Non-Invasive Milk Quality Monitoring using mm-wave Radar System

by

Jessica Katie Chong

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2024

© Jessica Katie Chong 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The Internet of Things (IoT) technologies is becoming increasingly valuable in industries for real-time monitoring of goods during manufacturing. This thesis explores the implementation of a low-cost electromagnetic radar solution to detect the quality of goods as they progress throughout the production, enabling companies to respond swiftly to potential quality issues that may arise. A specific application is demonstrated in the dairy industry, where milk can be monitored by tracking changes in the dielectric properties as deviations from a defined standard.

The dielectric properties of various concentrations of salt and sugar dissolved in distilled water are measured using an open-ended coaxial probe technique over the frequency range 200 MHz to 20 GHz. The conducted literature survey indicates a decrease in the value of the dielectric constant and an increase of the dielectric loss with an increase in concentration over this frequency range. Since the dielectric properties vary, the survey indicates that it is possible to develop a radar that can accurately detect these concentrations and the slight variances in a food product.

To investigate the relationship between milk fat content and dielectric properties, milk and cream samples with fat content ranging from 0.1-35% were tested using an open-ended coaxial probe connected to a vector network analyzer across frequency range 200 MHz to 20 GHz. The results revealed a linear decrease in both the dielectric constant and dielectric loss with increasing milk fat content, while other milk constituents also influenced these properties. Two regression models were developed to predict the fat content based on dielectric measurements, yielding promising accuracy.

Further measurements were extended up to 67 GHz to evaluate dielectric property variations at 60 GHz, the targeted frequency for radar implementation. Using this data, simulations were conducted in Ansys HFSS to validate the ability of a simple radar system to distinguish between different products based on their scattering parameter responses. Additionally, two electromagnetic millimetre-wave frequency-modulated continuous wave (FMCW) radars were tested on milk cartons, both stationary and moving on a conveyor belt. Data collected from the moving cartons were used to train and test a support vector classifier algorithm in Python, achieving 87.5% accuracy. The performance was compared with two alternative classification algorithms. This proof-of-concept demonstrates the potential for extending the proposed approach to other applications and industries, offering a versatile solution for real-time quality monitoring.

Acknowledgements

First and foremost, I have to thank my family and pet dog for their unwavering support and encouragement over the years. I am indebted to them for their moral, emotional, and financial support. This journey would not have been possible without them.

I would like to express my sincerest appreciation to Professor George Shaker, Professor Lan Wei, Professor Stefan Idziak for their support and guidance throughout this project.

Lastly, thank you to Victoria, for all her love and support.

Dedication

I would like to dedicate this to my family whose unwavering love and support have guided me to become the person I am today. This accomplishment is theirs as much as it is mine.

Table of Contents

Author's Declaration ii

Abstract iii

Acknowledgements iv

Dedication v

List of Figures ix

List of Tables xii

1 Introduction 1

1.1 Motivation 1

1.2 Research Objectives 2

1.3 Thesis Outline 3

2 Fluid Dairy Industry 5

2.1 Butterfat 5

2.2 Importance of Low-Cost, Real-time, Non-Invasive Quality Monitoring . . . 6

2.3 State of the Art Technologies 8

3 Dielectric Characterization Work **13**
3.1 Theory of Dielectrics . . . 13
3.1.1 Behaviour of Dielectrics . . . 13
3.1.2 Debye's Equations and Ionic Conductivity . . . 16
3.1.3 Theory Applied to Salt and Sugar Solutions . . . 20
3.2 Components of Milk . . . 22
3.2.1 Behaviour of the Dielectric Properties of Milk . . . 23
3.3 Hardware . . . 24
3.3.1 Scattering Parameters . . . 24
3.3.2 Open-Ended Coaxial Probe Fundamentals . . . 26
3.4 Experiments . . . 33
3.4.1 Sample Preparation . . . 33
3.4.2 Characterization of Salt and Sugar Solutions . . . 35
3.4.3 Temperature Study . . . 39
3.4.4 Characterization from 200 MHz - 20 GHz . . . 41
3.4.5 Linear Regression Analysis . . . 45
3.4.6 Characterization from 5 GHz - 67 GHz . . . 50

4 Radar Work **56**
4.1 Radar System Theory . . . 56
4.1.1 Radar Fundamentals and Radio System Building Blocks . . . 56
4.1.2 Radar Processing . . . 61
4.1.3 Dielectric Characterization using FMCW Radars . . . 66
4.1.4 Radar Hardware . . . 67
4.2 Full-Wave Electromagnetic Simulations . . . 68
4.2.1 FEM/FEBI Simulation . . . 69
4.2.2 SBR+ Simulation . . . 74
4.3 Machine Learning . . . 78

4.4 Proof of Concept . . . 82
4.4.1 Stationary Testing . . . 83
4.4.2 Conveyor Belt Testing . . . 96

5 Conclusion **104**
5.1 Summary . . . 104
5.2 Suggestions for Future Work . . . 105
5.2.1 Experimental Improvements . . . 105
5.2.2 Infineon Radar Thermal Characterization . . . 107
5.2.3 Sentaurus TCAD Electrothermal Simulation . . . 108

References **109**

APPENDICES **120**

A Modelling Behaviour using Mixture Equations **121**

B Regression Modelling in Python **126**

C SVM Classification in Python **131**

D Machine Learning Applied to Glass Jar Experiment **133**

E HFSS SBR+ Packaging Study **135**

F Infineon Radar Thermal Characterization **138**

G Gallium Nitride Power Amplifiers **141**
G.1 Gallium Nitride High Electron Mobility Transistors . . . 141
G.2 Sentaurus TCAD Electrothermal Simulation . . . 142
G.3 Multifinger GaN HEMT Electrothermal Simulation . . . 147

List of Figures

1.1 Real-time quality status reporting using radar. 2
1.2 Illustration of proposed radar data collection. 4

3.1 Plot of polarizability vs frequency. 15
3.2 Plot of Debye Model with ionic conductivity term. 19
3.3 Scattering parameters of a two-port network. 25
3.4 Coaxial probe with flange. 27
3.5 DAK-TL-P block diagram setup. 29
3.6 Setup of DAK-TL. 30
3.7 Salt and sugar solutions prepared with various concentrations. 35
3.8 Dielectric characterization of 2-10% salt solutions from 200 MHz to 20 GHz. 37
3.9 Dielectric characterization of 2-10% sugar solutions from 200 MHz to 20 GHz. 38
3.10 DAK-TL calibration comparison and temperature study. 40
3.11 Dielectric characterization of milk and cream from 200 MHz to 20 GHz. . . 42
3.12 Comparison of dielectric constant/loss of cream with FR filter. 44
3.13 Dielectric constant/loss vs butterfat content at 2450 MHz using a linear regression model. 47
3.14 Dielectric characterization of milk and cream from 5 GHz to 67 GHz. . . . 51
3.15 Dielectric constant/loss vs butterfat content at 60 GHz using a linear regression model. 54

4.1 Transmitter and receiver radio system. 60

4.2 Simple radar architecture. 61
4.3 Radar chirp and data cube. 64
4.4 Concept workflow describing data collection and processing. 65
4.5 Illustration of multiple dielectric slab problem. 68
4.6 Google ATAP SOLI radar chip with labelled antennas and packaging. . . . 69
4.7 Simplified waveguide radar system in HFSS. 73
4.8 Magnitude of reflections of various milk products at 60 GHz in HFSS. . . . 74
4.9 Carton model in HFSS SBR+. 75
4.10 Magnitude of reflections of various milk products at 60 GHz in HFSS SBR+. 77
4.11 Illustration of support vector machine. 80
4.12 Infineon radar and setup with wooden fixture. 84
4.13 Averaged magnitude of reflected responses of regular/core milk. 85
4.14 Averaged magnitude of reflected responses of microfiltered milk. 86
4.15 Milk carton positioned in front of SOLI for stationary test. 89
4.16 Magnitude of reflected responses from a 3.25% milk carton using SOLI radar. 90
4.17 Averaged magnitude of reflected responses using SOLI radar. 91
4.18 Preparation of glass jar experiment. 91
4.19 Magnitude of reflected responses from a 3.25% milk in a glass jar using SOLI radar. 92
4.20 Averaged magnitude of reflected responses with glass jar using SOLI radar. 93
4.21 Confusion matrices for stationary carton data. 95
4.22 Positioning of cartons on the conveyor belt. 96
4.23 Responses of 0.10% milk cartons on a conveyor belt. 98
4.24 Responses of different milk cartons on a conveyor belt. 99
4.25 Support vector machine confusion matrix. 102

A.1 Modelling the dielectric constant/loss vs butterfat content at 2450 MHz. . 124

D.1 Confusion matrices for stationary glass jar data. 134

E.1 Magnitude of reflection from cardboard block in HFSS SBR+. 136
E.2 Magnitude of reflection from glass block in HFSS SBR+. 136
E.3 Magnitude of reflection with 2 mm thick cardboard packaging in HFSS SBR+. 137
E.4 Magnitude of reflection with 2 mm thick glass packaging in HFSS SBR+. . 137

F.1 60 GHz radar system platform and thermal characterization. 139
F.2 Radar thermal characterization setup. 139
F.3 Heat map of 60 GHz radar. 140

G.1 Cross-section of a GaN HEMT. 142
G.2 IV curves computed using different temperature models. 144
G.3 Cross-Sectional heat map of GaN HEMT for $V_D = 10V$ and $V_G = 0V$. . . . 146
G.4 IV curves with different temperature models for 4-finger transistor. 148
G.5 Cross-sectional heat map and lattice temperature of 4-finger GaN HEMT. . 148
G.6 Electrostatic potential and lateral electric field. 149
G.7 Lattice temperature and total heat. 150
G.8 Electron mobility and electron velocity. 150

List of Tables

3.1 Classes of dielectrics and their observable polarization mechanisms 16
3.2 Recommended calibration settings from 200 MHz to 20 GHz 33
3.3 Concentrations of prepared salt solutions 34
3.4 Concentrations of prepared sugar solutions 34
3.5 Concentrations of prepared mixed creams. 36
3.6 Predicted fat content at 2450 MHz from dielectric constant using test data 48
3.7 Predicted fat content at 2450 MHz from dielectric loss using test data . . . 48
3.8 Predicted fat content of mixed cream samples at 2450 MHz 49
3.9 Predicted dielectric properties at 67 GHz 55

4.1 Infineon radar parameters 69
4.2 SOLI radar parameters 70
4.3 Calculated wavelength inside dielectric based on measured relative permittivity 74
4.4 Summary of machine learning techniques 82
4.5 Legend labels defined for data collection using Infineon radar 84
4.6 Machine learning predictions for stationary carton data 94
4.7 Support vector machine algorithm predictions 101

A.1 Summary of mixture equations 123

Chapter 1

Introduction

1.1 Motivation

With recent advancements in technology, the Internet of Things (IoT) has the potential to enable low-cost, non-invasive, real-time monitoring of food product quality throughout the production process [1], [2]. The real-time aspect allows for the early addressment of quality issues and reduces the chances of having production losses, supply shortages, and labour costs [2]. This would greatly improve supply and production planning. The product contents not meeting specifications and adulterated samples containing water, cleaning chemicals, or foreign objects in the product are a few issues that could arise. By using a frequency-modulated continuous wave (FMCW) radar to detect changes in a product's contents and the unique dielectric fingerprint of a sample, products that deviate too much from their reference sample can be flagged. Their results and timestamps can be made accessible over a network as illustrated in Figure 1.1. These radars are known for their high accuracy and sensitivity at close range [3]. Although there are many possible applications of this method in the food and beverage industries, the focus of this thesis is on its application to the fluid dairy industry.

