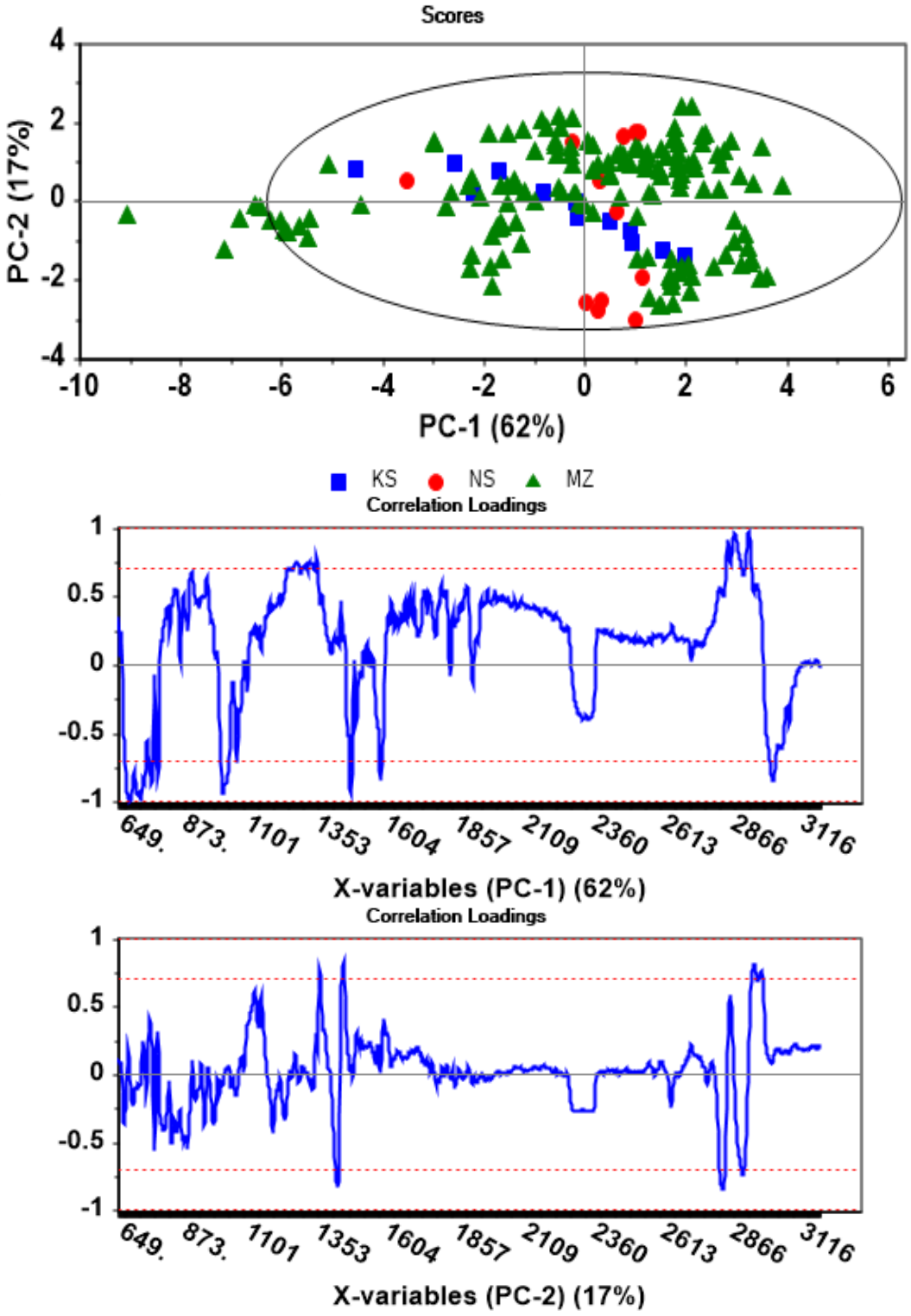

Figure 3.3: (*A) 2D-Score plot of PCA (KS-Kerosene Blended Samples, NS-samples collected from nearby state and MZ-samples collected from Mizoram) and (A) 1D-Correlation Loading Plot of PCA.*

After performing the analysis, we discovered that the known samples of kerosene blended samples are not separated from gasoline samples, there is no clear information on kerosene adulteration in this analysis. However, some of the Mizoram samples are significantly different from kerosene blended and other samples. It is also found that the major samples from Mizoram are all mixed up with the comparative samples from nearby states and kerosene blended samples. Furthermore, the unsupervised PCA chemometric technique does not clearly distinguish all of the samples, especially in the case of kerosene adulterated fuel.

3.3.1.4 SIMCA Analysis

SIMCA classification was used to determine the class of the samples and to investigate kerosene adulteration. Since there is no depot inside Mizoram, the samples taken from nearby states are all close to oil depots. It also appears to be of higher quality than Mizoram because all fuel samples have been obtained from neighboring states. So, the known kerosene blended samples and nearby state samples are used as model samples. The PCA two classes of 8 samples from Kerosene blended and 5 samples from nearby state samples were takes as the model samples. The model samples, also known as the training set are used to analyse 143 test samples or commercial samples of Mizoram. Table 3.1 and 3.2 shows the detailed information and performanced of the two classes of PCA models. The three chosen principal components exceed more than 90% of the total variability for both of the classes. Which is then used for building the sub model. All PCA sub model were build using NIPALS algorithm in order to classify large datat set of test samples. The Individual model of two classes were assembled by SIMCA.

Table 3.1: The percentage of variability covered by the PCA Kerosene blended class.

Principal Component number	% of Variation in PCA	% of Total Variation
1	77	77
2	14	91
3	5	96

Table 3.2: The percentage of variability covered by the PCA Nearby State Class.

Principal Component No.	% of variation in PCA	% of total Variation
1	90	90
2	5	95
3	4	99

Validation samples for both classes were created in order to validate the calibration model. A total of 7 validation samples, 4 samples for kersene blended class were prepared in the laboratory and 3 samples for, nearby state class were chosen. The classes of validation samples were predicted using calibration samples, which is summarized in Table 3.4 and 3.3. The calibrated model showed an excellent precission of classifation, as well as realiable specificity and sensitivity for classyfing the two classes of validation samples . However, the error rate of kereosene blended class is high, while the majority of gasoline properties are very similar to nearby state class. Despite the error rates, we also found that the calibration model had 100 % specificity and 75% sensitivity for both classes. It reveals that the model was capable of classify into two of the adopted classes.

Table 3.3: Classification result of SIMCA validation Set.

	No. of Samples	Classified as Kerosene blended class	Classified as Nearby State class
Kerosene Blended Class (KS)	4	4	4
Nearby State Class (NS)	3	0	3

Table 3.4: Confusion Matrix for SIMCA validation Set. True Positive Rate (TPR), False Positive rate (FPR), True Negative Rate (FNR), and False Negative Rate (FNR).

	No. of Samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
Kerosene Blend Class (KS)	4	1.00	0	1.00	1.00	1.00	0.50
Nearby State Class (NS)	3	1.00	0	1.00	0	1.00	1.00

The developed SIMCA model was applied to test samples obtained from the state of Mizoram. It was discovered that 108 samples were located inside the class of a nearby state, with no evidence of kerosene adulteration. However, 34 samples were found to be unclassified in both classes. It is suspected that other adulterants were present in those unassigned samples. The result of SIMCA classification is shown in Table 3.5.

Table 3.5: SIMCA classification result on commercial gasoline samples.

	Total No. of Samples	Classified as Kerosene Class	Classified as Nearby State Class	Unassigned Class
Commercial gasoline Samples	142	0	108	34

3.3.1.5 Partial Least Square Discriminant Analysis (PLS-DA)

To compare the PLS-DA with SIMCA model, another PLS regrssion model were developed. Two types of model were constructed for kerosene blended samples, 8 samples for calibration model ranging from 2-22 % v/v and another 4 samples for validation ranging from 6-24 %v/v. The four optimum latent variable (factor) were chosen for the model. It was able to explained 98% of the total variability.

In order to verify the calibration model, the validation data set was predicted by the calibration model. The perfomance of the model were shown Figure 3.4 and 3.5 , the predicted with deviation plot shows the validation samples that fall on 1 are predicted as Kerosene blended class. On the other hand, the samples that falls on -1 are predicted as Nearby State class. According to the predction result, the rate of error for kerosnene blended and Nearby State are found as 0 %. The information regarding the classification result are shown in table 6. It revealed that the two models has 100% Specificity and Sensitivity . It shows that the calibration model has an excellent specifity and sensititivty for classifying large data set of test samples. Table 3.6 show the comparison of the actual and predicted concentration of the known kerosene blended validation set. Which is also obtained at 97% confidence interval.

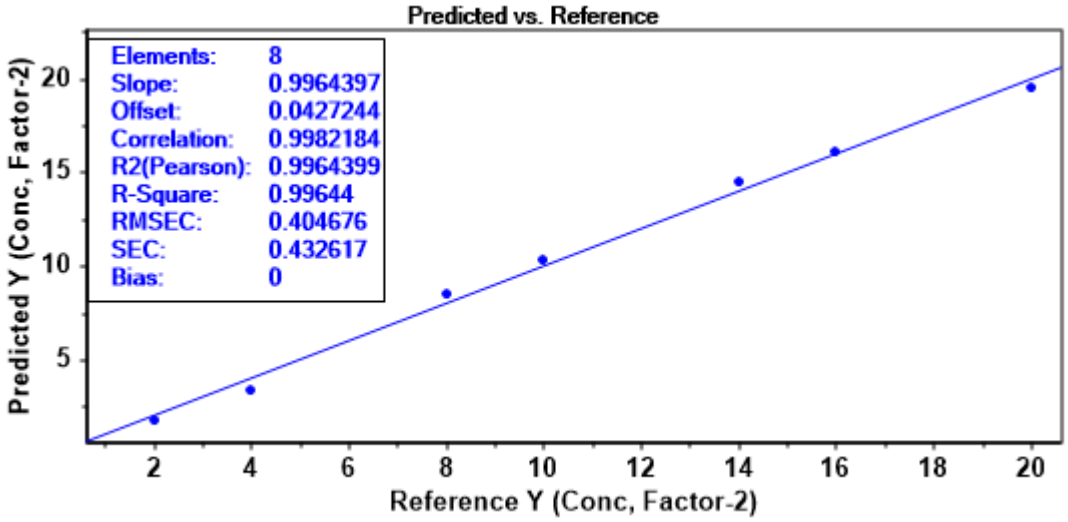

Figure 3.4: *The PLS-DA predicted vs reference plot of calibration model*

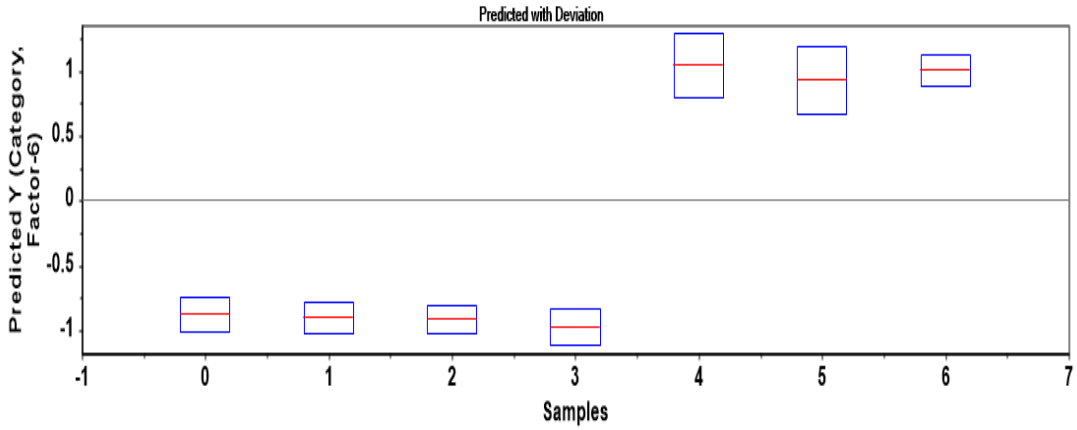

Figure 3.5: *The PLS-DA plot of Validation model predicted with deviation plot. (KS- Kerosene Blended and NS- Nearby State samples)*

Table 3.6: PLSDA classification result on validation Set. True Positive Rate (TPR), False Positive rate (FPR), True Negative Rate (FNR), and False Negative Rate (FNR).

	Number of samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
Kerosene blended Class (KS)	4	1.00	0	1.00	0	1.00	1.00
Nearby State Class (NS)	3	1.00	0	1.00	0	1.00	1.00

The classification of all the test samples were done by the PLS-DA model. Table 3.7 shows that the predicted result of 142 commecial samples or test samples. According to Table 8, it was clearly illustrated that 7 samples were predicted inside the group of kerosene blended samples. Another 84 samples are predicted inside the class of nearby state samples. Apart from this, 51 samples are unassigned to those two classes. The PLS-R calibration model with the high significant value of R-Square (0.996) and RMSEP (0.405) predicted concentration of the adulterted samples. Table 3.7 also showed the predicted result of kerosene concentration in adulterated samples, which is in a range of 0.3-13 % v/v. It is interesting that gasoline adulteration with kerosene in Brazil had the same maixum concentration of 13.4 % v /v(Teixeira et al. 2008). Therefore, we believed that the relatively small amount of kerosene is found to be in the average 6.78% v/v.

Table 3.7: PLSDA classification and PLS-R predicted results on commercial gasoline samples.

	Total number of samples	Classified as a kerosene blended class	Classified as a nearby state class	Unassigned Class	Predicted Concentration of Kerosene
Commercial gasoline Samples	142	7	84	51	0.3-13% v/v

The SIMCA and PLS-DA anlyses result shows that many of Mizoram samples are separated from Kerosene blended and nearby state class, while majority are still lie on nearby satate class. We believed that kerosene is not the only adulterant in Mizoram. It also shows that the supervised PLS-DA gives better classfication on Kerosene blended samples. Also, the information on samples classification is better on PLS-DA than PCA-SIMCA method.

3.3.2 Second Analysis

3.3.2.1 Principal Component Analysis.

PCA is an unsupervised technique that was used as an exploratory technique as well as for outlier discovery. The PCA results in Figure 3.6 shows the score plot, which explains 78% of the total variation. It shows that only the reformulated samples and a portion of the test samples can be differentiated from other samples. There are not any other clear distinctions between the samples. Therefore, the PCA result only demonstrates a slight degree of variation and class overlap rather than clearly demonstrating how the classes are varied.