An example of real-time monitoring found in literature during a production process is to determine the moisture content of grain [4]. According to [4], it is possible to develop a technology that can be installed on a production line to determine the moisture content of grains while being conveyed. This idea is adopted in this thesis. Conventional quality control methods for finished goods in the fluid dairy industry use batch testing where a carton or bag of milk is taken off a production line for sampling. The sample is representative of the product for the next few hours. The samples are then brought into a lab for testing

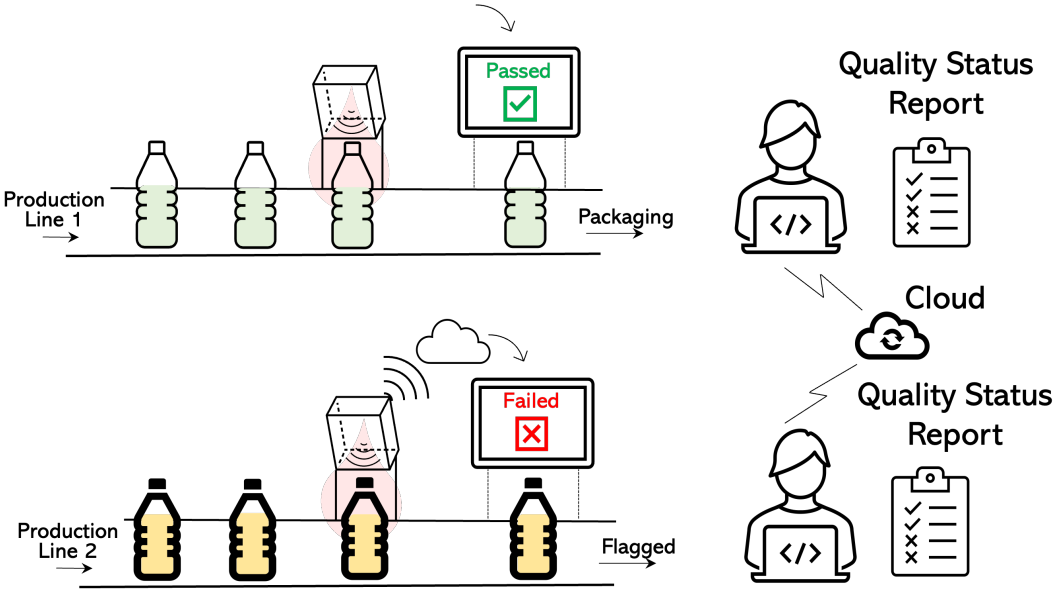


Figure 1.1: Real-time quality status reporting using radar.

using expensive and bulky equipment. Due to the nature of these instruments, they cannot be used for real-time monitoring on a production line, during transport, or to monitor milk during the pasteurization process. It can be used at the line, but the machines are quite costly. The global production of milk has been steadily increasing over the past few years and there has been an increasing need to test milk quality rapidly, non-invasively, and to provide real-time results.

1.2 Research Objectives

The purpose of this thesis is to demonstrate a proof-of-concept that liquid characterization can be performed using a low-cost, off-the-shelf mm-wave radar for the monitoring of milk products in packaging while being conveyed. This is a new technical contribution not yet attempted or demonstrated in literature. This thesis walks through the entire proof-of-concept step-by-step. This starts from examining the market and the need for this type of technology. It delves into the physics of whether or not this would be feasible, why exactly the selected technology was chosen, and features highly experimental work with extensive hours of data validation.

The concept is first examined by studying the dielectric properties of various concentrations of salt and sugar dissolved at room temperature. These are two common ingredients found in many food products. Next, the background knowledge will be applied to examine the dielectric properties of butterfat content with frequency. Butterfat, or milk fat, is a major constituent of milk and cream and is the part that would vary the most among these products. Differentiability of the dielectric properties of different concentrations of milk and cream samples will be the key factor when performing the measurements as they will provide a ground truth.

Since the radar is not a conventional dielectric characterization technique, the ground truth measurements provide confidence that the milks will be different between each other and distinguishable when working with the radar signals. After the initial characterization, the collected data is used in a few simulations utilizing a simple radar system. This is to simplify the radar setup and use a controlled environment to study the behaviour of the scattering responses. Then a mm-wave radar will be used to characterize milk cartons as they are stationary or moving along a conveyor belt. Finally, the radar signals are used to train a machine learning (ML) model for classification. To develop a golden sample or a reference model, all the possible variations and combinations of measured radar signals for a product must be used in the training stage.

Note, that although this thesis focuses on applying the radar to milk products for quality assurance and control, it is possible to extend this work to other liquids and industries. The proposed implementation of a radar on a production line is illustrated in Figure 1.2.

1.3 Thesis Outline

The structure of this thesis is as follows:

- **Chapter 2:** Literature survey and information about the dairy industry. The importance of a low-cost, real-time, non-invasive monitoring solution. State of the art technologies and current work being conducted for milk fat detection. Comparison of FMCW radar technology with a few state of the art technologies.
- **Chapter 3:** Background information regarding dielectric properties and modelling. Hardware and instrumentation used for dielectric characterization including setup and calibrations. Dielectric characterization results illustrating ground truth for differentiability. Selection of radar frequency based on the dielectric measurements.

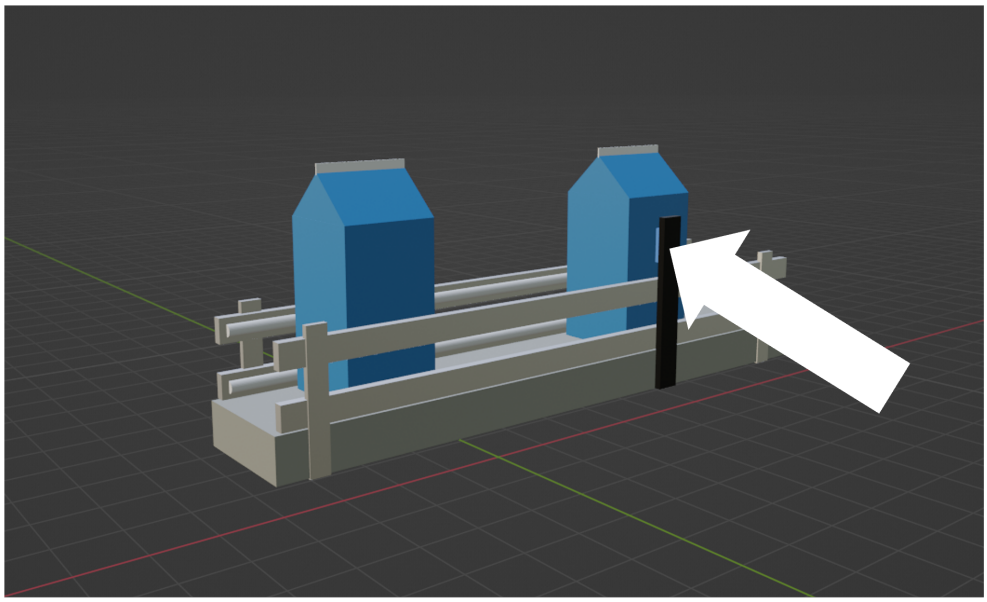


Figure 1.2: Illustration of proposed radar data collection.

- **Chapter 4:** Basics of radar systems and signal processing. Wave propagation through the classic multiple dielectric slab problem and basis of characterization using FMCW radars. Presentation of simulations, experimental results, and classification algorithms used for the proof of concept.
- **Chapter 5:** Recap and conclusion. Recommendations for future testing.

Chapter 2

Fluid Dairy Industry

2.1 Butterfat

In 2015, 497 million metric tons of milk were produced by cows and increased to 549 million metric tons in 2023 [5]. The total milk production volume in Canada was 94.5 million hectolitres in 2022 [6]. This is aligned with data from the Food and Agriculture Organization of the United Nations that illustrates a sharp increase in milk production in Canada starting in 2015 and a steady increase in global production since 1961 [7]. From [8], the largest milk output comes from Asia and increases in 2020 were mainly driven by India, China, Pakistan, and Turkey. There is a growing need for rapid quality check techniques from farm to product in terms of composition of fat, lactose, protein, microbacteria, etc. There are a multitude of literature work published in the last few years regarding the detection of milk fat, adulteration, and spoilage. This thesis primarily focuses on the detection of milk fat.

The general procedure and basic processing steps in a dairy processing plant are outlined in [9]. The milk travels from a supplier to a plant in large tankers with cooling systems. Once they arrive at the plant, the raw milk is tested for butterfat, lactose, protein, pH, microbacteria, and more. The butterfat is logged and recorded to pay the farmers. Milk with more butterfat is worth more than milk with less butterfat and varies with seasonality and the species and nutrition of a cow. For instance, raw milk is fattier in the winter than in the summer. More details regarding the value of butterfat content vs milk yield can be found in [10], [11]. Skim milk refers to 0.10% milk, partly-skimmed refers to 1.00% or 2.00% milk, and homogenized refers to 3.25% or whole milk. The milk is tested at every

stage of the production process to ensure the quality is maintained. This involves manually sampling of the milk to bring it back to the lab for testing.

As a key component in milk and the main determinant in how much milk is worth, butterfat or milk fat percentage is crucial to test and keep track of in the dairy industry. Although other ingredients and spoilage are just as crucial, the butterfat is where majority of the dollar value is. This is also the reason why homogenized whole milk is more expensive than skim milk and partly-skimmed milk. If the milk fat percentage is not meeting specifications or legal regulations, this becomes a quality issue. Companies and farms would also like to run their products as lean as possible for the best cost efficiency. This is the basis of why butterfat is selected as the sensing parameter for this thesis.

There is a lack of real-time monitoring because the sampling of milk is manual and testing methods are invasive. One may think that with modern technology, companies can see the overall composition of the milk as it flows through the piping system, into tanks, and fillers. However, this is far from the case. The importance of real-time monitoring is addressed in the following section. One may ask how often do butterfat issues occur and how useful this technology would be, but this is a difficult estimate as the issue varies from location to location and is highly dependent on many factors.

2.2 Importance of Low-Cost, Real-time, Non-Invasive Quality Monitoring

An example of an industry standard equipment is the MilkoScan FT120 and MilkoScan FT1 provided by FOSS Analytics. It is used to test for the composition of milk. These instruments use Fourier-transform mid-infrared (FTIR) spectrometry [12] and take roughly 30 seconds to test a single sample. This measurement technique is also applicable to wine, and their wine version of the machine is selling for approximately $45K CAD on [13]. It is safe to assume that the milk version of the machine is roughly the same price. This technique is based on an absorption spectrum using mid-infrared wavelengths and a cluster algorithm to identify, sort, and predict the peaks occurring in the spectrum. Cluster algorithm groups data based on other similar datasets. It requires little to no sample preparation, but is expensive and invasive.

Another example of an industry standard equipment is the Smart Trac Rapid Analysis System provided by CEM that uses nuclear magnetic resonance (NMR) to determine the percentage of milk fat content in a sample. It is typically used for creams and any thicker milk product. It provides a higher accuracy than the FTIR technique but requires a slightly

longer test time. The sample would need to be pipetted onto a square paper film and dried in a solids analyzer. This drying process takes time as it is removing the moisture on the paper film and leaving behind only the solid content. The sample would then be placed into another analyzer to check for the fat content by applying a magnetic field and radio-frequency signal. The absorption spectrum is analyzed again here. More details about this can be found in [14] and the instrument goes for roughly $105K CAD [15].

The two techniques are examples of what are used in industry and it is clear that the instruments are very expensive. This may not be affordable to smaller businesses or farms who need a quick and accurate method for testing the butterfat content in a milk sample. Because these two techniques are invasive and not real-time, larger companies could seek non-invasive real-time technologies to improve their supply chain. In high volume manufacturing, there could be numerous issues that occur leading to quality issues which then affect the supply chain. Imagine if there was an milk fat content issue with a tank such that it was reading 1.80% instead of 2.00%. Say that numerous fillers were running from that tank. In a hypothetical estimate of a worst-case scenario, it is easy to hit 50 pallets that need to be put on hold for rework. The amount of milk needed to be reworked depends on the size of the product. Each pallet contains 54 cases. Assuming that each pallet contained 2L cartons, there would be 9 cartons per case and there would be a total of 48,600 litres of milk put on hold for a single incident.

Say that the rate of a carton filler was 2858 cartons per hour, let the filler run at the wrong concentration for the time interval between testing and this would be a better estimate on a day to day basis. 2 hours can be taken as an example for the time elapsed between tests. This would result in recalls, waste, rework, packaging costs, labour costs, supply shortages, impact on brand name, and much more. Introducing a low-cost, non-invasive technology capable of real-time monitoring into the market could greatly benefit businesses, big and small. The exact impact of these issues and how often it occurs is difficult to quantify because it is highly dependent on many factors. Individual companies and farms could be interviewed in a future market research study to determine the impacts of the current issues.

Moreover, the real-time quality monitoring system will enable farmers and businesses of all sizes to run as lean as possible. A carton that an individual picks up from the grocery store may not necessarily be 2.00% as printed on the label. This could vary from 1.90% to 2.10% for example, and there are legal regulations on what the range for a 2.00% milk is. Companies and farmers could save money if they could run efficiently at 1.90%-2.00% instead of running at 2.10%. This could help control the amount of fat content they have in a product to reduce costs. Furthermore, if other ingredients are known, for example protein, it could help reduce costs even more. This one of the main objectives

that QuantWave Technologies, a startup by Prof. C. Ren's lab group from the University of Waterloo, aims to do [16]. They provide a brief case study on milk accessible on their webpage. This will be described more in the next section.

Additionally, because the two techniques described above are expensive, not every farm or company will have access to rapid spectroscopy techniques. According to [1], milk samples are picked up from farms and transported to labs where the milk fat content is tested for manually. From [2], a low-cost fat analysis method is through centrifugation in a butyrometer. This method is not automatic and requires isoamyl alcohol and a high concentration of sulfuric acid. A new technology for milk fat detection should be low-cost to benefit smaller companies and farms who cannot afford the best-in-line, industry standard testing methods. Traditional testing techniques have been described as time-consuming and prone to errors in dairy farming [1], [17].

The real-time monitoring of milk supply and products would greatly improve the supply chain, forecasting, and production planning. Current research on developing state of the art technologies will be discussed in the next section.

2.3 State of the Art Technologies

There are a handful of current state of the art technologies that can be found in literature, but only a few will be mentioned here in more detail. As mentioned in the previous section, two examples of state of the art technologies currently being used in industry are FTIR and NMR. There are some current works using different technologies such as dielectric spectroscopy, visible light spectroscopy, luminescence spectroscopy, and microwave microfluidic lab-on-a-chip technology applied to milk quality monitoring and compositional analysis. The proposed technology and its prior works will also be discussed, however it has not been used to quantify milk fat in a sample yet. Microwave and optical technologies are heavily favoured for their rapid, little to no sample preparation, minimal invasiveness, and potential for real-time monitoring.

Dielectric spectroscopy can be expensive due to the high cost of the vector network analyzer (VNA) needed. This technique is explored further in this thesis in Chapter 3 with its own literature survey to provide the ground truth for the radar work. There are multiple different types of dielectric characterization techniques such as resonators, time domain spectroscopy (TDR) or reflectometry, open-ended probes, free-space transmission, etc. [4]. The method selected for the work presented in this thesis is the open-ended coaxial probe technique. The basis of dielectric spectroscopy operates on the reflective and transmissive

properties of the liquid. Each sample has its own unique electromagnetic fingerprint. This technique also provides a working principle for microwave sensor designs, for example microwave resonator sensors [2], [18], which can be integrated with a miniaturized VNA front-end for measurements. Although the measurements are instantaneous, rapid, and require very little sample preparation, in-line, real-time monitoring is difficult to achieve because of its invasiveness.

Visible light spectroscopy is studied in [1], where the authors utilized an Arduino, spectrometer, Raspberry Pi and light sources for the solution. The light sources and their scattered light were studied individually and then compared for red, green, and blue visible lights. The experiment was conducted in a dark room to avoid other light interferences. The working principle of this technology is based off of the absorption and scattering of light using a spectrometer. The development of this solution has been stated in [1] to cost under $500. Yavari et al. concluded that because low-fat milk has less total solids in the milk compared to whole milk, there is less absorption into the liquid and there is a larger scattering response. There is a glaring downside with this technique and that the testing had to be done in a dark room, but may be suitable for monitoring the milk fat content in tanks since it would be dark.

Other milk quality monitoring systems using IoT can be found in [19], [20], [21], [22], and [17], but they focus on detecting adulterants in milk. Kumar et al. [19] utilize a pH sensor, conductivity sensor, and carbon dioxide sensor implemented with an Arduino UNO and a Back Propagation Neural Network for classification. Salinity, level, and spoilage of milk are tested for by Kalaiarasi et al. in [20] using an Arduino UNO, level sensor, gas sensor, and pH sensor. The Arduino was used with a WiFi module and a free software called Thing Speak to transfer data from the microcontroller storage to the cloud database. They mention that the data can be transferred every week or every month. These granular timestamps are not be fine enough to detect issues as they arise if the cloud database is updated only once a week.

M. Corradini's lab group at the University of Guelph studies non-invasive methods for the real-time monitoring of food quality, stability, and safety based on luminescence spectroscopy [23]. Her research group works towards developing big data models to predict nutrient stability and microbial quality which are used to forecast shelf-life and spoilage in a food product real-time. Corradini mentions that the application of luminescence spectroscopy is currently studied for several food products including the rennet-induced coagulation of milk. This technique depends on a molecule's photophysical properties and refers to the photoluminescent light that is emitted by the atom during UV or visible radiation testing.

C. Ren's lab group at the University of Waterloo has founded a startup called Quant-Wave Technologies [16] that aims to provide a multi-parameter, real-time, AI-powered IoT sensing platform. They have developed a smart platform that utilizes microwave sensing in a microfluidic system for liquid fingerprint characterization and compositional analysis. This aids companies in their decision-making process and improves their operational efficiency. After the characterization, their ML algorithm will identify and quantify the ingredients in the sample. It features real-time, in-line measurements and reports the fat, protein, solid non-fat, lactose, salts, total solids, water content, density, and freezing point of a milk sample to optimize quality control testing.

A handful of work has been done in soda, oil, blood glucose, and ammonium gas sensing on the basis of dielectric characterization using FMCW radars within the lab group and outside of the lab group. Considering the cost of a one of these FMCW radars is less than $20 for a single unit [24], this could be an attractive price point for farmers and businesses. Because of its longer wavelength it has the capability of penetrating packaging and is not foreign to the food and beverage industry. Visible light spectroscopy does not have this capability if the packaging is solid and opaque. Some industries use FMCW radars for continuous fill tank level monitoring [25], [26]. Integrating into a currently used technology could also help minimize changes and introductions of new and existing processes. This technology has the potential for low-cost, real-time compositional analysis in a tank or on a production line as packaged milk products are conveyed. The exact limitations of the radar are currently an ongoing study.

FMCW radars are not unheard of in industries thanks to their continuous non-invasive scanning abilities. Multiple startups have been developed around this technology for tank level monitoring and are easily searchable online. They are typically used to monitor the amount of liquid in a tank as they provide accurate information about its levels, volumes, and fill rates. The application of using them to monitor liquids for quality assurance and quality control purposes is becoming purposes is an ongoing study. Using them in the millimetre wave region is a new emerging trend. In fact, there is a patent belonging to the University of Waterloo for food quality testing using millimetre waves as found in [27].

Some work regarding oil characterization has been tested at the University of Waterloo using said patented technology [28]. In this work, different types of oil were tested on whispering-gallery-mode resonators. After the initial characterization using the resonators and proving that the different oils are differentiable, the oil samples were tested using the SOLI. The successful demonstration using the SOLI to differentiate oils on a plastic petri-dish can be found in [29] as a YouTube video. Studies were also conducted for glucose level sensing and detection using the DAK-TL and the SOLI in [30], [31], [32]. The SOLI was also demonstrated to work as a low-cost gas sensor in [33]. The work performed by

previous lab members of the group have laid a foundation for the dielectric characterization using the SOLI radar.

The characterization of milk was examined in a lower band from 1.8 GHz to 4.6 GHz and reported in [18] by our research group using resonance measurements. It was conducted with two prototypes of complementary split-ring resonators (CSRR) with a two-port VNA for a proof-of-concept to differentiate between 4 different milk fat concentrations. Examining these resonance sensors are outside the scope of this thesis, but it can be noted that prior work has been done with them to ensure differentiability but not at mm-wave frequencies.

Although not liquid characterization but rather general material characterization, Hegazy's work [34] features the usage of a 77 GHz Texas Instruments FMCW radar. The work describes different theoretical models that could be used to extract the permittivity and loss tangent of high-density polyethylene and polymethyl methacrylate of different thicknesses. The two methods used to extract the dielectric properties are the maxima and minima method and the curve fitting method. The work also features a method to calibrate the FMCW radar to de-embed the any magnitude and phase errors from the surrounding environment. This is done by measuring the reflection from the air, the reflection from a metallic sheet, and the reflection from a metallic sheet shifted backwards to account for the phase shift. Investigating the calibration procedure presented and adapting it for this project may be promising.

In 2023, Chen et al. [35] developed a convolution neural network to classify different liquids with a 77 GHz radar in real-time. The work conducted by the team is similar in nature to that of this thesis, but they tackle a less challenging problem. They test 8 different liquids with 97.33% accuracy: water, laundry detergent, oil, cola, milk, liquor, vinegar, and NaCl. In nature, these 8 liquids are very different from one another and in theory it would be expected to be much easier to differentiate these different solutions. Chen [35] mentions that the algorithm becomes confused when 2 or more liquids are similar. It is possible that hazardous liquids will leak their way into a finished product as they describe, but as it does so the chemical composition of the liquid is also changing, yet provided no evidence of solving or addressing the mixing problem.

This chapter outlined the importance of butterfat in the dairy industry, the need for a low-cost, real-time, non-invasive monitoring solution, and a literature survey on industry standard instrumentation and current state of the art. The proposed technology features a low-cost stand-alone system or one that could be integrable with current FMCW radar technologies used in a facility in tanks that has the capability of characterizing liquids non-invasively. Although milk fat detection is the main focus for now, theoretically this

should also work for foreign liquids and adulteration as the dielectric fingerprints would change. The next chapter two chapters will discuss the physics in the sensing parameter of the proposed technology. Since there is no accessible instrument to determine the exact milk fat percentage of a milk or cream sample, the percentage printed on the cartons were taken as is.

Chapter 3

Dielectric Characterization Work

This section delves into the physics behind the dielectric properties of liquids in general and further applies it to milk and cream. The Debye model is discussed. A discussion on mixing models can be found in Appendix A. Dielectric characterization is used in this project for initial testing to lay a ground truth. The end goal is to perform a similar characterization analysis using a mm-wave radar, but the sample should first be validated to ensure it is indeed differentiable before proceeding with the radar systems. Characterization measurements are presented.

3.1 Theory of Dielectrics

3.1.1 Behaviour of Dielectrics

Consider a parallel plate capacitor with a linear dielectric slab consisting of polar molecules inserted between the two plates. Upon applying an electric field, the molecules will experience a torque that forces them to rotate themselves in the direction of the electric field. The electrical energy passing through the capacitor is stored in the polarized molecules. The easier the molecules are to polarize, the higher the dielectric constant. This can be seen in Equation 3.1 for linear dielectrics [36]

$$\vec{P} = \epsilon_0(\epsilon_r - 1)\vec{E} \tag{3.1}$$

where ϵ_0 is the permittivity of free space and ϵ_r is the relative permittivity. In a time-

varying field, the relative permittivity will be a complex quantity and is defined as

$$\epsilon_r^* = \epsilon' - j\epsilon'' \tag{3.2}$$

The real part of the relative permittivity is known as the dielectric constant which quantifies the ability of a material to store electrical energy. The imaginary part of the relative permittivity is known as the dielectric loss which describes how much energy is dissipated in a material from internal friction, heat, or other physical processes [36]. A high dielectric loss will lead to a large power loss.