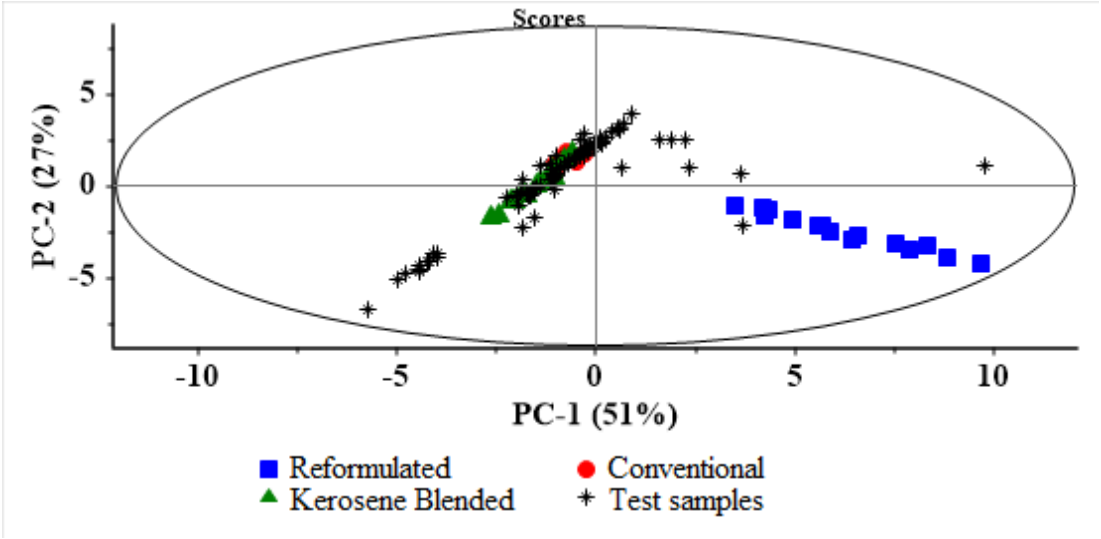

Figure 3.6: *PCA Score Plot of the first two components*

3.3.2.2 Classification of LDA Technique

The supervised technique of LDA is a well-known and useful pattern recognition approach that characterises a separation hyperplane by calculating linear discriminant functions, resulting in optimal class discrimination. In this work, the LDA in combination with PCA was obtained for all samples in the spectral region of 3200–650 cm^{-1}. Where the first two PCs almost explained 90% of the total variation for training samples. Which is then used for the classification of gasoline samples based on their quality evaluation. The training sets are then conducted with 28 samples (7 conventional, 10 reformulated, and 11 kerosene blended), and the remaining 14 samples are introduced to the validation set (4 conventional, 5 reformulated, and 5 kerosene blended) in order to predict 78 unknown samples. The training sets are used to predict the validation samples by using the LDA technique. The discrimination plot with 80% accuracy is projected on two PCs, as shown in Figure 3.7. The calibration

model has great specificity and sensitivity, as shown in the classification results of validation samples given in Tables 3.8 and 3.9. It has been proven that it is suitable for classifying a large data set of commercial gasoline samples. The classification of gasoline samples was done by training sets of LDA, which are shown in Table 3.10. The result shows that out of 78 test samples, 30 samples from conventional and 6 samples from reformulated samples are assigned to the normal class. An additional 3 samples from the reformulated are placed in the oxygenate class. In addition, 37 conventional fuel samples and 2 reformulated fuel samples were classified as contaminated with kerosene.

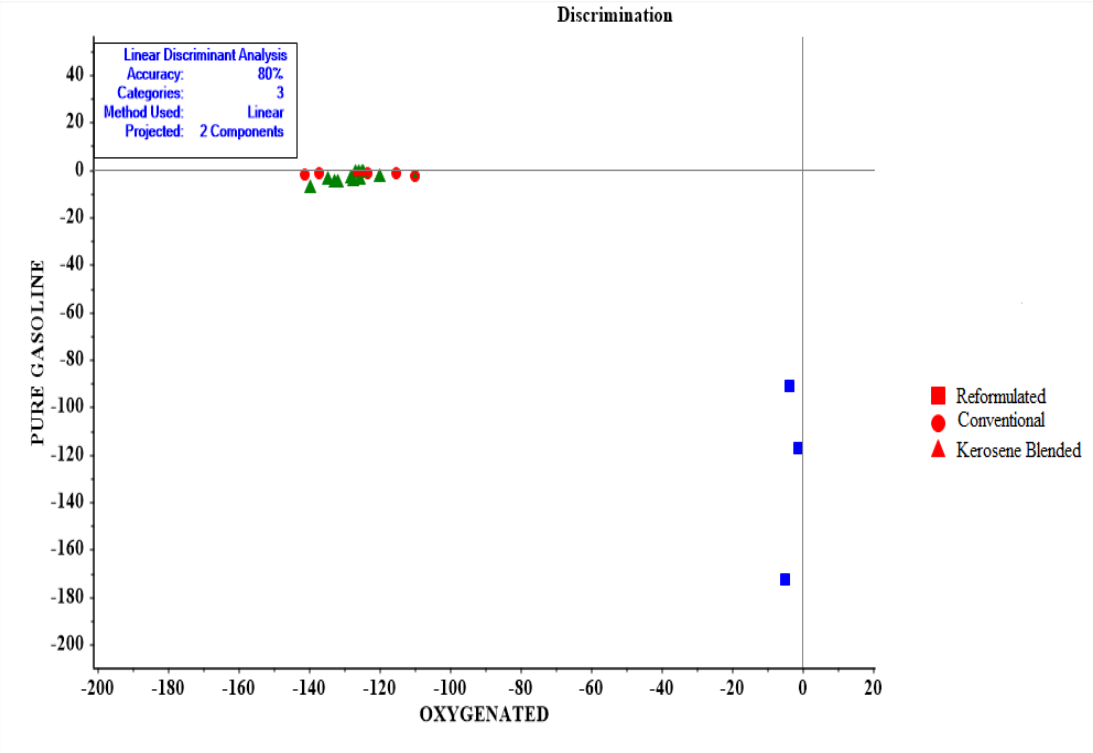

Figure 3.7: *The LDA Discrimination plot of Training Set*

Table 3.8: The LDA predicted result of validation sets.

	No of Samples	Classified as Normal Class	Classified as Oxygenate	Classified as Kerosene Blended Class
Conventional	2	1	0	1
Reformulated	2	0	2	0
Kerosene Blended	4	1	0	3

Table 3.9: The statistical parameters for LDA training set.

	No of Samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
conventional	2	0.50	0	1.00	0.50	0.50	0.50
Reformulated	2	1.00	0	1.00	0	1.00	1.00
Kerosene Blended	4	0.75	0	1.00	0.25	0.75	0.75

Table 3.10: The LDA classification result of commercial gasoline samples.

	Total No of samples	Classified as Normal Class	Classified as Oxygenate Class	Classified as kerosene contaminated
Conventional	67	30	0	37
	11	6	3	2

3.3.2.3 Classification of PLS-DA Technique

PLS-DA is one of the most recognised supervised classification algorithms in chemometrics. This method is based on partial least-squares regression of continuous predictor variables to find optimal latent variables with the highest covariance. The PLS-DA technique was also employed on the IR spectra of the same samples. In order to predict the unknown samples, the same training and validation sets were also conducted for the analysis. The performance of the training model is shown in Figures 3.8 (a) and (b). The predicted with deviation plot shows that the validation samples that fall on 1 are predicted as oxygenate or reformulate, samples falling on 0 are conventional gasoline, and -1 are kerosene blended samples. The classification results of validation samples are given in tables 3.11 and 3.12, with a 0% rate of error. It was found that both models had 100% Specificity and Sensitivity. Which proved that the calibration model is fit enough for data analysis. Therefore, all the commercial gasoline samples were analysed using the PLS-DA training model. From the classification result of 78 commercial gasoline samples, as shown in Table 3.13, it became clear that 54 conventional gasoline samples and 6 reformulated gasoline samples are classified in the normal class group. Another 4 reformulated samples are categorized as oxygenate class. However, 13 samples from conventional fuel and 1 sample from reformulated fuel are identified as adulterated with kerosene.

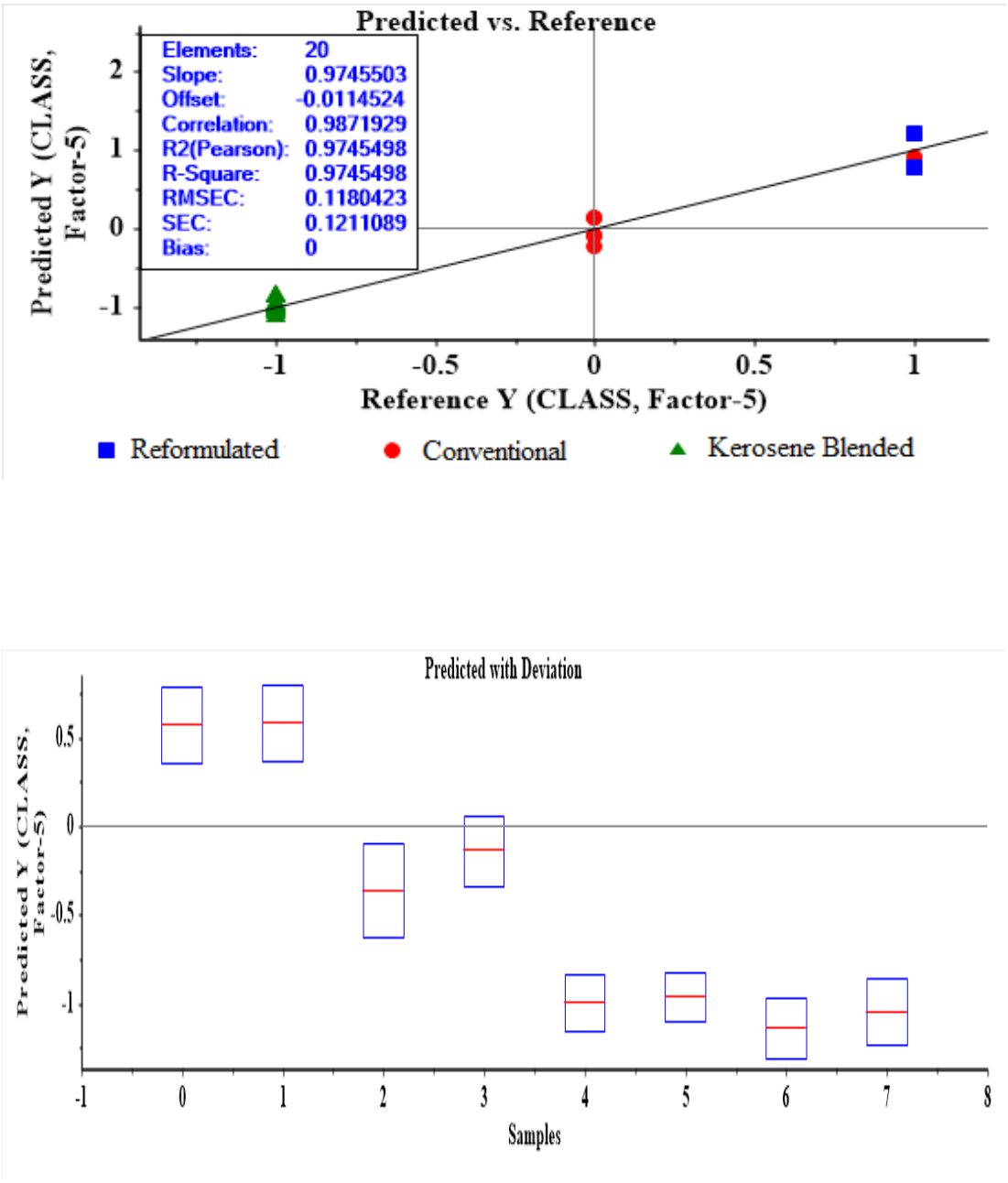

Figure 3.8: *The PLS-DA (a) predicted vs reference plot, (b) predicted vs deviation plot.*

Table 3.11: The predicted result of PLSDA validation set.

	Total No of samples	Classified as Normal Class	Classified as Oxygenate Class	Classified as kerosene contaminated Class
Conventional	67	54	0	13
Reformulated	11	6	4	1

Table 3.12: The statistical parameters of PLS-DA Training set.

	No of Samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
Conventional	2	1.00	0	1.00	0	1.00	1.00
Reformulated	2	1.00	0	1.00	0	1.00	1.00
Kerosene Blended	4	1.00	0	1.00	0	1.00	1.00

Table 3.13: The PLS-DA classification result of commercial gasoline samples.

	No of Samples	Classified as Normal Class	Classified as Oxygenate Class	Predicted as Kerosene Blended Class
Conventional	2	2	0	0
Reformulated	2	0	2	0
Kerosene Blended	4	0	0	4

3.3.3 Third Analysis

To detect adulteration, 176 petrol samples from three companies—IOC, HPC, and Bharat BPC—were analysed for quality control. Gasoline samples were adulterated with kerosene at concentrations ranging from 5% to 20% v/v to create standard sets. The samples were then divided into two groups: a 20% validation set to assess the model's performance, and an 80% calibration set to develop the analytical model.

3.3.3.1 Classification of SVM, LDA and PLS-DA Techniques

In this study, classification analysis was performed on all samples within the spectral range of 3200–650 cm^{-1}. The dataset was divided into a training set and a validation set. The training set consisted of 14 samples, while the validation set contained 5 kerosene samples, which were used to predict the 176 unknown samples. The training set was employed to predict/classify the validation samples, and the performance of the training set on the validation set is summarized in Table 3.14. This will likely involve metrics such as accuracy, precision, and other key indicators to evaluate the classifier's effectiveness on both new and unseen data. Table 3.15 illustrates that the specific parameters used in the training sets make them appropriate for classifying a large collection of commercial gasoline samples. This highlights that the classification models are dependable and can be applied to a wider range of sample types.

Table 3.14: validation sets of kerosene blended samples for SVMC, PCA-LDA and PLS-DA.

	No of Validation Samples	Classified as Kerosene Blended Class
SVMC	10	5
LDA	10	5
PLS-DA	10	5

Table 3.15: The statistical parameters for SVMC, PCA-LDA and PLS-DA training sets.

	No of Samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
Kerosene Blended samples	**10**	**1.00**	**0**	**1.00**	**0**	**1.00**	**1.00**

Based on the classification results, the SVMC and PCA-LDA techniques identified 54 samples as being contaminated with kerosene, while PLS-DA detected 32 samples as adulterated with gasoline. The classification outcomes were further predicted using PLSR, with the overall predicted results presented in the Table 3.16. The kerosene concentration was found to range from 0.5% to 29.3% v/v. Additionally, the average predicted results are illustrated in a bar chart, as shown in the Figure 3.9**.**

Table 3.16 Predicted Result on PLS-DA Classification

Technique	Kerosene Concentration
SVMC	0.6 to 29.3 % v/v
PCA-LDA	0.6 to 24.7 % v/v
PLS-DA	0.4 to 29.3 % v/v

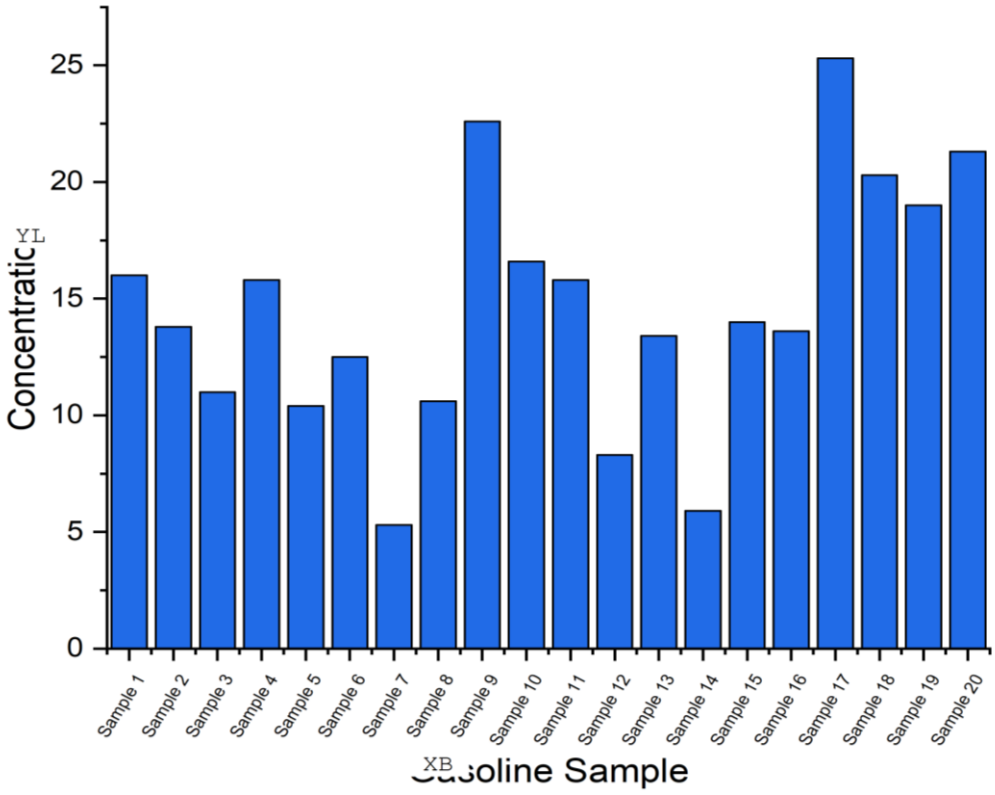

Figure 3.9: *Average predicted result of kerosene level in different gasoline stations*

3.4 Conclusion

In this chapter, the combination FTIR-ATR with multivariate analysis is used for classification and identification of gasoline adulteration with kerosene. In order to generate the qualitative relationships for each dataset, the methods of PCA, SIMCA, PLSR and PLS-DA analyses were examined. The PCA and SIMCA approach had a poor selectivity, it does not achieve on detection of adulterant samples. The PLS-DA model reveals its benefit for classification based on their specificity, sensitivity, and class distinction. According to the PLS-DA model results, 7 samples from Mizoram are found as contaminated with kerosene. The PLS regression model was predicted the considered adulterants with concentration range of 0.3-13% v/v with a good fitting value of R-Square and RMSE. In southeast Asia (including India), the quantity of blended kerosene in gasoline compared to diesel fuel was simply stated to be relatively small amount without knowing the average concentration. Therefore, we believed that the relatively small amount of kerosene is found to be in the average 6.78% v/v. Hence, FTIR combined with multivariate analysis provided reliable, quick classification and measurement of kerosene concentration in gasoline fuel.