It is important to note that although the real part of the relative permittivity is called the dielectric constant, the value varies with frequency. As previously mentioned, both the dielectric constant and dielectric loss are frequency dependent. This is due to the different types of polarization mechanisms being observable at certain frequencies as well as in certain materials. At zero frequency (DC), the measured dielectric constant is called the static dielectric constant. As frequency tends towards a relatively large number, the dielectric constant will tend towards a constant value called the high frequency dielectric constant. At optical frequencies, the high frequency dielectric constant will approach the index of refraction of the material squared [37]. Both values depend on the relaxation time of a material [38].

$$P_T = P_{int} + P_{dip} + P_{ion} + P_e \tag{3.3}$$

where P is a sum of all the dipole moments. For the purposes of this project, the interfacial polarization will not be examined. According to Figure 3.1 adapted from [39] of the general polarizability vs frequency plot, dipolar polarization is observed until about 10^9 Hz. Likewise, ionic polarization is observed until about 10^{13} Hz. In the microwave region of the electromagnetic spectrum, only the dipolar/orientational, ionic/atomic, and electronic polarizations can be observed.

Dipolar polarization or orientational polarization is inherent to molecules with a permanent dipole such as H2O or molecules with an induced dipole [38]. The equation for dipolar polarization given in [38] is

$$\vec{P}_{dip} = \frac{Np^2}{3kT}\vec{E} \tag{3.4}$$

Examining the above equation if temperature increases, the polarization decreases, and the static dielectric constant decreases. In fact, this is the only type of polarization

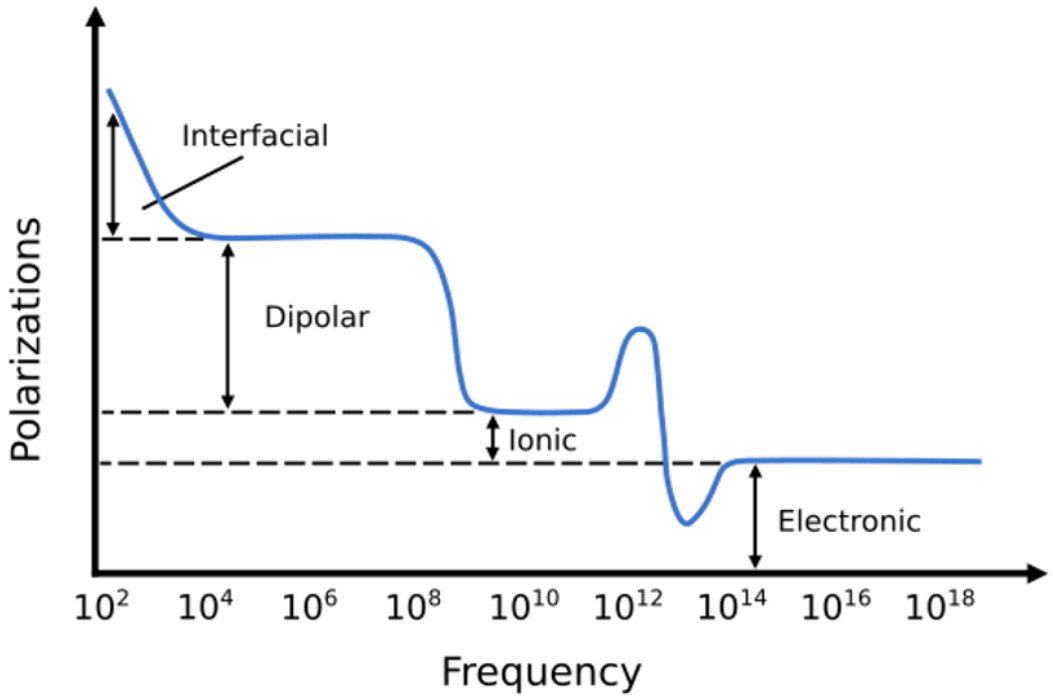


Figure 3.1: Plot of polarizability vs frequency. Adapted from [39].

that shows a temperature dependence in its equation where the polarization is inversely proportional to temperature [40]. Relaxation behaviour is also prevalent in this type of polarization. When an electric field is applied, the dipoles will orient themselves in the direction of the field. The time that it takes for them to rotate from an excited state to a new equilibrium or random distribution is called the relaxation time. The relaxation behaviour is an important concept that will be seen later when discussing the Debye equations. As frequency increases to about 10^9 Hz [39], the molecules cannot keep up with the rapidly oscillating field and will become nearly motionless in respect to rotations. The polarization will then no longer contribute to the total polarizability at this frequency and higher ones.

Ionic polarization applies to ionic crystals while atomic polarization occurs in polar covalent molecules [38]. Polar molecules are asymmetric [41] and have a permanent dipole [38]. Examples include salt and sugar. On the other hand, nonpolar molecules such as diatomic molecules are the opposite and only exhibit electronic polarization. Electronic polarization or optical polarization is inherent to all atoms because the applied field dis-

places the electrons of the material [41]. Both ionic/atomic and electronic polarizations can be described by resonance behaviour [38], [40] similarly to a mass on a spring. No equation for ionic/atomic polarization is given in [38], but electronic polarization is defined as

$$\vec{P}_e = N\left(4\pi\epsilon_0 R^3\right)\vec{E} \tag{3.5}$$

It is mentioned in [38] that a way to obtain the value of the ionic/atomic polarizability would be to subtract the dipolar and the electronic polarizabilities from the total polarizability. The three classes of dielectrics and details of each have been summarized in Table 3.1 from [41]

Table 3.1: Classes of dielectrics and their observable polarization mechanisms

Substance	**Observable Polarization Mechanisms**
Dipolar	1. Dipolar/orientational polarization 2. Ionic/atomic polarization 3. Electronic polarization
Polar	1. Ionic/atomic polarization 2. Electronic polarization
Non-polar	1. Electronic polarization

3.1.2 Debye's Equations and Ionic Conductivity

In deriving the Debye Equations, ionic/atomic polarization was neglected and only the electronic and dipolar polarizations were accounted for. The derivation can be found in [38] where the polarization build-up in a dielectric from 0 to some saturated polarization when an electric field is applied is examined. An alternative derivation can be found in [41] with a similar approach, but a decay function is used instead. Regarding the derivation in [38], it was assumed that the molecules are spherical, and the dielectric behaved as a linear

one. It was also assumed that an electric field of the form $E = Ae^{j\omega t}$ was applied. The resulting equations describe the behaviour of polar dielectric as a function of frequency and are defined as

$$\epsilon' = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + (\omega\tau)^2} \tag{3.6}$$

$$\epsilon'' = \frac{(\epsilon_s - \epsilon_\infty)\omega\tau}{1 + (\omega\tau)^2} \tag{3.7}$$

where ϵ_s is the static dielectric constant and ϵ_∞ is the high frequency dielectric constant. The two equations can be written together as a ratio to define the loss tangent

$$\tan\delta = \frac{\epsilon''}{\epsilon'} \tag{3.8}$$

Since the Debye equations are best used to describe relaxation behaviour found in dipolar and polar dielectrics, it will fall apart for nonpolar materials which only exhibit electronic polarization. Another aspect where the Debye equations will fall apart is when the material exhibits more than one relaxation time. Not many materials exhibit ideal Debye relaxation besides distilled water [38]. Because the Debye equations take into consideration dipolar and electronic polarizations, the model will start to break down at some high frequency. In [42], the single Debye model as introduced above is applicable for distilled water up to 50 GHz with a temperature ranging from 0°C to 30°C. It is mentioned that at 10 GHz there is an approximate 1% error, while at 50 GHz there is an approximate 5% error between measured and theoretical data.

The static dielectric constant, high frequency dielectric constant, and the relaxation time are all dependent on the "purity" of the water being measured. That is, with any concentration of impurity in the water, the parameters will be affected. The general trend seems to be as the static dielectric constant decreases, the relaxation time increases, and the high frequency dielectric constant increases as concentration increases. These trends will be further discussed in the literature survey where various papers are examined. It is mentioned in [43] that the solute disrupts the H-bond network of water when it is dissolved, therefore the static dielectric constant decreases regardless of the type of solute. This could be due to the water molecules bonding with the solute and forming hydration shells around them. With these hydration shells, the moment of inertia of the bound molecules will increase, decreasing the mobility of the free water molecules [44].

The relaxation time depends on temperature, shape, and axis of rotation [38]. It is possible to look at relaxation time both macroscopically and microscopically. The relaxation

time used in the Debye model is the macroscopic relaxation time [38]. For microscopic relaxation time, the equation given in [38] is

$$\tau_m = \frac{4\pi\eta a^3}{kT} \tag{3.9}$$

where η is the viscosity of the liquid and a is the molecular radius. Since there is a linear relationship between the relaxation time and viscosity, we can see that with an increase of viscosity there will be an increase in relaxation time. With an increase of concentration in a solution, the relaxation time as well as the dielectric loss will increase because there will be greater internal friction experienced by rotating molecules.

Some notable features of the ϵ'-ω and ϵ''-ω curves from [38] are that the maximum dielectric loss occurs at $\omega\tau = 1$ or $\omega = \frac{1}{\tau}$ and has a value of $\frac{\epsilon_s - \epsilon_\infty}{2}$. At this frequency, the dielectric constant takes on the value $\frac{\epsilon_s + \epsilon_\infty}{2}$. This frequency is called the relaxation frequency and may be denoted ω_p. With an increase of concentration in a solution, the static dielectric constant will decrease, and the high frequency dielectric constant will increase, giving the maximum dielectric loss attained a smaller value. Looking at the derivation of the Debye equations in [38], the polarization build-up and rate of change of polarization are

$$P(t) = P_\infty\left(1 - e^{-\frac{t}{\tau}}\right) \tag{3.10}$$

$$\frac{dP(t)}{dt} = \frac{1}{\tau}P_\infty e^{-\frac{t}{\tau}} \tag{3.11}$$

respectively. With an increase in relaxation time, there will be a faster polarization build-up. Likewise, there will be a slower polarization build-up with a decrease in relaxation time. The rate for the polarization to decay will be analogous to the rate of build-up. If there is an increase in concentration, there will be a quicker polarization decay, so the relaxation frequency will occur sooner.

An important consideration that needs to be accounted for aqueous ionic solutions is the fact that they are electrolytic. That is, if salt is dissolved into distilled water, the salt dissociates into ions and the solution becomes conductive. If sugar is dissolved into distilled water, the sugar does not dissociate into ions and the solution remains non-electrolytic and no ionic conductivity occurs. Ionic conductivity arises when there are free flowing ions present in a solution. There is a Debye model for aqueous ionic solutions that has been slightly modified to account for the ionic conductivity of a solution from [45]

$$\epsilon_r^* = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 - i\omega\tau} + \frac{i\sigma}{\omega\epsilon_0} \tag{3.12}$$

which can be separated into real and imaginary components giving the original dielectric loss an extra term

$$\epsilon'' = \frac{(\epsilon_s - \epsilon_\infty)\omega\tau}{1 + (\omega\tau)^2} + \frac{\sigma}{\omega\epsilon_0} \tag{3.13}$$

The first term of the dielectric loss is the loss due to the rotation of the dipoles and the second term is the loss due to ionic conductivity [46]. The behaviour of the dielectric constant and loss of an aqueous ionic solution can be illustrated in Figure 3.2 adapted from [38]. The loss tangent equation (Equation 3.8) can be modified accordingly.