In second analysis, the vibrational spectra of FTIR combined with chemometric tools were used to investigate Mizoram gasoline samples. The spectra of gasoline samples were analysed by supervised PCA-LDA and PLS-DA techniques. The analysis was carried out in two successive phases: First, classification algorithms were built with the three datasets as a training set; then, the performance of the constructed models was evaluated by analysing the classifications performed on the validation set. The obtained results were evaluated according to the values of

sensitivity and specificity in the classification of the spectra of the three datasets. According to the expected validation result in classification, PLS-DA has 100% accuracy for the training set. The LDA training set, on the other hand, exhibits less error. It indicates that the performance of the PLS-DA model is more sensitive than that of the LDA model. However, the PLS-R model with a good fitting value of R-Square and RMSE predicted result proved that the classification performed by LDA on gasoline samples is more accurate than PLS-DA, particularly in oxygenate or reformulate samples.

In the third analysis, a quality check was performed to detect adulteration across all samples. The classification analyses using SVMC, PCA-LDA, and PLS-DA showed identical results for SVMC and PCA-LDA. Consequently, all classified samples were predicted using PLSR. The results revealed that kerosene as adulterants was observed over the course of three years of study. However, the occurrence of kerosene adulteration was minimal throughout the research, likely because kerosene had been removed from the public distribution system and its price had become nearly equivalent to that of gasoline.

CHAPTER 4

DETERMINATION OF AROMATIC HYDROCARBONS AND OXYGENATE IN GASOLINE FUEL.

In this chapter, the classification and quantification of aromatic hydrocarbons and oxygenates in gasoline samples are carried out to evaluate fuel quality. Aromatic hydrocarbons, such as benzene, toluene, and xylene, are important indicators of fuel composition, affecting combustion efficiency, engine performance, and emissions. Oxygenates, like ethanol and MTBE, are added to gasoline to improve its combustion and reduce harmful emissions. By analysing the concentration of these compounds in various gasoline samples, the study aims to determine their levels and assess the impact of these components on the overall quality of the fuel. Higher concentrations of aromatic hydrocarbons, for example, can indicate lower fuel quality and may lead to increased pollutant emissions, while oxygenates can enhance combustion properties but, in excess, may alter engine behaviour.

To achieve this, FTIR spectroscopy with chemometric tools of PCA, SVM, PLSDA, SVMR, PLSR and PCR are employed to classify and quantify the aromatic hydrocarbons and oxygenates present in the gasoline samples. These techniques enable the identification of patterns in the data, facilitating the classification of gasoline samples based on their chemical composition. By quantitatively determining the concentration of each compound, the study provides insights into the fuel's quality, revealing how variations in aromatic and oxygenate content correlate with fuel performance and emissions.

4.1 Introduction

The quality of gasoline products is largely determined by the blending of various refinery streams, which directly influences engine performance. The precise formulation of gasoline ensures optimal combustion efficiency, engine smoothness, and fuel economy. Traditionally, this blending process takes into account a range of factors like volatility, octane rating, and sulfur content. However, in recent years, the compositional analysis of gasoline, particularly the levels of Benzene and total aromatics, has become increasingly important. These components are crucial because they can significantly affect both engine performance and environmental emissions. Benzene, a toxic compound, is regulated due to its carcinogenic properties, while the total aromatic content affects the fuel's volatility, combustion characteristics, and pollutant formation during combustion.

To standardize fuel quality and ensure compliance with environmental and performance requirements, different regions around the world have established specific gasoline specifications. In Europe, the EN-228 standard outlines the permissible limits for aromatic content and other properties. Similarly, the ASTM D4814 specification in the USA defines the requirements for gasoline used in spark-ignition engines. Japan follows the JIS K 2202 standard, and India has the IS 17943 (2022) standard, which also incorporates limits for aromatic hydrocarbons and benzene content. These specifications help to ensure that gasoline meets both performance criteria and environmental standards, reducing harmful emissions while maintaining fuel efficiency and engine longevity across different regions (Singh., 2003).

In India, Benzene, a common aromatic compound added to gasoline, is restricted in its concentration due to its carcinogenic properties. Under the current regulations in India, the maximum permissible limit for benzene in petrol (gasoline) is set at 1% by volume. This limit is enforced as part of the Bharat Stage (BS) emission norms, which are designed to regulate the environmental impact of vehicle emissions. Specifically, the benzene limit is applicable under the BS III, BS IV, and now BS VI standards, with each successive stage representing stricter emission requirements aimed at reducing air pollution and improving air quality across the country (Kumari *et al.,* 2023). This regulation helps ensure that gasoline remains within safety standards before it is marketed (Bonfim et al., 2012).

To produce different grades of gasoline, oxygenates are added to improve the octane rating and reduce air pollution. MTBE (Methyl Tert-Butyl Ether), an oxygenate additive in gasoline, serves dual purposes in improving fuel performance. It enhances the octane number, which boosts engine efficiency by preventing knocking during combustion. Additionally, MTBE helps to reduce exhaust emissions by promoting more complete combustion of the fuel, leading to lower levels of harmful pollutants such as carbon monoxide and hydrocarbons. This makes MTBE a valuable additive for meeting stricter environmental standards while improving overall fuel quality. However, its use is strictly regulated in the country due to environmental concerns, particularly its potential to contaminate groundwater (Prasad *et al.,* 2008). The allowable concentrations of aromatics and oxygenates in gasoline can vary by region and are subject to regulatory standards. Different areas or countries may have distinct

standards and testing methods for assessing gasoline quality, which can complicate the identification and measurement of these components.

This chapter assesses the effectiveness of multivariate techniques for addressing key issues in gasoline screening through FTIR spectroscopy. Various multivariate analysis methods, including Principal Component Analysis (PCA), Support Vector Machine Classification (SVMC), Linear Discriminant Analysis (LDA), and Partial Least Squares Discriminant Analysis (PLSDA), were utilized for quality-based classification. Additionally, Partial Least Squares Regression (PLSR), Support Vector Regression (SVR), and Principal Component Regression (PCR) were employed to quantify and identify aromatic hydrocarbons, oxygenates and kerosene adulterants in commercial gasoline.

4.2 Material and Methods

A total of 176 gasoline samples were collected from filling stations of three companies—IOC, HPC, and Bharat BPC—across various locations in Mizoram, a northeastern state in India. To ensure diverse sample compositions, these samples were gathered from different distributors and producers. Additionally, gasoline samples from the Indian Oil Corporation Ltd's Marketing Division and 11 samples from a nearby state were used as a training set for comparison. To measure the oxygenate content, 10 standard solutions were prepared by mixing conventional fuel with MTBE at concentrations of 1-10% v/v.

The solution of a five-compound mixture, consisting of Isooctane, Benzene, Toluene, and pure Xylene, was analysed in accordance with the ASTM D6277 standard. ASTM D6277 is a widely recognized method for the determination of

aromatics in gasoline using FTIR Spectroscopy. This standard provides a precise and reliable technique for analysing the concentration of aromatic compounds in fuel, such as benzene, toluene, and xylene, among others. By applying this method, the composition of the mixture was accurately determined, allowing for the quantification of individual aromatic hydrocarbons, which are critical for assessing fuel quality, engine performance, and environmental impact. In order to determine Benzene concentration, 25 standard solutions were prepared with concentrations ranging from 0.5-1.5% v/v, following ASTM D6277 standards.

The prepared solutions were then divided into two sets: a calibration set to develop the analytical model and a validation set to test the performance of this model. The summaries of calibration and validation sets were shown in Table 4.1.

TABLE 4.1: ***Summary of the calibration and validation sets of Benzene and MTBE***

	Concentration Range (V/v %)	**No. of Standard**
Calibration Standard Set of Benzene	0.5 – 1.5	20
Validation Standard Set of Benzene	0.5 – 1.5	5
Calibration Standard Set of MTBE	1 – 10	10
Validation Standard Set of MTBE	1 – 10	5

4.3 Result and Discussion

4.3.1 IR Spectra of gasoline

Figure 4.1 represent the original spectra of MTBE blended and Conventional gasoline samples in the region of 3200–650 cm^{-1}. It shows the original mean MIR spectra of the two classes of gasoline residue recorded in the range of 4000–650 cm^{-1}. It is possible to observe peaks corresponding to CH stretching of alkanes and alkenes (2800–3100 cm-^{1}), CH stretching of aromatic rings (1475–1600 cm1), CH bend of alkanes (1300–1500 cm^{-1}) and CH out-of-plane bend of alkenes and aromatic rings (650–1000 cm^{-1}). The broad absorption in the region of 1200-1000 cm^{-1} is a characteristic of functional group O-H stretching vibrations. So, the absorption peaks below 1000 cm^{-1} are mostly due to out-of-plane vibrations of aromatic and other cyclic hydrocarbons.

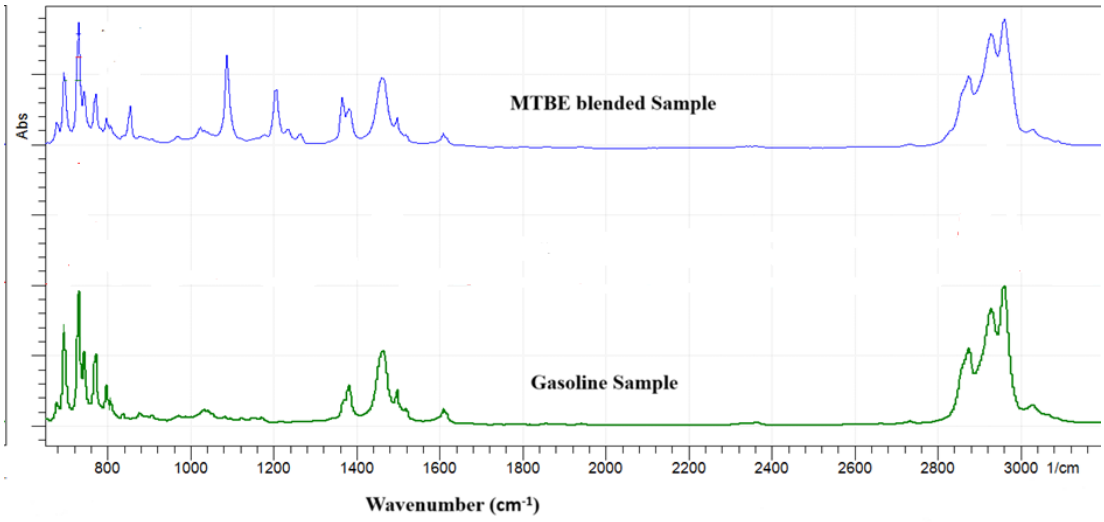

Figure 4.1: *IR Spectra of MTBE blended Sample and Conventional Gasoline sample at 650-3200 cm^{-1}.*

In the gasoline spectrum, the CO (carbonyl) mode is primarily characterized by absorption peaks at 1085 cm^{-1} and 1205 cm^{-1}. These peaks are mostly attributed to the presence of oxygenates in gasoline compounds, such as ethanol or MTBE (Methyl Tert-Butyl Ether), which are commonly used to improve combustion and reduce emissions. The oxygen-containing compounds in gasoline contribute to these distinctive absorption bands in the infrared (IR) spectrum.

Additionally, the C-H bending vibrations of aromatic compounds, such as benzene, are responsible for an absorption peak around 675 cm^{-1} in the IR spectrum. This absorption is specific to the aromatic ring structure, and its presence is used to identify and quantify aromatic hydrocarbons, such as benzene, toluene, and xylene, within gasoline. The distinct IR absorption features of these compounds allow for their accurate identification and analysis in fuel mixtures.

4.3.2 PCA Analysis

PCA is an unsupervised method that has been used for both outlier detection and exploration. The PCA results in Figure 4.2 shows the score plot, which explains 71% of the total variation. It only shows samples differences and similarities based on the supplier companies and identified 3 samples as outlier out of a total 176. Other than that, there are no obvious between the differences of the samples. Thus, rather of distinctly showing how the classes are varied, the PCA result only shows a little amount of variation and sample overlap. PCA is primarily used for reducing the dimensionality of the data, which can make is useful for visualizing high dimensional data and identifying patterns or grouping. However, it may not always be effective at

distinguishing between distinct classes or categories. Thus, PCA can be a valuable tool, but to classify or categorize the data other method like supervised technique of multivariate data analyses might be appropriate.

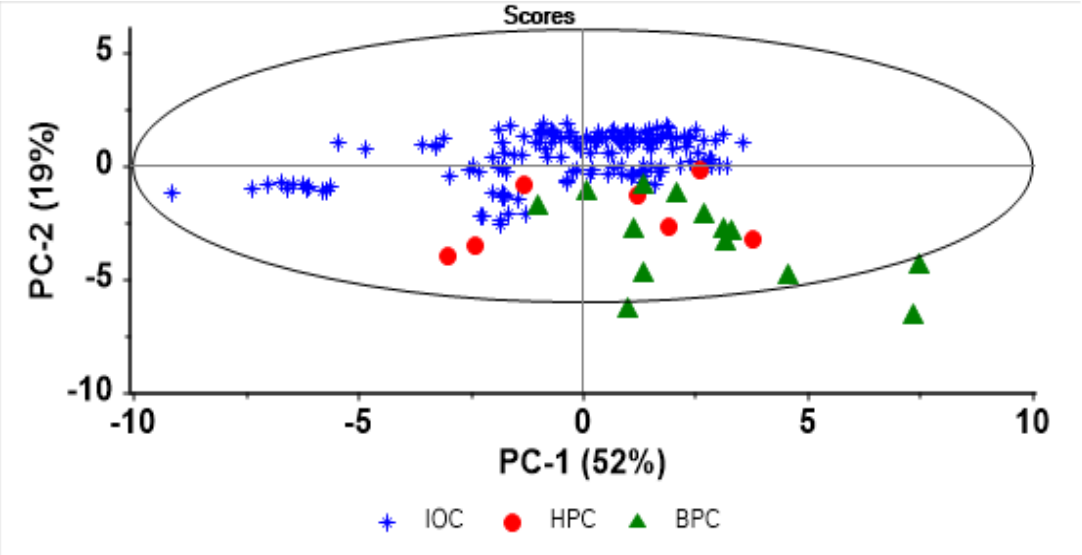

Figure 4.2: *PCA Score plot of all commercial gasoline Samples.*

4.3.3 Classification of SVM, LDA and PLS-DA Techniques

The supervised techniques of SVMC, PCA-LDA and PLS-DA are commonly used in chemometrics and data analysis for classification purposes, particularly in the context of identifying and classifying the samples. In this work, the classification analysis was obtained for all samples in the spectral region of 3200–650 cm^{-1}. The dataset was divided into training set and validation set. The training sets consist of 11oxygenated samples, while the validation set contained the remaining 5 oxygenated samples. In order to predict 176 unknown samples. The graphical plot of training sets in all three techniques is shown in Figure 3. Figure 4.3(a) represents the SVMC classification plot for the training set, with 100% training accuracy. SVMC is a

powerful classification algorithm used in machine learning. Figure 4.3(b) display the discrimination plot for PCA-LDA with 100% accuracy achieved using two principal components. PCA is used for dimensional reduction, while LDA is used for classification. Figure 4.3(c) shows the predicted vs reference plot of PLS-DA training sets with an R-Squared value of 0.99 and RMSEC value of 0.09. PLS-DA is often used for analysing spectral data and classifying samples. So that, the training set is used to predict/classify the validation samples, where the performance of training sets over validation sets are shown in table 4.2. This likely include metrices such as accuracy, precision, recall and other relevant measures to evaluate the classifier's performance on new, unseen data. Table 4.3 indicates that the specific parameters used in the training sets make them suitable for classifying a large data set of commercial gasoline samples. This suggests that the classification models are robust and may be applicable to a broader range of samples.