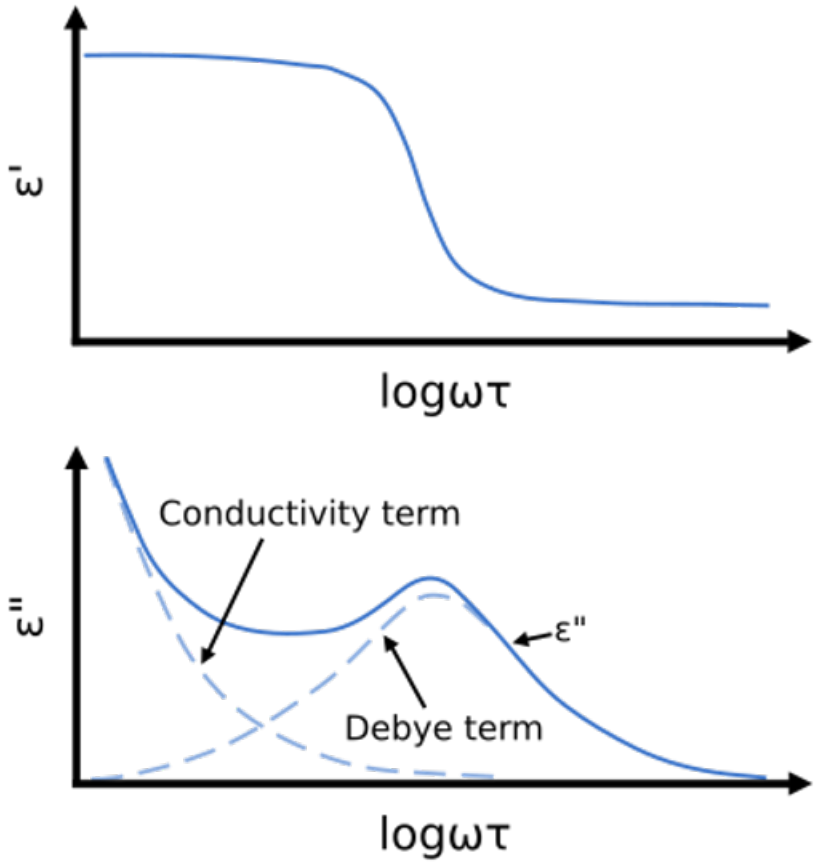


Figure 3.2: Plot of Debye Model with ionic conductivity term. Adapted from [38].

The conductivity, σ can be approximated to be equal to σ_s which arises due to the ionic conductivity of a solution at DC. In these cases, the Debye-Falkenhagen effect is assumed

to be negligible. The effect describes the increase in the conductivity as frequency increases [47] and arises due to AC currents [48]. There have been several models that have been developed to explain the Debye-Falkenhagen effect, but it has been a work in progress [49].

Mathematically, the loss due to ionic conductivity approaches zero as frequency tends towards infinity, so the term only greatly affects the dielectric loss at lower frequencies. An explanation for this phenomenon is provided by Gadani et al. [50] that seems to be describing the Debye-Huckel theory. Similar explanations are also found in [44], [49]-[51]. The Debye-Huckel theory describes the behaviour of strong electrolytes [47] such as salt dissolved in distilled water. The Debye-Huckel theory first introduced the concept of an ionic atmosphere in 1923 which has a similar idea to hydration shells [47]. The ionic atmosphere is defined as a spherical cloud of charge that describes the average distribution of ions in a region of space. Two phenomena arise from this theory called the asymmetric and the electrophoretic effects and together they lower the mobility of ions in a solution. The asymmetric effect describes the displacement of the central ion and the surrounding ionic atmosphere due to the presence of an applied electric field [47]. On the other hand, the electrophoretic effect describes the retardation force that is exerted on the central ion as it moves in the opposite direction of the ionic atmosphere [47]. There will always be a delay or lag of the ionic atmosphere as it tries to return to its equilibrium position.

As frequency increases, the asymmetric effect is reduced, and the central ion oscillates more quickly than at a lower frequency [44], [45]. The central ion eventually becomes motionless at some high frequency and the loss due to ionic conductivity term becomes negligible. In [38], it is mentioned that the conductivity of the solution does not decrease and only the loss due to ionic conductivity decreases. Interestingly, as frequency increases, the dielectric constant decreases, and the conductivity increases [38], [52]. Murphy and Morgan [52] mentioned that the increase in conductivity with frequency may be due to fewer restoring forces acting on the ions. Chandra and Bagchi [49] reported a similar idea, but about the Debye-Falkenhagen effect instead. Perhaps understanding this process will be important in understanding the variation of the high frequency dielectric constant with concentration.

3.1.3 Theory Applied to Salt and Sugar Solutions

Before delving into the components of milk and how the effect of butterfat content affects the dielectric properties of milk, the simpler topic of salt and sugar solutions will be discussed. Since milk has many constituents, it would be good to understand how the different constituents affect the dielectric properties. Thus, a common and simple solution

to begin with is salt and sugar water, which combines all the theory introduced so far. These are also two common components found in most foods including milk.

Salt water is commonly studied to develop radiometers for remote sensing of seawater [53]-[54]. It is also commonly studied to monitor water quality for agricultural activity [44]. To develop an accurate model for salt water, frequency, temperature, and salinity will need to be accounted for [53]. Stogryn [53] used regression methods to derive an accurate model for salt water at microwave frequencies. Unfortunately, due to the local interaction between molecules it is difficult to derive a physical model without regressions. Xu Yuan et al. [55] attempted to derive a physical model for saltwater at radio frequencies but achieved satisfactory results.

Another model based on Stogryn's regression models is the Klein and Swift model [54]. The main difference between the two models is that Stogryn's model is dependent on normality and a separate equation for normality based on salinity is given, while the Klein and Swift model only requires salinity [45]. Gadani et al. [45] compared the Stogryn and Klein and Swift models and determined that both model the dielectric loss of salt water well. However, the Stogryn model is better at modelling the dielectric constant above 0.9 GHz while the Klein and Swift model is better below 0.9 GHz. It is also reported in [45] that over the tested frequency range 200 MHz to 1.4 GHz that the dielectric constant decreases as salinity increases, and the dielectric loss of pure water increases as frequency increases, while the dielectric loss of saline water decreases as frequency increases. The decrease of dielectric loss of saline water can be modelled by the aqueous ionic Debye equation provided in the previous section and is explained due to the ionic conductivity of the solution. It can be noted that the conductivity increases as the concentration of salt increases in the solution.

An interesting comparison between salt water and sugar water is made by Meng et al. [51]. Concentrations of 5-25% salt in water and concentrations of 5-80% sugar, both in increments of 5%, are prepared and tested over the frequency range 200 MHz to 20 GHz. It is mentioned that the reason why 25% is the highest concentration of salt used is because the solution will become saturated by that concentration. Any extra salt added to the solution will give similar results. It appears that sugar water becomes saturated with a higher concentration of sugar than sugar. It can be seen in the graphs of dielectric loss vs frequency of various sugar concentrations in [51] that the maximum dielectric loss decreases with an increase in concentration. The occurrence of the peak also happens at an earlier relaxation frequency with the concentration of sugar increases. The conclusion of [51] is that an increase in the concentration of salt or sugar will lower the dielectric constant and increase the dielectric loss.

The work presented by these authors give inspiration to the regression modelling that may be performed given the dielectric properties are properly extracted from the radar measurements. A multivariable regression model could be implemented for the milk and cream data captured by DAK-TL as long as the concentration per component or constituent is known, but would serve little purpose for the end goal of this project. Albeit its complexity, if the dielectric properties were extracted from the radar, similar to Hegazy's work in [34], then the multivariable regression model would be more useful. The approach of simply developing a linear regression model based on the extracted radar data would also work, but is left for a future project and study.

Work performed for liquid characterization in the mm-wave region is an emerging field. Even at higher frequencies, the same physics should apply. Although different polarization processes and resonant frequencies occur, the concept of permittivity and losses can be extendable to mm-wave frequencies. With a first principles' understanding of the behaviour of dielectric properties of simple solutions, the knowledge can now be applied to more complex solutions such as milk and cream.

3.2 Components of Milk

There is great potential in using dielectric characterization to analyze the composition of milk. Because the behaviour of dielectric properties are heavily influenced by the chemical and physical properties of a sample, this makes it an attractive method. Milk is comprised of many constituents, where each one could greatly affect the dielectric properties of a sample. Raw milk is an inhomogenous solution with about 12.6% solids immersed in 87.4% water [9]. Out of these total solids, 3.6% of it is milk fat while the remaining constituents include caseins, whey proteins, lactose, and minerals [9].

Milk fat itself consists of 97-98% triacylglyercides (TAGs) and its remaining components are monoacylglycerol, diacylglycerol, cholesterol, phospholipids, and free fatty acids (FFA) [56], [57]. Kudra et al. [58] found no prior literature that measured the dielectric properties of pure milk fat and were the first. As a result, the data measured by the authors were determined to have similar permittivity and loss values to edible oils such as corn oil and sunflower oil. These experimental results are expected since edible oil is a mixture of triacylglycerides as well. Thus, milk can be treated as an emulsion of milk fat in water similarly to that of oil-in-water [56], [58], [59]. To validate the behaviour of the gathered experimental permittivity and loss data, an attempt to model milk as a two component solution of butterfat and water will be presented in a later section.

The quality of milk and its composition are subject to the age, nutrition, and species of a cow, as well as the lactation period, grazing, and seasonality [60]. These factors would impact the percentage per constituent would make up in a milk sample and no two cows will produce the exact same quality milk. O'Connor and Synnott [61] measured the dielectric properties of butter throughout the year showed that seasonality affected the values. They reported that the dielectric constant was highest in mid April and lowest in mid February. They also reported a sawtooth trend in the fluctuations of the dielectric constant throughout the year. Based on their findings, the percentage of total solids had little correlation to the fluctuations of dielectric constant with seasonality.

Overall, due to its many constituents, milk is a complex electrolyte solution. The behaviours of these solutions were not intensively studied until about 30 years ago [62]. Wright [62] explains that increasing the concentration of a solute leads to a decrease of a solution's relative permittivity because of the increased number of solvent molecules bounded to the solute. Wright [62] mentions that it is unclear whether this effect only arises due to bounded solvent molecules or if there are addition factors at hand and is a controversial topic. Understanding how the complex electrolytes behave in a solution would be an important first step in analyzing the solution as a whole. Unfortunately, it is outside the scope of this project.

3.2.1 Behaviour of the Dielectric Properties of Milk

Since fats generally have a long nonpolar hydrocarbon chain, there is little interaction with an electromagnetic wave, making them dielectrically inert [63]. In both fats and oils, only electronic polarization processes should be observed and should exhibit very little loss [64]. In the experiment conducted by Gouw and Vlugter [65], the triglycerides used needed to be free of any moisture for accurate measurement results. The addition of water to oil causes the dielectric constant to spike upwards because water has a high dielectric constant and exhibits more polarization processes than oil. Lizhi et al. [64] examined the effect of increasing moisture content on the dielectric constant of corn oil. A regression model to determine the dielectric constant of corn oil based on moisture content was made. It was mentioned that by monitoring the changes in the dielectric constant of oil during the production process, the moisture content of a sample may be able to be determined.

Determining the butterfat content of a sample is the first step towards implementing a radar for the compositional analysis of milk, because it is the most variable factor and one of the major constituents in the composition of milk. In the study conducted by Nunes et al. [66], homo, low-fat, and skim milk were purchased from a local supermarket and tested

over the frequency range 1 GHz to 20 GHz. Over this frequency range, with an increase in fat content, the dielectric constant and dielectric loss both decreased. In addition, the homo milk was diluted with water and measured. In comparison with the low-fat milk data, the dielectric constant and loss changed more rapidly with an increase of water than with an increase in butterfat. This suggests that the minor constituents and the non-fat components of milk play an important role in decreasing the dielectric constant or dielectric loss [66]. This also suggests that the fat content of milk can be estimated relative to the carbohydrate and protein content of a sample [66].

Zhu et al. [57] conducted a study using raw whole milk and raw skim milk over the frequency range 20 MHz to 4500 MHz. Over this frequency range, with an increase in butterfat, the dielectric constant decreased, and the dielectric loss decreased, similarly to the study conducted by Nunes et al. [66]. The decrease of the dielectric loss where the contribution due to ionic conductive losses dominates is due to a lack of mobility of the free-flowing ions caused by an increase in fat. It was determined that there was a linear relationship between the butterfat content and the dielectric properties. Using linear regression models, it is possible to determine the butterfat content from the dielectric constant and dielectric loss at a given frequency and temperature [57]. It was found that calculating the fat content from the dielectric loss of a sample gave a larger error than calculating from the dielectric constant. The error grew for predictions with less than 1% butterfat.

In this chapter, the basics of dielectric characterization using scattering parameters and the hardware used for the characterization will be discussed. There are multiple conventional characterization techniques as mentioned in [4]. Only the open-ended coaxial probe technique and mm-wave radars are explored in more detail for this project. There is little work done in literature using mm-wave radars for dielectric characterization as it is not a conventional technique.

3.3 Hardware

3.3.1 Scattering Parameters

There are several different kinds of parameter representations that can be used for characterization of microwave devices. Some examples include the transmission (ABCD) matrix, impedance and admittance matrices, hybrid parameters, and scattering parameters. Each set of parameters is best used different scenarios and schemes. The definition of scattering

parameters (S-parameters) are important for understanding the fundamentals of the open-ended coaxial cable and radar measurements. Consider a 2-port network as shown in Figure 3.3 with Port 1 (input) on the left and Port 2 (output) on the right of the device under test (DUT). The normalized values a and b refer to the incident power wave and reflected power wave respectively. The 2x2 matrix is called the scattering matrix ($[S]$ matrix) and is defined as a ratio of these normalized power waves with terminated transmission lines equal to the line channel impedance [67]. The generalized equation for a 2-port network can be written as

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}. \tag{3.14}$$

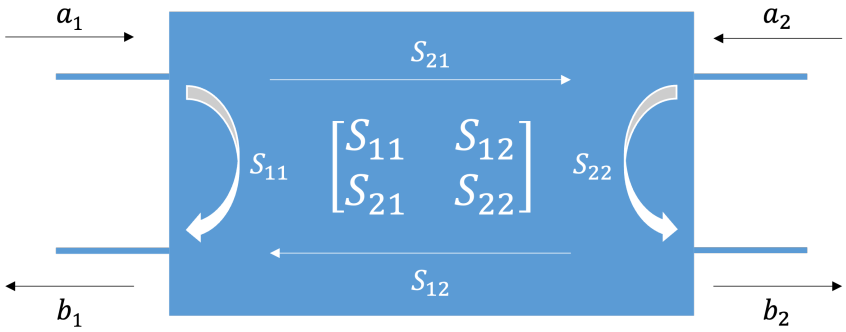


Figure 3.3: Scattering parameters of a two-port network.

The subscript in S_{ij} reads from Port j to Port i. Thus, S_{11} is the reflection coefficient seen at Port 1 when Port 2 is match terminated. S_{12} is the transmission coefficient from Port 2 to Port 1 when Port 2 is match terminated. They are complex-valued and comprise of a magnitude and phase, but most applications find the magnitudes to be of more interest. They are typically measured using a VNA, though it is possible to measure them through alternative methods. Measuring S-parameters are useful because they describe a linear relationship between the input and output signals. Unfortunately, they are not cascodable, that is,

$$[S_{AB}] \neq [S_A][S_B]$$

unless they are first converted into scattering transfer parameters, or T-parameters, or ABCD ones and then back again. This is important for de-embedding the open-ended coaxial probe that is done behind the scenes when making the dielectric measurements. It can be further applied to other de-embedding applications though. The S-parameters can

be summarized as follows:

$$S_{11} = \left.\frac{b_1}{a_1}\right|_{a_2=0} \quad \text{Input reflection coefficient}$$

$$S_{12} = \left.\frac{b_1}{a_2}\right|_{a_1=0} \quad \text{Transmission from Port 2 to Port 1}$$

$$S_{21} = \left.\frac{b_2}{a_1}\right|_{a_2=0} \quad \text{Transmission from Port 1 to Port 2}$$

$$S_{22} = \left.\frac{b_2}{a_2}\right|_{a_1=0} \quad \text{Output reflection coefficient}$$

The input and output reflection coefficients are commonly referred to as the input and output return losses. The above can be further generalized to an N-port network, similar to the other sets of parameters used for microwave characterization. In the case of the DAK-TL or open-ended coaxial probe, only the S_{11} coefficient will be utilized. This is further discussed in the next subsection with its advantages and disadvantages

When examining scattering responses detected by the radars, each antenna corresponds to a port. For example, if the radar setup comprised of 2 antennas (1 transmitter and 1 receiver), then the S_{11} would be the signal path from the transmitter back to itself. S_{21} would be the signal path from the transmitter to the receiver. The concept and ideas are similar to measuring materials using a two-port VNA, but the accuracy of the measurements would differ due to the hardware and the environment. Since the open-ended probe and VNA method are done in a closely controlled environment, measurements taken by the radar are affected by multiple uncontrollable factors which makes it a lot trickier to work with.

3.3.2 Open-Ended Coaxial Probe Fundamentals

Consider a single port network comprising of a coaxial line instead of the 2-port network from the previous section. The only measurable S-parameter in this case would be the S_{11} or reflection coefficient from the DUT. Instead of a DUT, say there is a dielectric specimen. The end of the coaxial line that comes into contact with the dielectric specimen has a sensor that measures the S_{11} [4]. This technique can be used for both solids and liquids, where in the case of a liquid the probe would be submerged [4]. A typical probe with a flange is illustrated in Figure 3.4, but there are many designs of flangeless probes

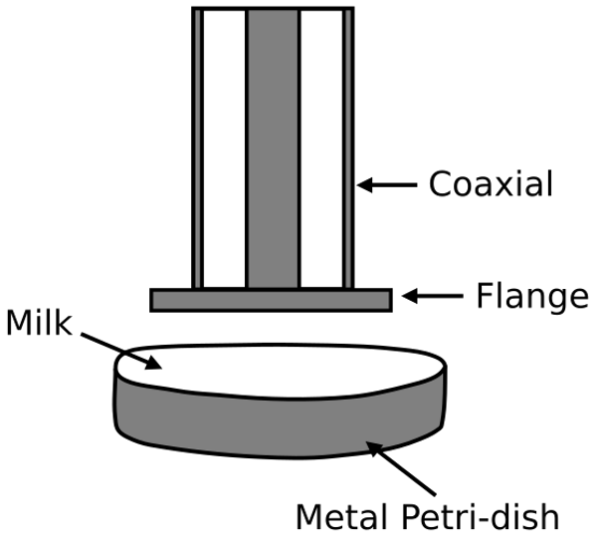


Figure 3.4: Coaxial probe with flange.

as well. The bead labelled in the diagram is a dielectric insert that shields the coaxial line from any sample debris.

Some advantages that come with the open-ended coaxial probe are that they are useful for broadband measurements [4], [38], require little to no sample preparation, and can be used for food, liquid, and other semisolid materials [38]. Albeit its advantages, this method has less precision than other characterization techniques especially at very low frequencies, very high frequencies, or for low permittivity and loss values [4], [38]. It is mentioned in [4] that more care must be taken to analyze fats and oils as they are low-loss materials. That being said, another care that must be exerted is in minimizing volumetric resonances. Since the size of the sensor varies considerably with frequency and wavelength, this means the penetration depth is greater with a larger sensor. This leads to volumetric resonances arising from the reflections at the boundaries of the sample [68], [69]. To reduce these boundary effects, an adequate sample size should be chosen.

It could be disadvantageous to use the open-ended coaxial probe method because only the S_{11} is measured [70]. Sometimes the reflection coefficient alone does not provide enough information about the sample under test and examining the transmission coefficient would be useful.

Pure water is a useful reference liquid for sensor calibration as it is the best characterized polar liquid [68]. Its dielectric properties can be simply modelled with a single-relaxation

Debye model described in the previous chapter. It is the reference liquid chosen for DAK-TL calibration because milk is mostly composed of water. More details of various reference liquids can be found in [68] such as ethanol, methanol, and ionic solutions.

In terms of measuring ionic solutions, there is a possible issue with electrode polarization that could cause inaccuracy and lesser precision. When measuring conductive samples, electrode polarization is a common effect that arises. Here, the ions in the solution float towards the electrode or sample surface to form an ionic double layer. At low frequencies, this masks the true characteristic of the material since there is a voltage drop across the double layer [71]. Because of the numerous factors that contribute to the issue such as the conductivity of a sample, temperature, structure and composition of an electrode, and any electrochemical interactions occuring [71], the issue becomes difficult to minimize. This was seen on the DAK-TL, however it may be extended to other open-ended coaxial probes in general. The DAK-TL probe is able to accurately measure 0.1M saline solutions or less without an issue. Improving their probes to address this issue is an ongoing research and development initiative at SPEAG.

Temperature also plays a profound impact on the dielectric properties of a sample. Although this thesis does not delve into these temperature dependencies, a quick temperature study is presented in the next chapter. Another challenge would be minimizing the air bubbles that form on the liquid. These air bubbles can affect the precision and accuracy of the measurement. It may be possible that the bubbles are too small to be detected as well. Liquids with low surface tension like alcohol don't have this issue, but for liquids like water, it presents a greater challenge [68].

Dielectric Assessment Kit for Thin Layers

The Dielectric Assessment Kit (DAK) product lineup offered by SPEAG performs dielectric characterization of solids, semi-solids, and liquids with high precision. The products use the previously described open-ended coaxial probe technique and calculates the reflectance coefficient, or S_{11} parameter, of the material under test [4], [44]. An algorithm in the DAK software then uses the reflection coefficient to calculate the complex permittivity, loss tangent, and conductivity of the material [72]. This project uses the DAK-TL, and the main difference between the DAK-TL and DAK probe system is the fact that the DAK can measure larger volumes of liquids. This would improve measurement accuracy and minimize volumetric resonances caused by the boundaries of the sample. This project uses the DAK-TL system which has the capability of measuring about 10-12 mL of liquid in a small metal petri-dish or small and thin solid specimens such as printed circuit boards (PCBs). Materials under test are to be placed below the coaxial sensor probe on top of the

little platform that is software controlled. A block diagram of the system setup is shown in Figure 3.5. A coaxial cable is connected at Port 1 of the VNA to the DAK-TL probe. The system is connected to a laptop or PC to use the DAK software. The probe movement is fully software controlled. The PC and VNA both need to be on the same subnetwork.

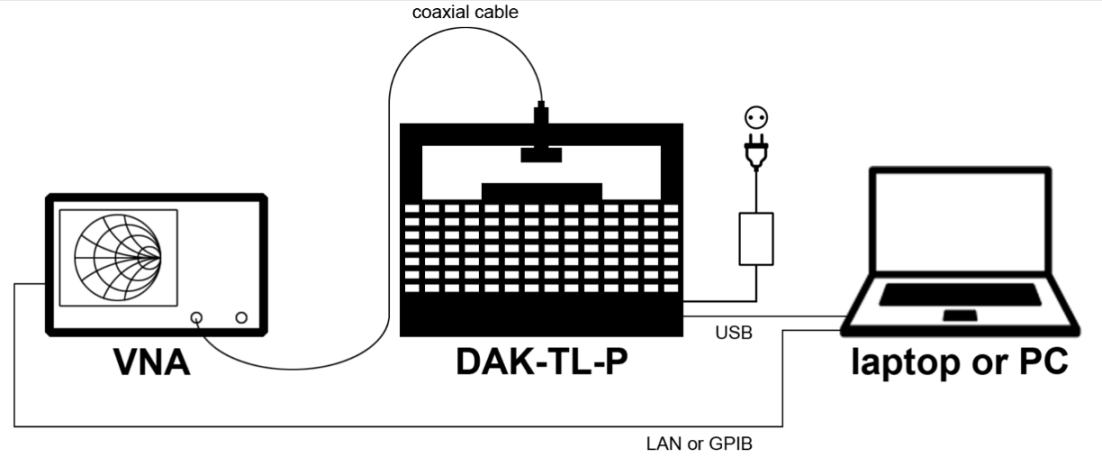


Figure 3.5: DAK-TL-P block diagram setup from [72].