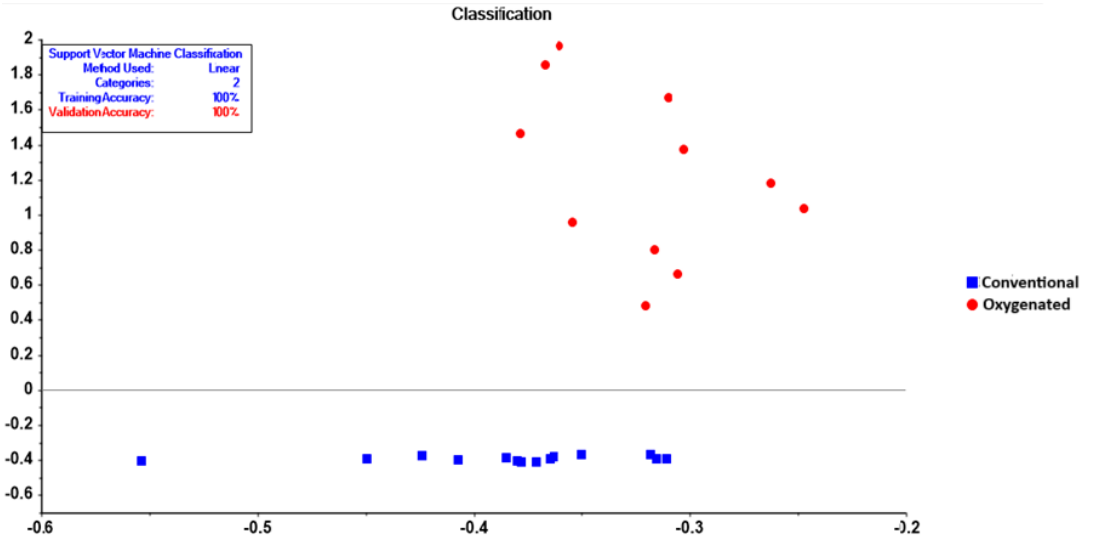

Figure 4.3(a): *Classification plot of SVMC Training Sets*.

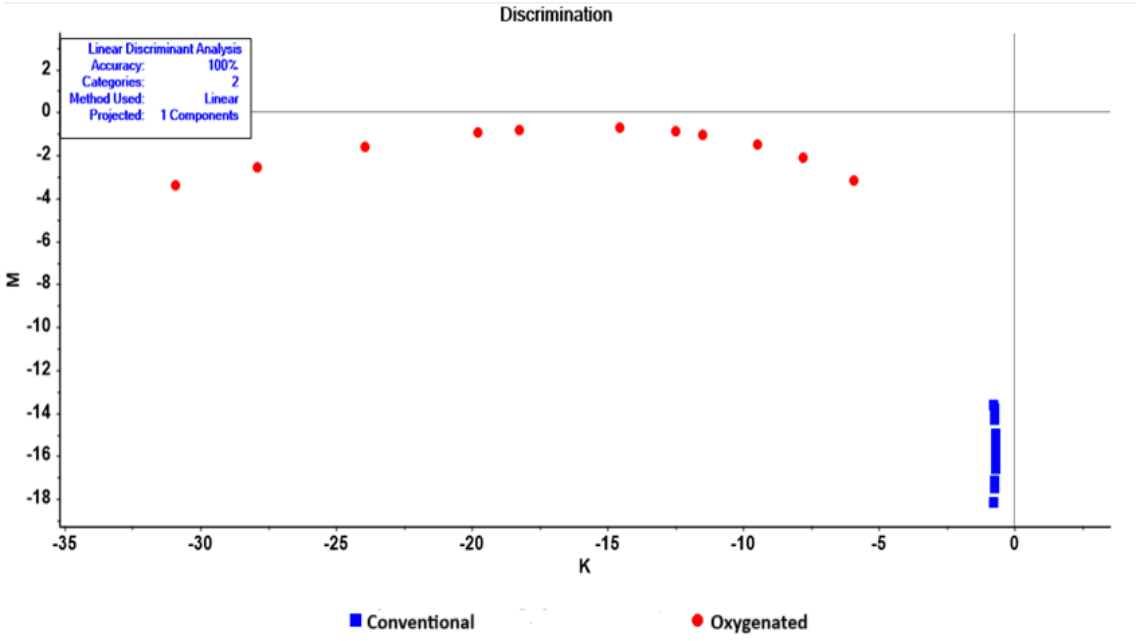

Figure 4.3(b): *Discrimination Plot of PCA-LDA Training Sets*

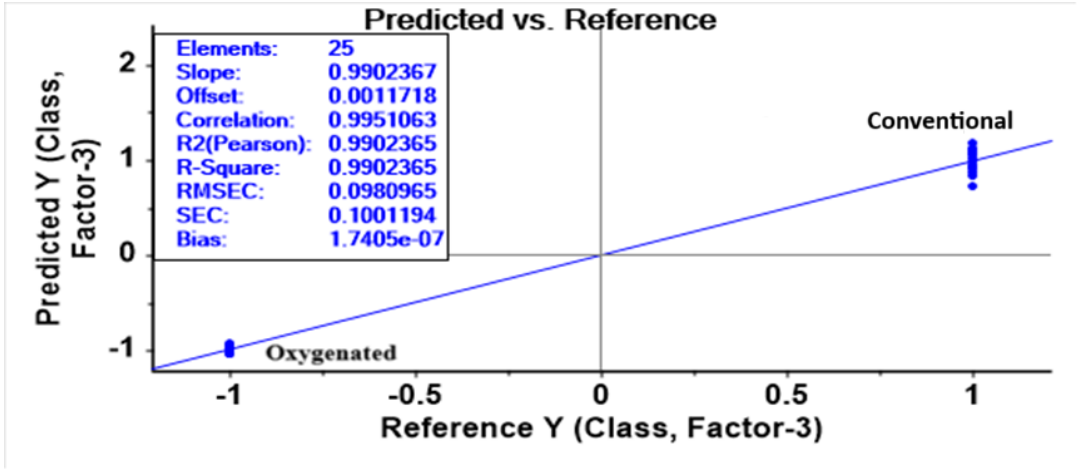

Figure 4.3(c): *Predicted vs Reference Plot of PLS-DA Training Sets.*

Table 4.2: The classification result of validation sets for SVMC, PCA-LDA and PLS-DA.

	No of Validation Samples	Classified as Oxygenate class
SVM	10	5
LDA	10	5
PLS-DA	10	5

Table 4.3: The statistical parameters for SVMC, PCA-LDA and PLS-DA training sets.

	No of Samples	TPR	FPR	TNR	FNR	Specificity	Sensitivity
Reformulated	10	1.00	0	1.00	0	1.00	1.00

The classification of gasoline samples using training sets were done by SVMC, PCA-LDA and PLS-DA, which are shown in Table 4.4, 4.5 and 4.6. According to the results in table 4.4, 5 samples of 14 BPC samples were found to be oxygenated, whereas the remaining 9 samples were found to be non-oxygenated. Furthermore, just one HPC sample was discovered to be oxygenated, while the remaining six samples were determined to be non-oxygenated. Regardless, the results in Table 4.5 demonstrate that all of the IOC test samples have been assigned as non-oxygenated class. In regard to HPC, two samples were discovered as oxygenated out of 7 collected, all of which were thought to be oxygenated and 5 samples from BPC are determined

as oxygenated, despite the fact that all samples are assumed to be oxygenated with MTBE. According to the classification Table 4.6, Out of seven HPC samples, only 2 samples are discovered to be oxygenated. Furthermore, 3 of BPC samples were also discovered as oxygenated.

Table 4.4: Classification result of SVMC

Name of Supplier Company	Total No. of Samples	Non-Oxygenated	Oxygenated
IOC	155	155	0
HPC	7	6	1
BPC	14	9	5

Table 4.5: Classification result of PCA-LDA.

Name of Supplier Company	Total No. of Samples	Non-Oxygenated	Oxygenated
IOC	155	155	0
HPC	7	6	2
BPC	14	9	5

Table 4.6: Classification Result of PLS-DA.

Name of Supplier Company	Total No. of Samples	Non-Oxygenated	Oxygenated
IOC	155	155	0
HPC	7	5	2
BPC	14	6	3

Moreover, the work done by El Orche *et al.*, in classification of raw milk using FTIR in combination with PLS-DA, PCA-LDA and SVM. The acceptable and accurate results were provided by PCA-LDA and PCA-SVM method than PLS-DA (El Orche *et al.,* 2021). Although, LS-SVM and PLS-DA were the effective method than PCA-LDA, PLS-LDA and SIMCA for distinguishing the paper relic types using ATR-FTIR Spectroscopy in the work done by Jingjing Xia *et al.*, (J. Xia *et al.,* 2019). Apart from this, Da Silva *et al.*, showed that the best result in the mid-IR region was achieved in LDA with 100% accuracy in test samples between LDA and PLS-DA (Da Silva *et al.,* 2014). It also said that the SVM constructed with NIR spectra showed better classification in full region. According to the classification results observed in this work in the mid-IR range of 650-3200 cm^{-1}, SVMC and PCA-LDA classification were similar in oxygenated samples. This means that when classifying the samples related to oxygenation using SVMC and PCA-LDA, the two methods produced comparable results. It suggests that either method could be used effectively for this specific classification task. In this study, it appears that SVMC, PCA-LDA and PLS-DA were evaluated and their performance was found similar in certain classification task.

4.3.4 Prediction of SVMR, PLSR and PCR

To explore the differences in Benzene and MTBE concentrations based on SVMC, PCA-LDA, and PLS-DA classification results. As a consequence of this, the SVMR, PLSR, and PCR regression models for benzene and MTBE blended samples were created. The high significant value achieved for R-Square and RMSEC, as shown in Table 4.7, shows calibration's predictive ability. On validation sets, it is then used

to forecast the concentration of three types of gasoline blended components. According to the validation results in Table 4.8, SVMR provided superior accuracy in Benzene prediction. In the case of MTBE, all three approaches perform equally well in terms of prediction analysis. It also clearly demonstrates that the calibration models used are reliable for predicting the concentration of blended components in gasoline. Consequently, the previously mentioned results have been used to predict the amount of MTBE in commercial gasoline samples categorised by SVMC, PCA-LDA, and PLS-DA. Other than this, the concentration of benzene had also been predicted for each sample.

Table 4.7: Statistical parameters obtained by SVMR, PLSR and PCR.

		R^2 calibration	R^2 cross validation	RMSEC	RMSECV
Benzene	SVMR	0.96	0.85	0.03	0.08
	PLSR	0.91	0.89	0.12	0.14
	PCR	0.92	0.89	0.12	0.14
MTBE	SVMR	0.99	0.99	0.05	0.15
	PLSR	0.99	0.98	0.12	0.26
	PCR	0.99	0.98	0.01	0.25

Table 4.8: Statistical Result on Validation Sets by SVMR, PLSR and PCR.

	Sample No.	Actual value	Predicted value		
			SVMR	PLSR	PCR
Benzene	1	1.2	1.2	1.2	1.2
	2	1.5	1.4	1.4	1.3
	3	0.75	0.87	0.94	0.94
	4	1	0.95	0.95	0.97
MTBE	1	4	5.6	5.6	5.6
	2	6	5.9	5.9	5.8
	3	10	9.7	9.7	9.7
	4	12	12.1	12.1	12.1
	5		8.1	8.1	8.1

Based to the predicted results on Tables 4.9, 4.10 and 4.11 the classification done by SVMC and PCA-LDA in MTBE was found to be in the range of 2.8 - 11.9% v/v, but the classification done by PLS-DA was found to be in the range of -2.5 - 8% v/v. In addition, the quantity of Benzene concentration is predicted for all samples, which are discovered in the range of -0.1 - 0.25% v/v, which does not exceed the Government's limit.

Table 4.9: Predicted Result on SVMC Classification.

	MTBE	Benzene
SVMR	3 to 11.4 % v/v	-0.03 to 0.25 % v/v
PLSR	2.8 to 11.6 % v/v	-0.1 to 0.13 % v/v
PCR	2.8 to 11.9 % v/v	-0.08 to 0.18 % v/v

Table 4.10: Predicted Result on LDA Classification.

	MTBE	Benzene
SVMR	3 to 11.4 % v/v	-0.03 to 0.25 % v/v
PLSR	2.8 to 11.6 % v/v	-0.1 to 0.13 % v/v
PCR	2.8 to 11.9 % v/v	-0.08 to 0.18 % v/v

Table 4.11: Predicted Result on PLS-DA Classification.

	MTBE	Benzene
SVMR	-2 to 7.7 % v/v	-0.03 to 0.25 % v/v
PLSR	-2.5 to 8 % v/v	-0.1 to 0.13 % v/v
PCR	-2.5 to 8 % v/v	-0.08 to 0.18 % v/v

Tables 4.9, 4.10, and 4.11 displayed that the quantity of MTBE concentrations is comparable on the classification done by SVMC and PCA-LDA, but PLS-DA indicates substantial differences. As a consequence, the classification and predicted concentration results reveal that the SVMC classification approach superior to the PCA-LDA and PLS-DA methods. In addition, the PCA-LDA performed better than the PLS-DA. In terms of prediction approaches, PLSR and PCR are more comparable than SVMR. Although SVMR's prediction is also extremely near to PLSR and PCR. These findings also demonstrate the significance of SVMC and PCA-LDA as classification methods. As demonstrated in the results, using the regression approaches of SVMR and PLSR significantly improves classification performance. As a result,

the PLSR method was found to be the most reliable of the three quantification procedures.

The average predicted concentrations of Benzene and MTBE in all samples, determined through PLSR, are shown in Figures 4.4 and 4.5. Each sample corresponds to a different gasoline filling station. Figure 4.4 shows the average benzene concentration across various stations, with the highest average concentration being around 0.10 v/v and the lowest at 0.01v/v. This indicates that the benzene levels detected remain within the limits set by current government standards.