For this project, the DAK-TL was used and an image of the setup can be found in Figure 3.6. The system was connected to a Keysight Power Network Analyzer N5227A/B. There are two possible probes for measurements over different frequency ranges: the DAK-TL 3.5 probe (200 MHz – 20 GHz) and the DAK-TL 1.2 E (5 GHz – 67 GHz) probe. The calibration, sample preparation, and measurements using both probes are presented in a later chapter. It should be noted that with the small volumes of liquid being tested, the DAK-TL will not be as accurate as if the samples were tested using the DAK. However, for the proof of concept in this project, it is sufficient.

A yearly calibration of the probe is done by SPEAG where several dielectric materials are measured and compared with their target values. The company takes into account the measurement uncertainty in the lossy materials measured and the effect of temperature deviations. When calibrating the DAK-TL, the probe itself is de-embedded from the measurements and the S_{11} is normalized to the reference plane at the end of the probe. Thus, the quality of the short circuit at the end of the probe is also assessed as part of the yearly calibration [72]. Unlike the Keysight VNA used, SPEAG did not publish an error model for the DAK-TL or have made their error models readily accessible. It would be a great asset to understand the possible errors that could arise from using the open-ended

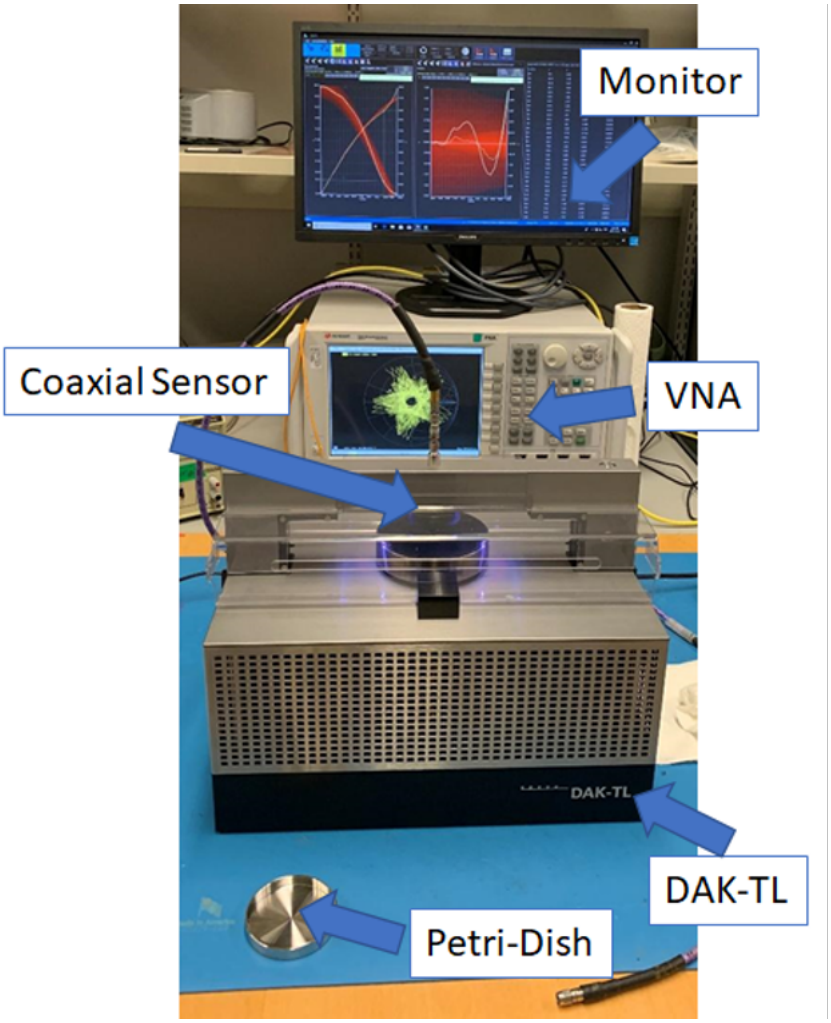


Figure 3.6: Setup of DAK-TL.

coaxial probe method along with the DAK Professional Handbook of best practices [72] when performing measurements and interpreting results.

De-Embedding

When the VNA is calibrated, it is usually calibrated such that the reference plane for measurement is set to the end of the transfer cable. Generally when performing measurements with the open-ended coaxial probe, the cable is connected to the probe, but

the measurement plane does not change from the end of the cable. In the characterization measurement, there would be introduced inaccuracy because the probe would then be considered part of the measurement results. Since S-parameters are not cascadable, it is not possible to divide the results between the sample under test and the probe, as the S-parameter will only give a measurement based on the whole picture. Thus, to have accurate characterization of the sample under test, the probe itself must be removed mathematically from this picture by using scattering transfer parameters (T-parameters) or an ABCD matrix. T-parameters are defined similarly to S-parameters [73]

$$\begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}. \tag{3.15}$$

but they are cascadable, that is they are multipliable

$$[T_{measured}] = [T_{probe}][T_{sample}].$$

The relationship between the S-parameter and T-parameter matrices from [73] can be written as

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \frac{T_{12}}{T_{22}} & \frac{T_{11}T_{22}-T_{12}T_{21}}{T_{22}} \\ \frac{1}{T_{22}} & -\frac{T_{21}}{T_{22}} \end{bmatrix} \tag{3.16}$$

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} \frac{-S_{11}S_{22}-S_{12}S_{21}}{S_{22}} & \frac{S_{11}}{S_{21}} \\ -\frac{S_{22}}{S_{21}} & \frac{1}{S_{21}} \end{bmatrix} \tag{3.17}$$

Using general matrix theory, it is well known that

$$[T][T]^{-1} = [T]^{-1}[T] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{3.18}$$

Thus, this property can be used to obtain the T-parameters of the desired sample under test

$$[T_{probe}]^{-1}[T_{probe}][T_{sample}] = [T_{sample}].$$

The calibration method will provide additional equations to solve for the T-parameter matrix of the probe. In the 1-port error model, there are 3 possible system errors that need to be accounted for. These system errors are the directivity errors (E_d) or leakage of the output signal from the directional coupler in the VNA, reflection tracking (E_r) errors,

and source mismatch (E_s) errors [74]. To obtain these error terms, 3 well-defined reflection measurements are needed. These standards are short ($\Gamma = -1$), open ($\Gamma = 1$), and load ($\Gamma = 0$). However, having perfect short-open-load calibrations are only possible in an ideal world. VNA manufacturers offset the imperfections and deviations as part of a standard model taken into consideration during the calibration process [74].

Once de-embedded, the T-parameter matrix of the measured sample can be converted to a S-parameter one using the relationship described above. This is not explicitly done in this project, but the DAK software may either use a similar approach or calibrate the VNA such that the measurement plane is at the end of the flange on the probe.

DAK-TL Calibration Procedure

There are two possible ways to calibrate the DAK-TL from the DAK software. The desired frequency range is first selected, then either the Open-Short-Load (OSL) method, which was used for this project, or being a Open-Load-Load is chosen. For the open part of the OSL procedure, the TL-controller was set to 400 N solid and the steel plate was used as the platform. For the short, the TL-controller was also set for 400 N solid but a copper tape was placed on the platform for the probe to come into contact with. The tape was twisted under contact with the probe to ensure that the short was good. To double-check the quality of the short and open of the calibration, ensure that the S_{11} reading is on the far left and far right of the Smith chart provided in the DAK calibration process for the short and open respectively. To calibrate for the load, the TL-controller is set to 3 N liquid. The distance of the probe was set to 4 mm, which is the distance between the platform and the bottom of the probe. This distance was determined after several repetitions with 10 mL of distilled water to ensure the bottom of the probe was fully submerged in the liquid without the liquid overflowing. The steel metal tray previously used should be changed to the black aluminum tray used for liquids. 20°C water can be selected as the material. As mentioned earlier, the temperature of the liquid does not play a large difference in the measurements because of the small volume being tested. The calibration is complete. It is good practice to measure distilled water again to compare with the target values in case there were air bubbles introduced in the calibration. Furthermore, it would be good to verify another material such as methanol to ensure the calibration is valid.

3.4 Experiments

From the background section, it was discussed it is theoretically possible to differentiate milk samples of different butterfat concentration. This section will aim to apply the theory and obtain experimental evidence that the samples are different from one another. Salt and sugar solutions are first measured for intuition to observe how the salt and sugar molecules in milk could affect its dielectric behaviour. After, the results for milk and cream from 0.10% to 35% are presented and discussed. A linear regression model is developed and shows the linear relationship between the amount of milk fat content in a sample and its dielectric constant or loss. A comparison of the data to predictive mixing models for suspended fat in water is further discussed in Appendix A.

The setup and calibration of the DAK-TL was previously presented. Both the DAK-TL 3.5 probe and the DAK-TL 1.2 E probes are used. For the low band, the instrument was calibrated according to SPEAG's recommended frequency and resolution settings for the DAK-TL 3.5 probe shown in Table 3.2. A smaller resolution is used at lower frequencies because the solutions being tested are more dispersive due to the dipolar/orientational polarization in water being the major contributor. For the high band, i.e. DAK-TL 1.2 E, it was calibrated over 5 - 67 GHz with 250 MHz steps. Note that measurements are most accurate over a smaller range of frequencies. It is possible to only calibrate around the 60 GHz range, but this test was to check and confirm the feasibility of implementing a radar at any point in the tested range. For instance, a few radars could be used in an array where each one is operating at a different frequency.

Table 3.2: Recommended calibration settings from 200 MHz to 20 GHz

Frequency Range	**Resolution**
200 - 295 MHz	5 MHz
300 - 6,000 MHz	50 MHz
6,250 MHz - 20,000 MHz	250 MHz

3.4.1 Sample Preparation

Salt and Sugar Solutions

Samples of 2%, 4%, 6%, 8%, and 10% of salt or sugar were measured using a Sartorius BP 2100 S scale with accuracy up to a hundredth of a gram. The uncertainty per measurement

was approximately ±0.02g. Samples were stored in plastic watertight Qorpak containers and sealed with Parafilm M as shown in Figure 3.7 to prevent evaporation from occurring. Windsor Iodized Table Salt and Redpath Special Fine Granulated Sugar were dissolved separately into Equate Distilled Water. To prepare a 2% solution, 1 gram of either salt or sugar was measured into a sample bottle. The bottle was then filled with distilled water until the scale read 50 grams. This procedure of preparing weight to weight concentrated solutions was repeated for the remaining concentrations of salt and sugar. The samples were then swirled until the solute was dissolved and sat at room temperature. None of the samples contained both salt and sugar in the solution. The weights of salt or sugar to the weight of distilled water to prepare the solutions are tabulated in Tables 3.3 and 3.4.

Table 3.3: Concentrations of prepared salt solutions

Concentration	**Molarity (mol/L)**	**Weight (g)**	**Weight of Distilled Water (g)**
2%	0.342	1.00	50.01
4%	0.684	2.00	50.01
6%	1.027	3.00	50.00
8%	1.369	4.00	50.00
10%	1.711	5.00	50.01

Table 3.4: Concentrations of prepared sugar solutions

Concentration	**Weight (g)**	**Weight of Distilled Water (g)**
2%	1.00	50.00
4%	2.00	50.01
6%	3.00	50.01
8%	4.00	50.01
10%	5.00	50.01

The concentrations were selected from 2-10% in increments of 2% to draw a clear trend between the data. If the differences between concentration were too small, the DAK-TL may not have the capability or resolution to differentiate the samples. If such a precise instrument is unable to pick up these slight variances, then the characterization with the radar system may be more difficult. Testing salt and sugar is not the end goal, but as part of a liquid or beverage, such concentrations are so small that it is likely that the radar will not have the sensitivity to detect it.



Figure 3.7: Salt and sugar solutions prepared with various concentrations.