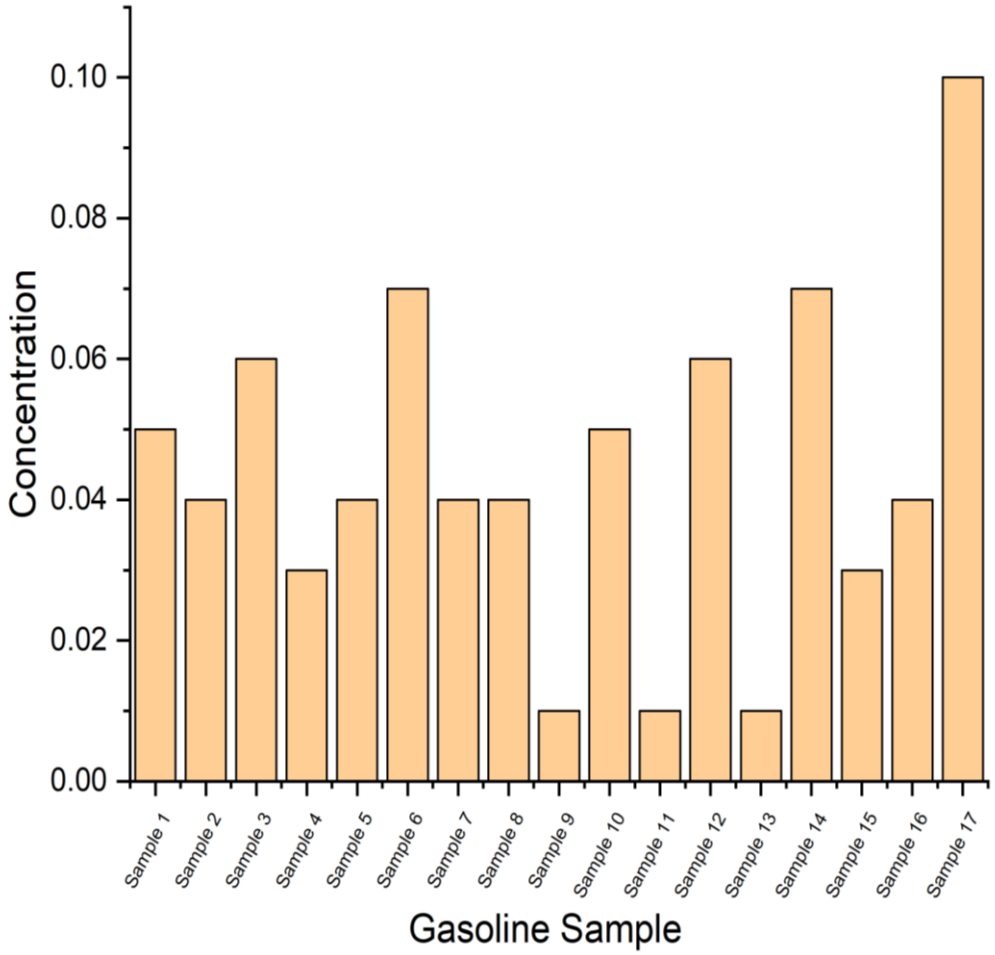

Figure 4.4: *Average predicted result of Benzene concentration in gasoline stations*.

Figure 4.5 shows the average MTBE concentration from seven different gasoline stations. As the graph indicates, the bar for sample 1 is zero, signifying that it does not contain an oxygenator; this sample was obtained from Indian Oil Corporation Ltd. With the exception of sample 1, the remaining six stations used oxygenators in gasoline samples, which is supplied by Hindustan Petroleum Corporation Ltd and Bharat Petroleum Corporation Ltd. The graph clearly demonstrates that the highest detected concentration of oxygenator in the state remains within the government's regulatory limits, even though the average maximum concentration is 10 v/v.

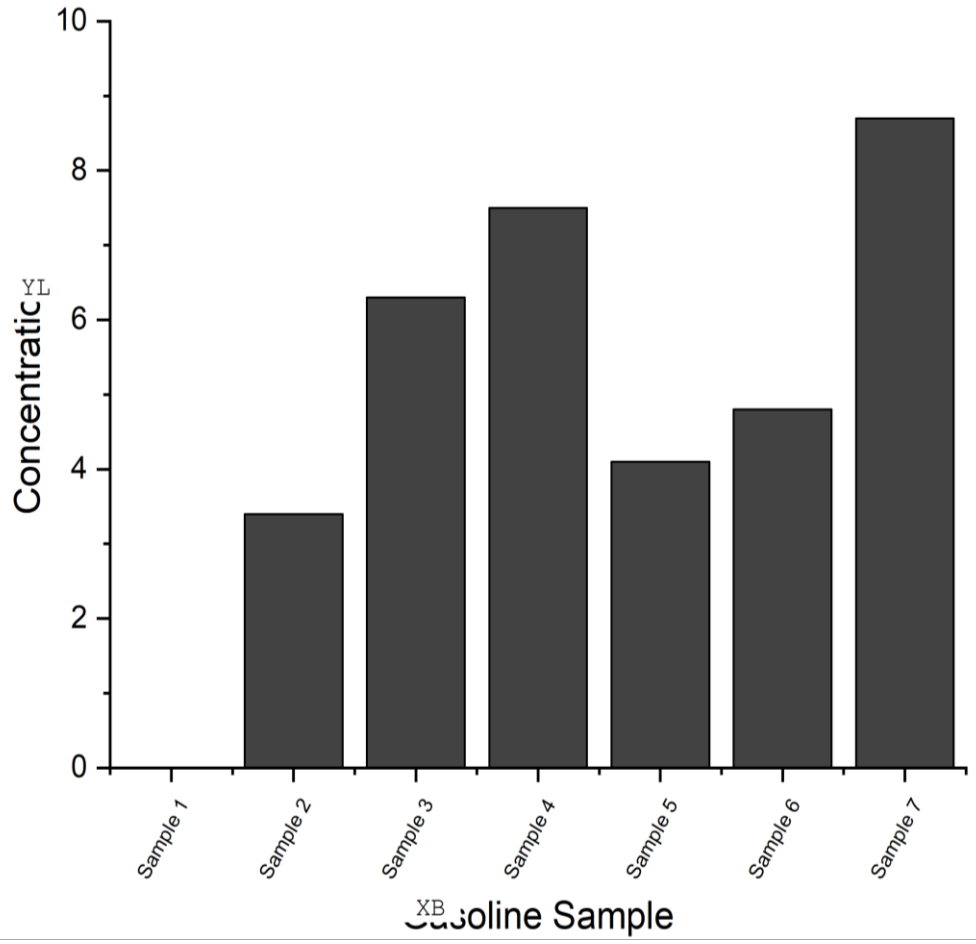

Figure 4.5: *Average predicted result of MTBE concentration in gasoline stations.*

4.4 Conclusion

This study focussed on the rapid identifying and classifying gasoline samples by using ATRFTIR spectroscopy and various chemometric techniques. Six different chemometrics methods namely, SVMC, PCA-LDA, PLS-DA, SVMR, PLSR and PCR were employed for both classification and quantification purposes. For classifying gasoline sample, SVMC and PCA-LDA were found to be effective methods, particularly in distinguishing the oxygenated gasoline samples. These methods showed comparatively satisfactory classification accuracy. In terms of quantifying gasoline samples, PLSR and PCR were found to be more comparable to each other, and these methods outperformed SVMR. However, it is noted that SVMR still had prediction that were very closed to PLSR and PCR. Among all the methods, the SVMC classification methods was highlighted as a good choice for classifying oxygenated samples, indicating its strong discrimination capabilities. These findings provide valuable insights into the use of chemometric techniques for analysing and classifying gasoline samples, which can have implications for quality control monitoring in the petroleum industry.

CHAPTER 5

SUMMARY AND FUTURE PROSPECT

5.1 Summary

This chapter discusses the use of FTIR-ATR combined with multivariate analysis to classify and identify gasoline adulteration with kerosene. Various methods, including Principal Component Analysis (PCA), Soft Independent Modelling of Class Analogy (SIMCA), Partial Least Squares Regression (PLSR), and Partial Least Squares Discriminant Analysis (PLS-DA), were evaluated for generating qualitative relationships within the datasets. The PCA and SIMCA methods showed limited selectivity and were ineffective in detecting adulterated samples. In contrast, the PLS-DA model proved effective for classification due to its high specificity, sensitivity, and ability to distinguish between classes. Using the PLS-DA model, 7 samples from Mizoram were identified as contaminated with kerosene. The PLS regression model accurately predicted the concentration of adulterants in the range of 0.3-13% v/v, demonstrating a good fit with high R-Square and low RMSE values.

In Southeast Asia, including India, kerosene is known to be blended in gasoline in relatively small amounts compared to diesel, with an average concentration estimated at 6.78% v/v. The study concludes that FTIR combined with multivariate analysis offers a reliable and rapid method for both classifying and measuring kerosene concentration in gasoline.

In this study, FTIR spectra combined with chemometric tools were used to analyse gasoline samples from Mizoram. The analysis employed supervised Principal

Component Analysis-Linear Discriminant Analysis (PCA-LDA) and Partial Least Squares Discriminant Analysis (PLS-DA) techniques. The study was conducted in two phases: first, classification models were developed using three datasets as training sets; then, the models' performance was assessed by evaluating their accuracy on a validation set. The results showed that the PLS-DA model achieved 100% accuracy on the training set, indicating its high effectiveness in classification. In comparison, the LDA model exhibited slightly more errors, suggesting that PLS-DA was more sensitive. However, the Partial Least Squares Regression (PLS-R) model, with strong R-Square and low RMSE values, demonstrated that the LDA classification was more accurate than PLS-DA, especially for distinguishing between oxygenate or reformulated samples.

However, in the analysis of rapid identification and classification of gasoline samples using ATR-FTIR spectroscopy combined with various chemometric techniques. Six chemometric methods—Support Vector Machine Classification (SVMC), Principal Component Analysis-Linear Discriminant Analysis (PCA-LDA), Partial Least Squares Discriminant Analysis (PLS-DA), Support Vector Machine Regression (SVMR), Partial Least Squares Regression (PLSR), and Principal Component Regression (PCR)—were employed for classification and quantification purposes.

For classifying gasoline samples, SVMC and PCA-LDA were particularly effective, especially in distinguishing oxygenated gasoline samples, and both demonstrated satisfactory classification accuracy. SVMC and PLS-DA performed similarly when classifying adulterated gasoline samples. In terms of quantification,

PLSR and PCR were found to be comparable and outperformed SVMR, though SVMR's predictions were close to those of PLSR and PCR. Among all the methods, SVMC stood out as a strong choice for both classifying oxygenated and adulterated samples due to its robust discrimination capabilities. These findings offer valuable insights into the use of chemometric techniques for analysing and classifying gasoline, which can enhance quality control in the petroleum industry.

The research indicates a variation in gasoline quality between samples obtained from Mizoram and those from a nearby state, suggesting that adulteration may be occurring within Mizoram. The classification analysis reveals this quality difference. Additionally, the study found that the levels of aromatic hydrocarbons and oxygenates in the gasoline do not exceed the country's permissible limits. Moreover, kerosene was not identified as the primary adulterant.

5.2 Future Prospects

- GC-MS is known for its reliability in classification and determination, although it is costly and time-consuming. As a more efficient and economical alternative, FTIR-ATR combined with chemometric tools is employed for classifying and quantifying gasoline in Mizoram. Of the five techniques tested, LDA demonstrated the highest accuracy in classification, making it a dependable choice for future research.
- In this study, only the presence of kerosene as an adulterant was investigated. The findings reveal that while kerosene is present in the samples, it is not the primary adulterant in the state. This may be because the government does not include kerosene in the public distribution

system. Future research could explore other potential adulterants to gain a more comprehensive understanding.

- This research focuses exclusively on the Aizawl area. Given the numerous new gasoline filling stations throughout the state, including in the city, it would be valuable to expand the study to cover all of Mizoram for a more wide-ranging analysis.
- The Government of India notified the National Policy on Biofuels on 4 June 2018. Under the Ethanol Blended Petrol (EBP) Programme, the policy set an indicative target of achieving 20% ethanol blending in petrol by 2030. The progress of this target should be regularly monitored, and the impact on fuel quality and engine performance should also be carefully evaluated.

REFERENCES

Abdi, Hervé. 2010. "Partial Least Squares Regression and Projection on Latent Structure Regression (PLS Regression)." *WIREs Computational Statistics* 2 (1): 97–106. https://doi.org/10.1002/wics.51.

Ali, Hendratta N. 2017. "Fundamentals of Petroleum Geology." In *Springer Handbook of Petroleum Technology*, edited by Chang Samuel Hsu and Paul R. Robinson, 321–57. Springer Handbooks. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-49347-3_9.

Balabin, Roman M., Ravilya Z. Safieva, and Ekaterina I. Lomakina. 2010. "Gasoline Classification Using near Infrared (NIR) Spectroscopy Data: Comparison of Multivariate Techniques." *Analytica Chimica Acta* 671 (1–2): 27–35. https://doi.org/10.1016/j.aca.2010.05.013.

Behrenbruch, Peter, and Thivanka Dedigama. 2007. "Classification and Characterisation of Crude Oils Based on Distillation Properties." *Journal of Petroleum Science and Engineering* 57 (1–2): 166–80. https://doi.org/10.1016/j.petrol.2005.10.016.

Bezerra, Andressa Cristina De Mattos, Danieli De Oliveira Silva, Gustavo Henrique Machado De Matos, Josuel Pereira Dos Santos, Claudio Neves Borges, Landulfo Silveira, and Marcos Tadeu Tavares Pacheco. 2019. "Quantification of Anhydrous Ethanol and Detection of Adulterants in Commercial Brazilian Gasoline by Raman Spectroscopy." *Instrumentation Science & Technology* 47 (1): 90–106. https://doi.org/10.1080/10739149.2018.1470535.

Bonfim, Ricardo R., Maria I.R. Alves, and Nelson R. Antoniosi Filho. 2012. “Fast-HRGC Method for Quantitative Determination of Benzene in Gasoline.” *Fuel* 99 (September):165–69. https://doi.org/10.1016/j.fuel.2012.04.027.

Correia, Radigya M., Eloilson Domingos, Vagne M. Cáo, Brenda R.F. Araujo, Sthefany Sena, Layla U. Pinheiro, André M. Fontes, et al. 2018. “Portable near Infrared Spectroscopy Applied to Fuel Quality Control.” *Talanta* 176 (January):26–33. https://doi.org/10.1016/j.talanta.2017.07.094.

D02 Committee. n.d. “Test Method for Determination of Benzene in Spark-Ignition Engine Fuels Using Mid Infrared Spectroscopy.” ASTM International. Accessed December 18, 2024. https://doi.org/10.1520/D6277-07R22.

Da Silva, Michelle Patrícia Ferreira, Lívia Rodrigues E Brito, Fernanda Araújo Honorato, Ana Paula Silveira Paim, Celio Pasquini, and Maria Fernanda Pimentel. 2014. “Classification of Gasoline as with or without Dispersant and Detergent Additives Using Infrared Spectroscopy and Multivariate Classification.” *Fuel* 116 (January):151–57. https://doi.org/10.1016/j.fuel.2013.07.110.

Dadson, J., S. Pandam, and N. Asiedu. 2018. “Modeling the Characteristics and Quantification of Adulterants in Gasoline Using FTIR Spectroscopy and Chemometric Calibrations.” Edited by Gene Hall. *Cogent Chemistry* 4 (1): 1482637. https://doi.org/10.1080/23312009.2018.1482637.

Doble, Philip, Mark Sandercock, Eric Du Pasquier, Peter Petocz, Claude Roux, and Michael Dawson. 2003. “Classification of Premium and Regular Gasoline by Gas Chromatography/Mass Spectrometry, Principal Component Analysis and

Artificial Neural Networks." *Forensic Science International* 132 (1): 26–39. https://doi.org/10.1016/S0379-0738(03)00002-1.

El Orche, Aimen, Amine Mamad, Omar Elhamdaoui, Amine Cheikh, Miloud El Karbane, and Mustapha Bouatia. 2021. "Comparison of Machine Learning Classification Methods for Determining the Geographical Origin of Raw Milk Using Vibrational Spectroscopy." Edited by Ana Domi Nguez Vidal. *Journal of Spectroscopy* 2021 (December):1–9. https://doi.org/10.1155/2021/5845422.

ESMAP. n.d. "Catching Gasoline and Diesel Adulteration." *THE WORLD BANK 2002*.

Fodor, G. E., K. B. Kohl, and R. L. Mason. 1996. "Analysis of Gasolines by FT-IR Spectroscopy." *Analytical Chemistry* 68 (1): 23–30. https://doi.org/10.1021/ac9507294.

Gautam, Kislaya, Xiaohong Jin, and Marion Hansen. 1998. "Review of Spectrometric Techniques for the Characterization of Crude Oil and Petroleum Products." *Applied Spectroscopy Reviews* 33 (4): 427–43. https://doi.org/10.1080/05704929808002061.

Haney, Marshall. 1923. "Petroleum." In *The Scientific Monthly*, Vol. 17:548-561 (14 pages). No. 6. https://www.jstor.org/stable/3693096.

Héberger, Károly. 2008. "Chemoinformatics—Multivariate Mathematical–Statistical Methods for Data Evaluation." In *Medical Applications of Mass Spectrometry*, 141–69. Elsevier. https://doi.org/10.1016/B978-044451980-1.50009-4.

Keenan, J.J., S.H. Gaffney, D.A. Galbraith, P. Beatty, and D.J. Paustenbach. 2010. "Gasoline: A Complex Chemical Mixture, or a Dangerous Vehicle for Benzene

Exposure?" *Chemico-Biological Interactions* 184 (1–2): 293–95. https://doi.org/10.1016/j.cbi.2010.01.016.

Khanmohammadi Khorrami, Mohammadreza, Mina Sadrara, and Mahsa Mohammadi. 2022. "Quality Classification of Gasoline Samples Based on Their Aliphatic to Aromatic Ratio and Analysis of PONA Content Using Genetic Algorithm Based Multivariate Techniques and ATR-FTIR Spectroscopy." *Infrared Physics & Technology* 126 (November):104354. https://doi.org/10.1016/j.infrared.2022.104354.

Kumari, Poonam, Gazal Garg, Daya Soni, and Shankar G. Aggarwal. 2023. "Measurement of Benzene and Other Volatile Organic Compounds: Implications for Its Inhalation Health Risk Associated with the Workers at a Fuel Station in Delhi." *Asian Journal of Atmospheric Environment* 17 (1): 7. https://doi.org/10.1007/s44273-023-00007-8.

Kümmerer, K. 2011. "Emerging Contaminants." In *Treatise on Water Science*, 69–87. Elsevier. https://doi.org/10.1016/B978-0-444-53199-5.00052-X.

Leal, Ana L., Artur M. S. Silva, Jorge C. Ribeiro, and Fernando G. Martins. 2020. "Using Spectroscopy and Support Vector Regression to Predict Gasoline Characteristics: A Comparison of ^{1}H NMR and NIR." *Energy & Fuels* 34 (10): 12173–81. https://doi.org/10.1021/acs.energyfuels.0c01889.

Lee, Don-Min, Dong-Hyun Lee, and In-Ha Hwang. 2018a. "Gasoline Quality Assessment Using Fast Gas Chromatography and Partial Least-Squares Regression for the Detection of Adulterated Gasoline." *Energy & Fuels* 32 (10): 10556–62. https://doi.org/10.1021/acs.energyfuels.8b02368.

———. 2018b. “Gasoline Quality Assessment Using Fast Gas Chromatography and Partial Least-Squares Regression for the Detection of Adulterated Gasoline.” *Energy & Fuels* 32 (10): 10556–62. https://doi.org/10.1021/acs.energyfuels.8b02368.

Li, Dongmei, and Xiaodong Han. 2011. “Methyl Tert-Butyl Ether.” In *Reproductive and Developmental Toxicology*, 617–21. Elsevier. https://doi.org/10.1016/B978-0-12-382032-7.10046-3.

Liu, R.X., J. Kuang, Q. Gong, and X.L. Hou. 2003. “Principal Component Regression Analysis with Spss.” *Computer Methods and Programs in Biomedicine* 71 (2): 141–47. https://doi.org/10.1016/S0169-2607(02)00058-5.

Lu, Xingcai, Dong Han, and Zhen Huang. 2011. “Fuel Design and Management for the Control of Advanced Compression-Ignition Combustion Modes.” *Progress in Energy and Combustion Science* 37 (6): 741–83. https://doi.org/10.1016/j.pecs.2011.03.003.

Mabood, Fazal, Farah Jabeen, Manzor Ahmed, Javid Hussain, Saaida A.A. Al Mashaykhi, Zainb M.A. Al Rubaiey, Saim Farooq, et al. 2017. “Development of New NIR-Spectroscopy Method Combined with Multivariate Analysis for Detection of Adulteration in Camel Milk with Goat Milk.” *Food Chemistry* 221 (April):746–50. https://doi.org/10.1016/j.foodchem.2016.11.109.

Mahmodi, Korosh, Mostafa Mostafaei, and Esmaeil Mirzaee-Ghaleh. 2019. “Detection and Classification of Diesel-Biodiesel Blends by LDA, QDA and SVM Approaches Using an Electronic Nose.” *Fuel* 258 (December):116114. https://doi.org/10.1016/j.fuel.2019.116114.

Mishra, Puneet, Alessandra Biancolillo, Jean Michel Roger, Federico Marini, and Douglas N. Rutledge. 2020. “New Data Preprocessing Trends Based on Ensemble of Multiple Preprocessing Techniques.” *TrAC Trends in Analytical Chemistry* 132 (November):116045. https://doi.org/10.1016/j.trac.2020.116045.

Moreira, L. S., L. A. d’Avila, and D. A. Azevedo. 2003. “Automotive Gasoline Quality Analysis by Gas Chromatography: Study of Adulteration.” *Chromatographia* 58 (7–8): 501–5. https://doi.org/10.1365/s10337-003-0065-z.

Pereira, Rita C. C., Vinicius L. Skrobot, Eustáquio V. R. Castro, Isabel C. P. Fortes, and Vânya M. D. Pasa. 2006. “Determination of Gasoline Adulteration by Principal Components Analysis-Linear Discriminant Analysis Applied to FTIR Spectra.” *Energy & Fuels* 20 (3): 1097–1102. https://doi.org/10.1021/ef050203e.

Prasad, P. Ravi, K. S. Rama Rao, K. Bhuvaneswari, N. Praveena, and Y. V. V. Srikanth. 2008. “Determination of Ethanol in Blend Petrol by Gas Chromatography and Fourier Transform Infrared Spectroscopy.” *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 30 (16): 1534–39. https://doi.org/10.1080/15567030701258444.

Sales, Rafaella De F., Luan Cássio Barbosa-Patrício, Neirivaldo C. Da Silva, Lívia Rodrigues E Brito, Maria Eduarda Fernandes Da Silva, and Maria Fernanda Pimentel. 2023. “Gasoline Discrimination Using Infrared Spectroscopy and Virtual Samples Based on Measurement Uncertainty.” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 303 (December):123248. https://doi.org/10.1016/j.saa.2023.123248.

Singh, A. 2003. “Determination of Benzene and Total Aromatics in Commercial Gasolines Using Packed Column GC and NMR Techniques.” *Fuel* 82 (1): 23–33. https://doi.org/10.1016/S0016-2361(02)00157-6.

Skrobot, Vinicius L., Eustáquio V. R. Castro, Rita C. C. Pereira, Vânya M. D. Pasa, and Isabel C. P. Fortes. 2005. “Identification of Adulteration of Gasoline Applying Multivariate Data Analysis Techniques HCA and KNN in Chromatographic Data.” *Energy & Fuels* 19 (6): 2350–56. https://doi.org/10.1021/ef0500311.

Swick, Derek, Andrew Jaques, J.C. Walker, and Herb Estreicher. 2014. “Gasoline Risk Management: A Compendium of Regulations, Standards, and Industry Practices.” *Regulatory Toxicology and Pharmacology* 70 (2): S80–92. https://doi.org/10.1016/j.yrtph.2014.06.022.

Teixeira, L, F Oliveira, H Dossantos, P Cordeiro, and S Almeida. 2008. “Multivariate Calibration in Fourier Transform Infrared Spectrometry as a Tool to Detect Adulterations in Brazilian Gasoline.” *Fuel* 87 (3): 346–52. https://doi.org/10.1016/j.fuel.2007.05.016.

Vempatapu, Bhanu Prasad, and Pankaj K. Kanaujia. 2017. “Monitoring Petroleum Fuel Adulteration: A Review of Analytical Methods.” *TrAC Trends in Analytical Chemistry* 92 (July):1–11. https://doi.org/10.1016/j.trac.2017.04.011.

Vitale, Raffaele, Marina Cocchi, Alessandra Biancolillo, Cyril Ruckebusch, and Federico Marini. 2023. “Class Modelling by Soft Independent Modelling of Class Analogy: Why, When, How? A Tutorial.” *Analytica Chimica Acta* 1270 (August):341304. https://doi.org/10.1016/j.aca.2023.341304.

Wiedemann, L, L Davila, and D Azevedo. 2005. “Adulteration Detection of Brazilian Gasoline Samples by Statistical Analysis.” *Fuel* 84 (4): 467–73. https://doi.org/10.1016/j.fuel.2004.09.013.

Xia, Jingjing, Jixiong Zhang, Yueting Zhao, Yangming Huang, Yanmei Xiong, and Shungeng Min. 2019. “Fourier Transform Infrared Spectroscopy and Chemometrics for the Discrimination of Paper Relic Types.” *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 219 (August):8–14. https://doi.org/10.1016/j.saa.2018.09.059.

Xia, Qi, Lei-ming Yuan, Xiaojing Chen, Liuwei Meng, and Guangzao Huang. 2019. “Analysis of Methanol Gasoline by ATR-FT-IR Spectroscopy.” *Applied Sciences* 9 (24): 5336. https://doi.org/10.3390/app9245336.

Xu, Jigang, Shujun Liu, Ming Gao, and Yonggang Zuo. 2023. “Classification of Lubricating Oil Types Using Mid-Infrared Spectroscopy Combined with Linear Discriminant Analysis–Support Vector Machine Algorithm.” *Lubricants* 11 (6): 268. https://doi.org/10.3390/lubricants11060268.

Yousefinejad, Saeed, Leila Aalizadeh, and Fatemeh Honarasa. 2016. “Application of ATR-FTIR Spectroscopy and Chemometrics for the Discrimination of Furnace Oil, Gas Oil and Mazut Oil.” *Analytical Methods* 8 (23): 4640–47. https://doi.org/10.1039/C6AY00051G.

Yu, Xiao-Zhang, and Ji-Dong Gu. 2006. “Uptake, Metabolism, and Toxicity of Methyl Tert-Butyl Ether (MTBE) in Weeping Willows.” *Journal of Hazardous Materials* 137 (3): 1417–23. https://doi.org/10.1016/j.jhazmat.2006.04.024.

BRIEF BIO-DATA OF LALBIAKTLUANGA

Name : Lalbiaktluanga

Father's Name : Laltawnluia

Date of Birth : 25 - 08 - 1991

Address : House No. M1, Pehlawn, Aizawl District
Mizoram, Pin- 796011.

Subject of Specialization : Laser and Spectroscopy

Educational Qualification :

H.S.L.C	: 2006	First Division	M.B.S.E.
H.S.S.L.C.	: 2009	Third Division	M.B.S.E.
B.Sc. (Physics)	: 2012	First Division	Annamalai University
M.Sc. (Physics)	: 2016	First Division	M.Z.U.
Pre-Ph.D Course	: 2017	First Division	M.Z.U.

LIST OF PUBLICATIONS AND ACTIVITIES

I. Journals:

1. J. Lalramnghaka, H.H. Thanga & Lal Biaktluanga (2022), Evaluation of gasoline fuel quality using FTIR spectroscopy and multivariate technique: a case study in Aizawl city, *Petroleum Science and Technology.* 41(5):1-23, DOI:10.1080/10916466.2022.2091596. **0.404 (2024) SJR**

2. LAL BIAKTLUANGA, J.LALRAMNGHAKA, H.H.THANGA (2020). Determination of Benzene and Methyl-Tert-Butyl Ether (MTBE) Content in Gasoline Fuel by Partial Least Square Regression Applied to FTIR Spectra. *International Journal of Advances in Science Engineering and Technology.* Volume-8, Issue-4 p (46-52), DOI: IJASEAT-IRAJ-DOIONLINE-20385. **0.177 (2024) SJR**

3. Lalbiaktluanga, J. Lalramnghaka, B. Lalremruata, R. Lalrempuia , H.H. Thanga (2023). Detection and estimation of adulterated gasoline fuel in India using FTIR-ATR spectroscopy with chemometric methods. *Infrared Physics and Technology,* 137 (2024) 105119, DOI: https://doi.org/10.1016/j.infrared.2024.105119. **0.595 (2024) SJR**

4. Lal Biaktluanga, Josef Lalhruaitluanga, J. Lalramnghaka and H.H. Thanga (2024). Identification and Measurement of Adulteration in Mizoram Gasoline Using FTIR Spectroscopy and Chemometric Methods. *Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and*

Technology Edition.), Vol: 43 Issue: 01-2024: DOI:10.5281/zenodo.10522287. **0.256 (2024) SJR**

5. Lalbiaktluanga and H.H. Thanga (2024). PCA and PLS-DA techniques in conjunction with an FTIR Spectrometer for assessing the quality of gasoline fuel. *European Chemical Bulletin.* Eur. Chem. Bull. 2024, 13(Regular Issue 03), 95 – 104: DOI: 10.53555/ecb/2024.13.03.09.

6. Lal Biaktluanga, Josef Lalhruaitluanga, J. Lalramnghaka and H.H. Thanga (2024). Analysis of gasoline quality by ATR-FTIR spectroscopy with multivariate techniques. *Results in Chemistry,* Vol: 8, June 2024, 101575: DOI: https://doi.org/10.1016/j.rechem.2024.101575. **0.605 (2024) SJR**

II. Conference Proceedings:

1. Lalbiaktluanga, H.H. Thanga, J.Lalramnghaka (2020), Classification of Gasoline quality in Aizawl by FTIR-ATR combined with PCA Multivariate Data Analysis, *proceeding of Asian Society for Academic Research- international conference, Hyderabad, India,* p(1-4).
2. Lalbiaktluanga, H.H. Thanga, J.Lalramnghaka (2020), Determination of Benzene and Methyl-tert-Butyl Ether (MTBE) in gasoline fuel by Partial Least Square Regression applied to FTIR Spectra, *proceeding of GSRD - international conference, New Delhi, India,* p(42-46).

III. Book Chapter:

IV. Conferences/ Workshop Attended:

1 Multidisciplinary International Seminar on “A Perspective of Global Research Process: present scenario and future challenges”, organized by Manipur University, Manipur (Imphal). (2019)

2 International Conference on Gas, Oil and Petroleum engineering held in Hyderabad, India organized by Asian Society for Academic Research. (2020)

3 International Conference on Oil, gas and Petrochemistry held in New Delhi, India organized by Global Society for Research and Development. (2020)

4 National Conference on Emerging Trends in Physics (NCETP 2021) organized by Department of Physics, Tezpur University, Assam, India.

5 Mizoram Science Congress (MSC) organized by Mizoram Science, Technology and Innovation Council (MISTIC), Directorate of Science and Technology (DST) Mizoram, Govt. of Mizoram. (2022)

6 International conference on Recent Advances in mathematical, Physical and Chemical Sciences organized by School of Physical Sciences, Mizoram University. (2024)

PETROLEUM SCIENCE AND TECHNOLOGY
https://doi.org/10.1080/10916466.2022.2091596

Taylor & Francis
Taylor & Francis Group

Check for updates

Evaluation of gasoline fuel quality using FTIR spectroscopy and multivariate technique: a case study in Aizawl city

J. Lalramnghaka, H. H. Thanga, and Lal Biaktluanga

Department of Physics, Mizoram University, Aizawl, Mizoram, India

ABSTRACT

Fuel quality control had become increasingly important due to the direct relationship between fuel quality and vehicle emissions. Variability of gasoline fuel from one refilling station to another is analyze using Fourier Transform Infrared spectroscopy (FTIR) and Principal Component Analysis (PCA). The result indicates that HC chain length, oxygenate, toluene and aromatic content are the major variables that cause sample variation and grouping. Determination of kerosene and Methyl tert-Butyl Ether (MTBE) concentration in gasoline fuel is done by using partial least squares (PLS) regression multivari-ate technique as applied to FTIR-ATR spectral data of the test samples. PLS calibration models were developed for the prediction of kerosene and MTBE and validated with independent set of known samples. Accordingly, among 179 tested samples, 103 samples contain kerosene (% v/v) less than 10%, 61 samples contain 11–20%, 11 samples contain 21–30% and 4 samples exceed 30% resulting that majority of gasoline fuel in the study area are adulterated. It is observed that 17 samples are MTBE blended with varying concentrations in the range of 1–7.34% v/v. This study provides an inexpensive and effective method to investigate gasoline quality, detection of adulter-ants and oxygenator in areas where adulteration of fuel is quite popular.

KEYWORDS

Fourier transform infrared (FTIR) spectroscopy; principal component analysis (PCA); partial least square regression (PLS-R); gasoline; oxygenator; kerosene; adulteration

1. Introduction

Air pollution in some India cities is a serious health issue. A tremendous increase in air pollution led by vehicular tail-pipe emission has become a major concern and is recognized as the major source of air pollution in many cities of India (Ghose, Paul, and Banerjee 2004; Goyal et al. 2006). According to the WHO database, cities in India leads the list of the world's most polluted cities, with 14 of the 15 cities featuring in the list are in India (Khan and Hassan 2020). In this situation, road transport sector is estimated to contribute as high as 70% of the total pollution (Shrivastava,

CONTACT H.H. Thanga hthanga@rediffmail.com Department of Physics, Mizoram University, Aizawl, 796004 Mizoram, India.

International Journal of Advances in Science Engineering and Technology, ISSN(p): 2321 –8991, ISSN(e): 2321 –9009 Volume-8, Issue-4, Oct.-2020, http://iraj.in

DETERMINATION OF BENZENE AND METHYL-TERT-BUTYL ETHER (MTBE) CONTENT IN GASOLINE FUEL BY PARTIAL LEAST SQUARE REGRESSION APPLIED TO FTIR SPECTRA

[1]LAL BIAKTLUANGA, [2]J.LALRAMNGHAKA, [3]H.H.THANGA

[3]Department of Physics, Mizoram University, University, Aizawl-796004, Mizoram, India
E-mail: [1]bitakawlni@gmail.com, [3]hthanga@rediffmail.com

Abstract - The pollution created by blending of chemical compounds in gasoline fuel is one of the major concerns in nowadays, a mixture of over 200-derived chemical compounds is found in gasoline. As a result of the obvious direct impact on quality affected by blended compounds, fuel quality assurance has already become extremely important. The current study describes the determination of blended compoundsin gasoline within thestate of Mizoramsamples. Principal Component Analysis (PCA)and Partial Least Squared Regression(PLSR) approaches were used to evaluate a total of 177 samples. Therefore, the PLSR with high R-square and RMSEC significance values are used to quantify the samples. In addition, the PLSRwith high accuracy in validation is reliable for the present work.

Keywords - Gasoline, FTIR, PCA, PLSR

I. INTRODUCTION

Methyl Tert-Butyl Ether (MTBE) and aromatic compounds are presently the blended chemical compounds in gasoline fuels to increase the fuel performance in many countries. Among the aromatic compounds, Benzene is found as the most common and most hazardous additive compound and has been used as an anti-knock agent to increase octane ratings[13]. Due to the presence of Benzene in gasoline fuel, the emission produced by a motor engine contains more Carbon Monoxide (CO). Being a toxic air pollutant and a well-known carcinogen, it is regulated in many countries to limit the concentration level at 1% in volume[1]. Apart from Benzene, Methyl tert-butyl Ether (MTBE) is also used as oxygenators which are chemical compound containing more than one or more oxygen atoms in their chemical structure and are used as octane enhancers[4] as well as to reduce the emission of Carbon Monoxide. Along with Benzene, MTBE is also a part of an environmental concern because the leakage from underground storage tanks can be soluble in water and does not biodegrade easily, so this is why it is regulated in many countries as well. However, due to the environmental pollutant the limit concentration level of MTBE is placed 1-15% in volume[2]. Therefore, it is necessary to determine the content of Benzene and Oxygenator in gasoline sample. IR spectroscopy coupled with Multivariate Analysis is an ideal method for quantifying the concentration of aromatic compounds and oxygenator in gasoline fuel. The ASTM D6277-07[1] and ASTM D5845-95[2] methods are used to determine the content of Benzene and MTBE in gasoline compounds; it employs the Principal Component analysis (PCA) to classify gasoline samples based on its quality and Partial Least Square Regression (PLS) which is used to predict quantity of the measured chemical compounds by analyzing standard calibration set and gasoline sample in collaboration with FTIR-ATR spectroscopy[3].

In Multivariate Analysis, PCA takes information carried by original variable and project into smaller latent variable called Principal Component. It is mainly used to reveal the hidden structure of large data set; it provides the relationship between samples and variables as well as the similarities and differences of samples caused by the measured variable. In PLS models, the PLS components are very similar to Principal Components, both the X and Y Matrices are simultaneously used to find the latent variable in X-matrices that will predict the latent variable in Y-Matrices. FTIR techniques coupled with Principal Component Analysis (PCA) and Partial Least Square Regression (PLS) were especially found successful for determining Benzene and MTBE in gasoline by using mid-IR (600-4000 cm^{-1}) with the help of Attenuated Total Reflectance (ATR)[4]. The main objective of this investigation is to classify the gasoline fuel base on its quality and to determine the amount of Benzene and Oxygenator content in gasoline fuel obtained from Filling stations in Aizawl area.

II. MATERIAL AND METHOD

Gasoline samples were collected from different filling stations in Aizawl area and various areas outside Aizawl; the total of 177 gasoline samples collected is of normal and standard quality. The solution of five compound mixtures i.e., Isooctane, Benzene, Toluene and pure Xylene mixture were analyzed according to ASTM D6277[1]. The chemical compounds used for analyzing the concentration of Benzene and Methyl-tert-Butyl-Ether (MTBE) were obtained from Sisco Research Laboratories Pvt. Ltd. With these chemical compounds, calibration and validation standard sets were prepared gravimetrically as described in

Infrared Physics & Technology 137 (2024) 105119

Contents lists available at ScienceDirect

Infrared Physics and Technology

journal homepage: www.elsevier.com/locate/infrared

Research Paper

Detection and estimation of adulterated gasoline fuel in India using FTIR-ATR spectroscopy with chemometric methods

Lalbiaktluanga [a], J. Lalramnghaka [a], B. Lalremruata [a], R. Lalrempuia [b], H.H. Thanga [a,*]

[a] Department of Physics, Mizoram University, India
[b] Department of Chemistry, Mizoram University, Aizawl 796004, Mizoram, India

ARTICLE INFO

Keywords:
Gasoline
Kerosene
Adulteration
FTIR
PCA
SIMCA
PLS-DA

ABSTRACT

In this work, Fourier Transform Infrared Attenuated Total Reflectance (FTIR-ATR) spectroscopy was combined with multivariate techniques, such as Principal Component Analysis (PCA), Soft Independent Modelling of Class Analogy (SIMCA), and Partial Least Squares Discriminant Analysis (PLS-DA), to classify and quantify the gasoline adulteration with kerosene. The Calibration models of Kerosene blended on gasoline fuel of 12 standard samples (2–24% v/v) were prepared. The chemometric tools were used to compare 142 samples from Mizoram and 12 samples from neighbouring states. The PCA and SIMCA approaches did not detect adulteration. The PLS-DA classification model found 7 samples were adulterated, with excellent specificity and sensitivity. The PLS regression model had the ability to predict adulterant concentrations with prediction errors less than 2% for all adulterants. The predicted concentration range of the PLS-R model is 0.3–13 % v/v, with high significant values of R-Square (0.996) and RMSE (0.405).

1. Introduction

Gasoline adulteration, by mixing unwanted substances into petroleum-based fuels, is one of the most acute problems in developing countries. Adulteration is described as the illegal or criminal insertion of any foreign substance into a product, causing it to fails to comply with its specified specifications. Adulterants are foreign compounds that, when introduced, alter and reduce the quality of the base transport fuels. The basic purpose of adulteration is to deceive consumers into purchasing poor fuel in order to increase their earnings while avoiding the hazardous consequences of this action [1,2]. Some of the adverse effects of adulteration included engine malfunctions, higher tailpipe pollution emissions, and health issues [3,4].

In India, transportation fuels are frequently adulterated with cheaper products, by-products, or waste hydrocarbon streams for monetary advantages. According to the Supreme Court (SC) of the Central Government in the year 2016, one of the most common types of adulteration in the country is combining kerosene with gasoline and diesel [5]. The amount of kerosene added to diesel can range from 20 to 30 %, while only a small amount is mixed with automobile gasoline [6]. Kerosene, also referred to as public distribution system (PDS), is a subsidised product provided by the government to those who are economically deprived [1,4]. As a result, PDS kerosene is mostly used throughout the country for fuel adulteration. Gasoline adulteration detection and identification are challenging tasks since the solvents used for this purpose are complex combinations of hundreds of distinct molecules, some of which might be found in unadulterated gasoline [7,8]. Detection of adulterants in automotive fuel using Gas Chromatography (GC)/Mass Spectrometry is considered to be among the most accurate and reliable analytical methods. However, there are several disadvantages, such as the high cost and the requirement of more than 30 min to obtain a single chromatogram, as well as sometimes requiring trained technicians for analysis [9–11].

Vibrational spectroscopy (NIR, MIR, and RAMAN) in combination with chemometric methods has proven to be a powerful tool for analysing gasoline fuels. [12–14]. Detection and estimation of gasoline adulteration have been done by multivariate methods like Principal Component Analysis (PCA), Partial Least Squares Discriminant Analysis (PLS-DA) and PLS regression, which were applied for statistical analysis of NIR spectral data [15]. Raman spectroscopy can also be used to evaluate the quality of fuels in a remote, rapid, and non-destructive manner without the need for reagents [16]. However, a Raman spectrometer is much more expensive than FTIR. FTIR spectroscopy combined with multivariate calibration analysis has been shown to be a

* Corresponding author.
E-mail address: hthanga@rediffmail.com (H.H. Thanga).

https://doi.org/10.1016/j.infrared.2024.105119
Received 30 January 2023; Received in revised form 21 December 2023; Accepted 1 January 2024
Available online 9 January 2024

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN: 1671-5497
E-Publication: Online Open Access
Vol: 43 Issue: 01-2024

IDENTIFICATION AND MEASUREMENT OF ADULTERATION IN MIZORAM GASOLINE USING FTIR SPECTROSCOPY AND CHEMOMETRIC METHODS

LAL BIAKTLUANGA
Department of Physics, Mizoram University, Aizawl, Mizoram, India.

JOSEF LALHRUAITLUANGA
Department of Physics, Mizoram University, Aizawl, Mizoram, India.

J. LALRAMNGHAKA
Department of Physics, Mizoram University, Aizawl, Mizoram, India.

H.H. THANGA*
Department of Physics, Mizoram University, Aizawl, Mizoram, India.
* Corresponding Author Email: hthanga@rediffmail.com

Abstract

Gasoline adulteration remains a global problem due to its economic, health and environmental impacts. The gasoline adulteration monitoring methods used in the present study were not effective in most developing countries due to the associated implementation costs. Therefore, a fast, reliable, and cheaper method of vibrational spectroscopy in combination with chemometric tools was developed. In this study, sensitive FTIR spectroscopy was proposed in combination with multivariate techniques to analyse adulteration in Mizoram gasoline fuels. Standard solutions were prepared by mixing the gasoline with different proportions of kerosene and methyl tert-butyl ether (MTBE), which are then used for model calibration. The chemometric techniques used were principal component analysis-linear discriminant analysis (PCA-LDA), classification of partial least squares discriminant analysis (PLS-DA) and the regression method (PLS). The developed PCA-LDA classification model has an overall error rate of 12.5%, while PLS-DA has an accuracy of 100% for model. The PLS regression models of kerosene and MTBE with a high significance value of R^2 0.996 and 0.986, RMSEP 0.275 and 0.291 were able to predict adulterant and MTBE concentrations, respectively. However, PCA-LDA achieved better accuracy than PLS-DA in classification analysis, especially on oxygenated gasoline samples.

Keywords: Gasoline, Adulteration, PCA-LDA, PLS-DA, PLSR.

1. INTRODUCTION

Fuel adulteration remains a problem across the world, particularly in developing countries. It refers to the illegal practice of blending of gasoline fuel with impurities or lower-grade quality, cheaper products, and solvents in attempts of increasing profits. It also often results in lower engine performance, leading to the harmful effects of engine tailpipe emissions on the environment. Hence, fuel quality monitoring has become more rigorous not just due to economic and social concerns, but also due to environmental concerns (Vempatapu & Kanaujia, 2017). To prevent fuel adulteration and ensure quality control, the government of India has implemented various measures and actions in accordance with the Marketing Discipline Guidelines (MDG) and a Dealership Agreement substituting between Retail Outlet (RO) dealers and public sector Oil Marketing Companies (OMCs).

PCA AND PLS-DA TECHNIQUES IN CONJUNCTION WITH AN FTIR SPECTROMETER FOR ASSESSING THE QUALITY OF GASOLINE FUEL

Lalbiaktluanga[1] And H.H. Thanga[2*]

Abstract

The primary aim of the study is to assess gasoline quality in Mizoram and compare it with samples from neighboring states (Assam) and laboratory kerosene-blended samples. A total of 174 samples were collected, including those from Mizoram and Assam, and laboratory-blended paraffin samples were categorized according to their quality characteristics. Principal Component Analysis (PCA) was used to analyses the data and identify patterns or trends among the samples. However, it seems that PCA alone may not be sufficient for precise classification. The Partial Least SquaresDiscriminant Analysis (PLS-DA) was also employed as a method for classification. PLS-DA is a supervised method that may offer more accurate classification results compared to PCA. The study highlights the importance of ensuring gasoline quality in Mizoram, especially in the face of adulteration practices. The use of advanced analytical techniques like PLS-DA enhances the ability to detect and classify adulterated samples accurately. This research could potentially inform regulatory measures to combat fuel adulteration and ensure consumer safety and satisfaction.

Keywords: Mizoram, Assam, Kerosene, PCA, PLS-DA.

[1,2*]Department of Physics, Mizoram University, Aizawl-796004, Mizoram, INDIA

***Corresponding Author:** H.H. Thanga
*Department of Physics, Mizoram University, Aizawl-796004, Mizoram, INDIA (hthanga@rediffmail.com)

DOI: 10.53555/ecb/2024.13.03.09

Results in Chemistry 8 (2024) 101575

Contents lists available at ScienceDirect

Results in Chemistry

journal homepage: www.sciencedirect.com/journal/results-in-chemistry

Analysis of gasoline quality by ATR-FTIR spectroscopy with multivariate techniques

Lal Biaktluanga, Josef Lalhruaitluanga, J. Lalramnghaka, H.H. Thanga *

Department of Physics, Mizoram University, University, Aizawl 796004, Mizoram, India

ARTICLE INFO

Keywords:
Gasoline
FTIR
Chemometrics
Benzene
MTBE
Adulteration

ABSTRACT

In this paper, chemometric methods were used for exploratory analysis, categorization, and quantification of gasoline fuel using Fourier Transform Infrared Spectroscopy (FTIR) data. During exploratory analysis, Principal Component Analysis (PCA) was employed, and to categorise the gasoline samples, Support Vector Machine Classification (SVMC), Linear Discrimination Analysis (LDA), and Partial Least Squares Discriminant Analysis (PLS-DA) were used. The concentration of Benzene, Methyl Tert-butyl Ether (MTBE), and Public Distribution System (PDS) Kerosene were determined using the Partial Least Squares Regression (PLSR), Principal Component Regression (PCR), and Support Vector Machine Regression (SVMR). All of the chemometric models had 100% accuracy and high R-square and RMSEC significance values. The SVM classification techniques were identified as a suitable choice for both classifying oxygenated and adulterated samples among all approaches. Both PLSR and PCR can also be suitable choices for quantification when dealing with high dimensional data. As a result, the research reported in this study has the potential to become a significant tool for improving gasoline sample discrimination and as a resource for quality control in the petroleum industry.

Introduction

In Mizoram (Northeast India), petrol or gasoline is the most commonly used transportation fuel for vehicles. Among the 54 gas stations are Indian Oil Corporation Ltd. (IOC), which is the largest petrol supplier with 46 petrol stations across the state of Mizoram. The other major companies are Hindustan Petroleum Corporation Ltd (HPC) and Bharat Petroleum Corporation Ltd. (BPC), also called reformulated gasoline (RFG). RFG is a gasoline blend with oxygenator that burns cleaner than regular gasoline and reduces the formation of smog and toxic pollutants in the air. However, the quality of gasoline products is generally controlled by blending different refinery streams which in turns is reflected by engine performance. So that, in order to be marketed, Benzene, the most widely added aromatic compound to gasoline, is limited in India due to its carcinogenic chemical substance [1]. To create a different grade of gasoline, oxygenator is also to fuel added chemical compounds to enhance octane rating and reduce air pollution. Methyl *tert*-butyl ether (MTBE) is the most often used oxygenator. Additionally, it is strictly limited inside the nation due to environmental concerns, mostly pertaining to ground water pollution [2]. The permissible levels of aromatics and oxygenates in gasoline may vary by region and are subject to regulation. Different regions or countries may have different standard and methods for testing gasoline quality, further complicating the identification and quantification of these components.

On the other hand, the other main problems of gasoline transport fuel are adulteration. Adulteration is the unlawful addition of any foreign material to a product that causes it to fail beyond its original specifications. It is common in the motor vehicle fuels (gasoline and diesel) sectors in developing countries, causing fuel quality to become poorer. Adulteration of petroleum-based automotive fuels is indeed a serious with several negative consequences. Adulterated fuels with impurities or lower quality substances can harm the engine's performance, can lead to long term damage to various engine components. Adulterated fuels often burn less cleanly, leading to higher emissions of harmful pollutants. This contributes to air pollution and possess to both the environment and living beings. Some of the different substances that blended gasoline adulteration are PDS Kerosene, naphtha, rubber solvents, aromatics, thinners, white spirits, turpentine, and lubricants [3]. To combat gasoline adulteration, government and regulatory bodies must enforced stringent regulations, conducted regular testing, and invested in advanced detection method such as Chromatography and Spectroscopy. Nevertheless, adulteration continues to be challenging to

* Corresponding author.
E-mail address: hthanga@rediffmail.com (H.H. Thanga).

https://doi.org/10.1016/j.rechem.2024.101575
Received 18 December 2023; Accepted 1 June 2024
Available online 2 June 2024

CLASSIFICATION OF GASOLINE QUALITY IN AIZAWL BY FTIR ATR COMBINED WITH PCA MULTIVARIATE TECHNIQUES

[1]LALBIAKTLUANGA, [2]H.H. THANGA, [3]J. LALRAMNGHAKA

Department of Physics, Mizoram University, Tanhril-796004, Aizawl, Mizoram
E-mail: bitakawlni@gmail.com

Abstract -
The transportation of our choices in daily life can have a huge impact on air quality and environment. Air pollution is one of the major concerns now a days especially in large cities of our country. Majority of air pollution is created by motor vehicles; it gives half of carbon monoxide emission. To reduce the carbon monoxide emission and increase the engine performance some additives compounds are added to fuel although it can also affect the quality of gasoline fuel. Among the additive compounds, oxygenator and aromatic compounds are most commonly used for reducing carbon monoxide emission and to increase octane ratings. Apart from these, the adulterations can also affect the quality of gasoline fuel and still remain global worry due to its environmental, health and economic effect. The present work covers the classification of gasoline quality obtained from different filling stations in Aizawl (Mizoram) and some parts outside Aizawl area. The analysis is done by using the Fourier Transform Infra-Red (FTIR)-ATR technique in mid-IR region, the variability of gasoline samples is investigated on Principal Component Analysis (PCA) of Multivariate technique. Spectra of gasoline samples in the region between 650-3200 cm^{-1} are then divided into three groups and variation based on their quality is obtained using the IR mode assignment. The result indicates that HC chain length, Oxygenate, Toluene and Aromatic contents are the major cause of variation. The discussion of IR mode assignment in the region of 650-3200 cm^{-1} are also included.

Keywords - Fourier Transform Infra-Red (FTIR)-ATR, Principal Component Analysis (PCA), HC Chain Length

I. INTRODUCTION

The transportation of our choices in daily life can have a huge impact on air quality and environment. Air pollution is one of the major concerns now a days especially in large cities of our country. The emission produced by motor engine is one of the majors caused of air pollution, along with the emission the blended compound in gasoline fuel can pollute the environment. Methyl Tert Butyl Ether (MTBE) and aromatic compounds are presently the blended chemical compounds in gasoline fuels to increase the fuel performance in many countries. Among the aromatic compounds, Benzene is found as the most common and most hazardous additive compound and has been used as an anti-knock agent to increase octane ratings. Due to the presence of Benzene in gasoline fuel, the emission produced by a motor engine contains more Carbon Monoxide (CO). Being a toxic air pollutant and a well-known carcinogen, it is regulated in many countries. Apart from Benzene, Methyl tert-butyl Ether (MTBE) is also used as oxygenators which are chemical compound containing more than one or more oxygen atoms in their chemical structure and are used as octane enhancers[4] as well as to reduce the emission of Carbon Monoxide. Along with Benzene, MTBE is also a part of an environmental concern because the leakage from underground storage tanks can be soluble in water and does not biodegrade easily, so this is why it is regulated in many countries as well. Therefore, it is necessary to classify the quality of gasoline fuels to decrease the pollution created in the environment. IR spectroscopy coupled with Multivariate Analysis is an ideal method for distributing the quality of gasoline fuels, it employs the Principal Component analysis (PCA) to classify gasoline samples based on its quality. In Multivariate Analysis, PCA takes information carried by original variable and project into smaller latent variable called Principal Component. It is mainly used to reveal the hidden structure of large data set; it provides the relationship between samples and variables as well as the similarities and differences of samples caused by the measured variable. In PLS models, the PLS components are very similar to Principal Components, both the X and Y Matrices are simultaneously used to find the latent variable in X-matrices that will predict the latent variable in Y-Matrices. FTIR techniques coupled with Principal Component Analysis (PCA) were found successful for distributing the quality of gasoline fuel using mid-IR (600-4000 cm-1) with the help of Attenuated Total Reflectance (ATR)[4]. The main objective of this investigation is to classify the gasoline fuel based on its quality and to determine the amount of Benzene and Oxygenator content in gasoline fuel obtained from Filling stations in Aizawl area.

International Journal of Advances in Science Engineering and Technology, ISSN(p): 2321 –8991, ISSN(e): 2321 –9009 Volume-8, Issue-4, Oct.-2020, http://iraj.in

DETERMINATION OF BENZENE AND METHYL-TERT-BUTYL ETHER (MTBE) CONTENT IN GASOLINE FUEL BY PARTIAL LEAST SQUARE REGRESSION APPLIED TO FTIR SPECTRA

[1]LAL BIAKTLUANGA, [2]J.L.ALRAMNGHAKA, [3]H.H.THANGA

[3]Department of Physics, Mizoram University, University, Aizawl-796004, Mizoram, India
E-mail: [1]bitakawlni@gmail.com, [3]hthanga@rediffmail.com

Abstract - The pollution created by blending of chemical compounds in gasoline fuel is one of the major concerns in nowadays, a mixture of over 200-derived chemical compounds is found in gasoline. As a result of the obvious direct impact on quality affected by blended compounds, fuel quality assurance has already become extremely important. The current study describes the determination of blended compoundsin gasoline within thestate of Mizoramsamples. Principal Component Analysis (PCA)and Partial Least Squared Regression(PLSR) approaches were used to evaluate a total of 177 samples. Therefore, the PLSR with high R-square and RMSEC significance values are used to quantify the samples. In addition, the PLSRwith high accuracy in validation is reliable for the present work.

Keywords - Gasoline, FTIR, PCA, PLSR

I. INTRODUCTION

Methyl Tert-Butyl Ether (MTBE) and aromatic compounds are presently the blended chemical compounds in gasoline fuels to increase the fuel performance in many countries. Among the aromatic compounds, Benzene is found as the most common and most hazardous additive compound and has been used as an anti-knock agent to increase octane ratings[13]. Due to the presence of Benzene in gasoline fuel, the emission produced by a motor engine contains more Carbon Monoxide (CO). Being a toxic air pollutant and a well-known carcinogen, it is regulated in many countries to limit the concentration level at 1% in volume[1]. Apart from Benzene, Methyl tert-butyl Ether (MTBE) is also used as oxygenators which are chemical compound containing more than one or more oxygen atoms in their chemical structure and are used as octane enhancers[4] as well as to reduce the emission of Carbon Monoxide. Along with Benzene, MTBE is also a part of an environmental concern because the leakage from underground storage tanks can be soluble in water and does not biodegrade easily, so this is why it is regulated in many countries as well. However, due to the environmental pollutant the limit concentration level of MTBE is placed 1-15% in volume[2]. Therefore, it is necessary to determine the content of Benzene and Oxygenator in gasoline sample. IR spectroscopy coupled with Multivariate Analysis is an ideal method for quantifying the concentration of aromatic compounds and oxygenator in gasoline fuel. The ASTM D6277-07[1] and ASTM D5845-95[2] methods are used to determine the content of Benzene and MTBE in gasoline compounds; it employs the Principal Component analysis (PCA) to classify gasoline samples based on its quality and Partial Least Square Regression (PLS) which is used to predict quantity of the measured chemical compounds by analyzing standard calibration set and gasoline sample in collaboration with FTIR-ATR spectroscopy[3].

In Multivariate Analysis, PCA takes information carried by original variable and project into smaller latent variable called Principal Component. It is mainly used to reveal the hidden structure of large data set; it provides the relationship between samples and variables as well as the similarities and differences of samples caused by the measured variable. In PLS models, the PLS components are very similar to Principal Components, both the X and Y Matrices are simultaneously used to find the latent variable in X-matrices that will predict the latent variable in Y-Matrices. FTIR techniques coupled with Principal Component Analysis (PCA) and Partial Least Square Regression (PLS) were especially found successful for determining Benzene and MTBE in gasoline by using mid-IR (600-4000 cm^{-1}) with the help of Attenuated Total Reflectance (ATR)[4]. The main objective of this investigation is to classify the gasoline fuel base on its quality and to determine the amount of Benzene and Oxygenator content in gasoline fuel obtained from Filling stations in Aizawl area.

II. MATERIAL AND METHOD

Gasoline samples were collected from different filling stations in Aizawl area and various areas outside Aizawl; the total of 177 gasoline samples collected is of normal and standard quality. The solution of five compound mixtures i.e., Isooctane, Benzene, Toluene and pure Xylene mixture were analyzed according to ASTM D6277[1]. The chemical compounds used for analyzing the concentration of Benzene and Methyl-tert-Butyl-Ether (MTBE) were obtained from Sisco Research Laboratories Pvt. Ltd. With these chemical compounds, calibration and validation standard sets were prepared gravimetrically as described in

PARTICULARS OF THE CANDIDATE

NAME OF THE CANDIDATE : LALBIAKTLUANGA

DEGREE : DOCTOR OF PHILOSOPHY

DEPARTMENT : PHYSICS

TITLE OF THESIS : EVALUATION OF GASOLINE QUALITY IN AIZAWL AND CONTENT OF AROMATIC AND OXYGENATE COMPOUNDS

DATE OF ADMISSION : 11.08.2017

APPROVAL OF RESEARCH PROPOSAL

1. BOS : 06.06.2018

2. SCHOOL BOARD : 24.04.2018

4. MZU REGISTRATION No.: 102 of 2015

5. Ph.D REGISTRATION No. : MZU/Ph.D./1141 of 24.04.2018

6. EXTENSION (IF ANY) : No. 16-2/MZU(Acad)/21/227-228, 11.08.2025

(Prof. B Lalremruata)

Head

Department of Physics

ABSTRACT

EVALUATION OF GASOLINE QUALITY IN AIZAWL AND CONTENT OF AROMATIC AND OXYGENATE COMPOUNDS

AN ABSTRACT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

LALBIAKTLUANGA

MZU REGISTRATION NO.: 102 of 2015

Ph.D. REGISTRATION NO.: MZU/Ph.D./1141 of 24.04.2018

DEPARTMENT OF PHYSICS

SCHOOL OF PHYSICAL SCIENCES

FEBRUARY, 2025

EVALUATION OF GASOLINE QUALITY IN AIZAWL AND CONTENT OF AROMATIC AND OXYGENATE COMPOUNDS

By

Lalbiaktluanga

Department of Physics

Name of Supervisor: Dr. Hranghmingthanga

Submitted

In partial fulfilment of the requirement of the Degree of Doctor of Philosophy in Physics of Mizoram University, Aizawl

Abstract

This abstract provides an overview of the quality assessment of gasoline sourced from Mizoram, a northeastern state of India. In addition to quality analysis, the study also investigates the levels of aromatic hydrocarbons and oxygenates. A comparison of various chemometric methods is conducted to explore their potential for future applications, particularly in areas where standard gas chromatography (GC) analysis is not available. This highlights the need for faster and more cost-effective techniques for gasoline quality analysis, which could be beneficial for future research.

The study focuses on fuels from Mizoram, its neighbouring state, and laboratory kerosene blends. The aim is to assess fuel quality by making comparisons between conventional, reformulated, and kerosene-blended fuels, taking into account the transportation timing of fuels into Mizoram by road. The analysis involves categorizing all the samples and comparing their quality to evaluate the available fuels in Mizoram.

To enhance the accuracy and predictive power of FTIR spectroscopy in gasoline quality assessment, multivariate data analysis methods such as PCA, SIMCA, PLSDA, and LDA are utilized. Combining FTIR with these multivariate techniques offers a robust, efficient, and non-destructive approach to monitoring gasoline quality. This integrated method provides precise classification and predictive capabilities, improving quality control and ensuring that gasoline complies with required standards for performance, safety, and environmental concerns.

The classification and quantification of aromatic hydrocarbons and oxygenates in the gasoline samples help evaluate their quality. Aromatic hydrocarbons, such as benzene, toluene, and xylene, are crucial indicators of fuel composition, affecting combustion efficiency, engine performance, and emissions. Oxygenates like ethanol and MTBE are added to enhance combustion and reduce emissions. The study aims to measure these compounds in different gasoline samples, assessing their concentrations and their impact on fuel quality. Higher aromatic hydrocarbon levels can indicate lower fuel quality and higher pollutant emissions, while oxygenates can improve combustion but may affect engine behaviour if present in excess.

To measure oxygenate content, 10 standard solutions were prepared by mixing conventional fuel with MTBE in concentrations ranging from 1-10% v/v. Similarly, for benzene concentration measurement, a five-compound mixture was analysed in accordance with ASTM D6277 standards. This method provides a precise technique for determining the concentration of aromatic compounds, such as benzene, toluene, and xylene, in fuel. The composition of the mixture was accurately analysed, enabling the quantification of individual aromatic hydrocarbons, which are important for assessing fuel quality and environmental impact. For benzene concentration, 25 standard solutions were prepared with concentrations ranging from 0.5-1.5% v/v, following ASTM D6277.

The study used FTIR spectroscopy combined with chemometric tools such as PCA, SVM, PLSDA, SVMR, PLSR, and PCR to classify and quantify the aromatic hydrocarbons and oxygenates in the gasoline samples. These techniques identify patterns in the data, allowing for the classification of gasoline samples based on their

chemical composition. By determining the concentration of each compound, the study provides insights into fuel quality, revealing how variations in aromatic and oxygenate levels affect fuel performance and emissions.

The results show a quality variation between gasoline samples from Mizoram and a neighbouring state, suggesting possible adulteration within Mizoram. Classification analysis revealed this difference in quality, and it was found that the levels of aromatic hydrocarbons and oxygenates in the gasoline did not exceed the permissible limits set by the country. Kerosene was not identified as the primary adulterant.

Furthermore, the study found that SVM and PCA-LDA were particularly effective in classifying gasoline samples, especially for distinguishing oxygenated gasoline. Both methods showed good classification accuracy. SVM and PLS-DA performed similarly in classifying adulterated gasoline samples. For quantification, PLSR and PCR were found to be comparable and outperformed SVMR, although SVMR's predictions were close to those of PLSR and PCR. Among all the methods, SVM was found to be particularly strong in classifying both oxygenated and adulterated samples due to its strong discriminatory power. These findings provide valuable insights into the use of chemometric techniques for gasoline analysis and classification, contributing to enhanced quality control in the petroleum industry.