Milk and Cream

A few cartons of cream were purchased from a local No Frills of 0.10%, 1%, 2%, 3.25%, 5%, 10%, 18%, and 35% butterfat. Small samples were prepared in watertight Qorpak containers and transported to the microwave lab in Waterloo for testing. There was no access to a machine or device that could identify the exact concentration of each constituent in a sample. The butterfat content of a sample was taken as printed on the carton. However, there could be slight variations with the butterfat content of a sample, for example, 5% cream may have a bit more or less than 5% butterfat in it. It can be noted that creams have added stabilizers to stabilize the fat content compared to milk, but this would have little effect on the dielectric properties.

To test the validity of the regression model when applying to a new data set, new samples were mixed by weight from the 5%, 10%, 18%, and 35% creams using a Sartorius BP 2100 S scale to create 6%, 8%, 12%, 16%, and 20% creams. To obtain 6% cream, 30g of 5% cream and 7.5g of 10% cream were measured separately and then mixed. The measured values are displayed in Table 3.5 and the uncertainty per measurement was about ± 0.02g.

3.4.2 Characterization of Salt and Sugar Solutions

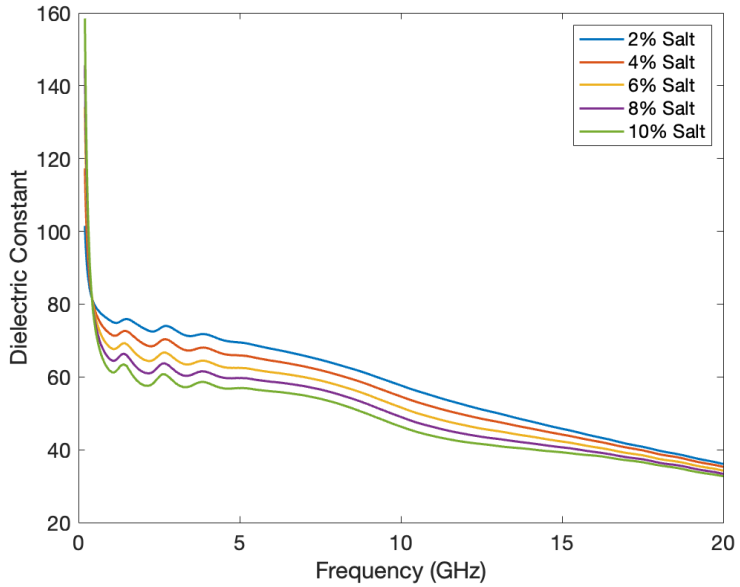
Each sample was tested 2 times and the averaged dielectric constant and loss results for 2-10% concentrations of salt and sugar are shown in Figure 3.8 and Figure 3.9. In Figure 3.8, small oscillations as well as the spike towards infinity at low frequencies can be observed.

Table 3.5: Concentrations of prepared mixed creams.

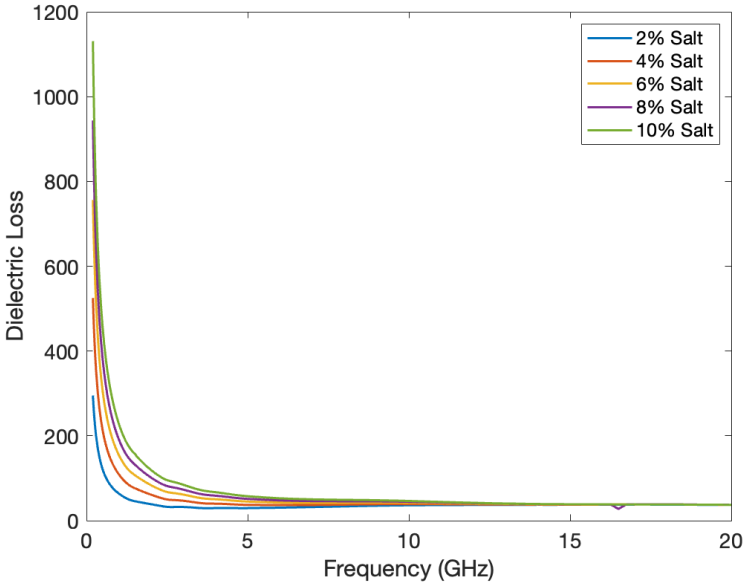
Mixed Percentage	**Sample 1**	**Required Weight (g)**	**Measured Weight (g)**	**Sample 2**	**Required Weight (g)**	**Measured Weight (g)**
6%	5%	30	30.02	10%	7.5	7.51
8%	5%	16.8	16.80	10%	25.2	25.21
12%	10%	36	36.00	18%	12	12.01
16%	10%	10.5	10.50	18%	31.5	31.31
20%	18%	32.54	32.55	35%	4.34	4.35

The large spike was caused by the electrode polarization that arises from conductive solutions. Because the probe can only measure saline solutions up to 0.1M, the spike is more evident in the measured data because of the greater than 0.1M concentrations. As the sugar solutions are not conductive, these spikes are not visible at low frequencies. The sugar solutions have small oscillations similar to the salt solutions, which further suggests these are due to volumetric resonances. The resonances arise from the reflections at the boundaries of the sample, but these effects are negligible if a large enough volume of sample is used previously discussed in the hardware section. To improve the measurement accuracy and eliminate the resonances, the classic DAK system should be used instead, which has the capability of measuring liquids of larger volumes.

In both cases of the salt and sugar solutions, the dielectric constant is shown to decrease with an increase in solute concentration. When salt dissociates into distilled water, ionic conductivity arises at DC, aligning with general physics knowledge that the conductivity of a saline solution increases with an increase in salinity.

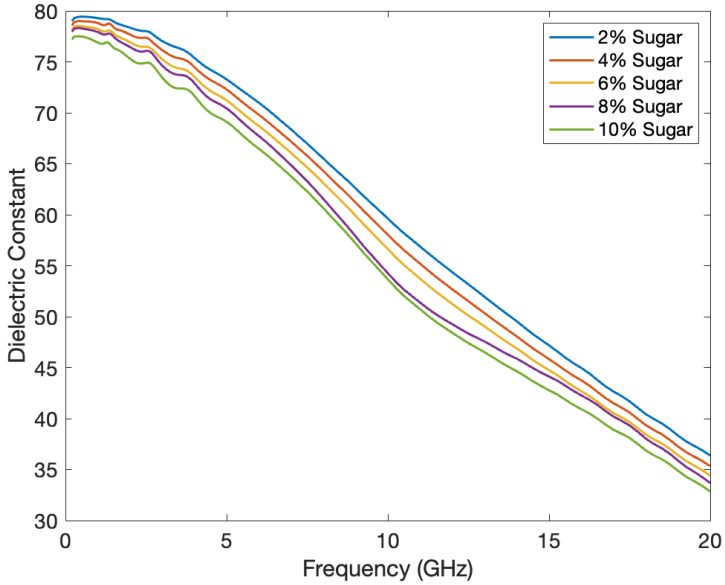


(a) Dielectric constants of salt solutions.

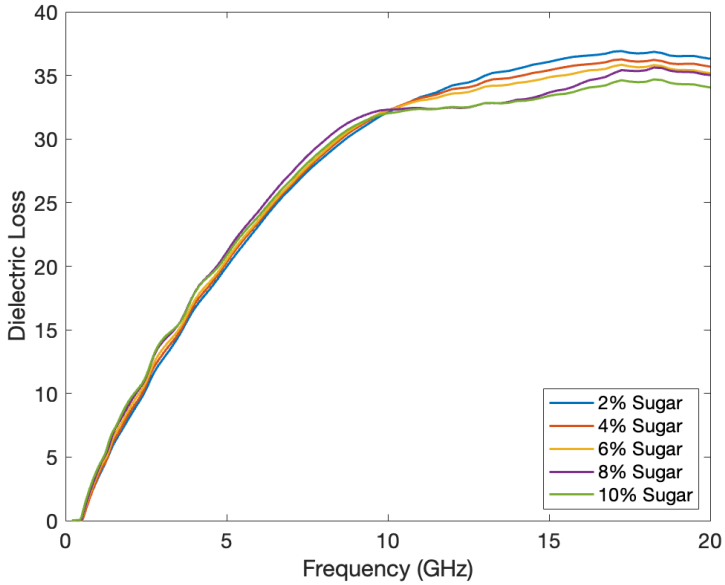


(b) Dielectric losses of salt solutions.

Figure 3.8: Dielectric characterization of 2-10% salt solutions from 200 MHz to 20 GHz.



(a) Dielectric constants of sugar solutions.

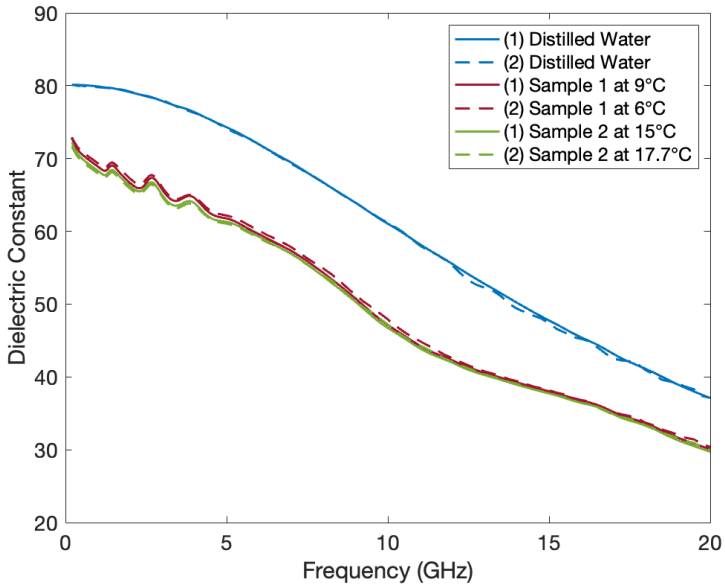


(b) Dielectric losses of sugar solutions.

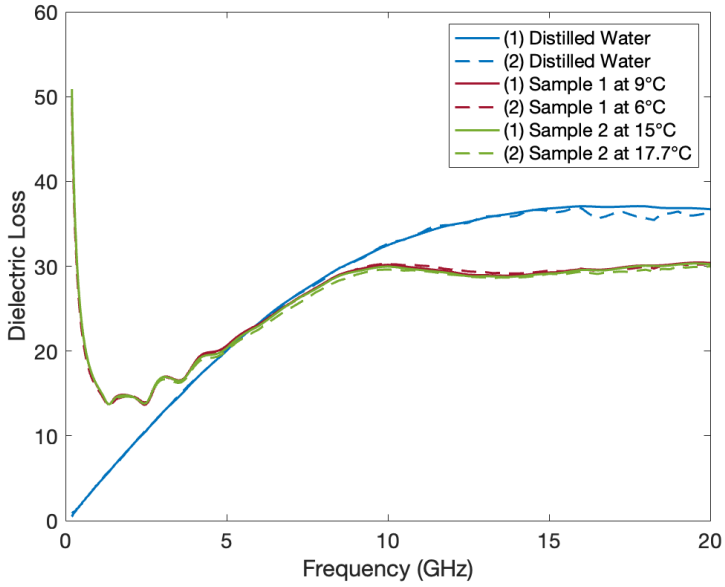
Figure 3.9: Dielectric characterization of 2-10% sugar solutions from 200 MHz to 20 GHz.

3.4.3 Temperature Study

Since the dielectric properties of a sample are temperature dependent, a brief temperature study was conducted on the DAK-TL before testing all the milk and cream samples. Distilled water and two 5% cream samples at 9°C and 15°C were tested, both tested two times for repeatability for each calibration. After, the calibration was wiped and re-calibrated with the exact same settings and procedure. This was done so that there would be confidence in the calibration of the machine and to eliminate the possibility of temperature affecting the results. Distilled water and the same two 5% cream samples, this time at 6°C and 17.7°C due to temperature fluctuations, were tested again. The results are shown in Figure 3.10.



(a) Dielectric constant temperature study.



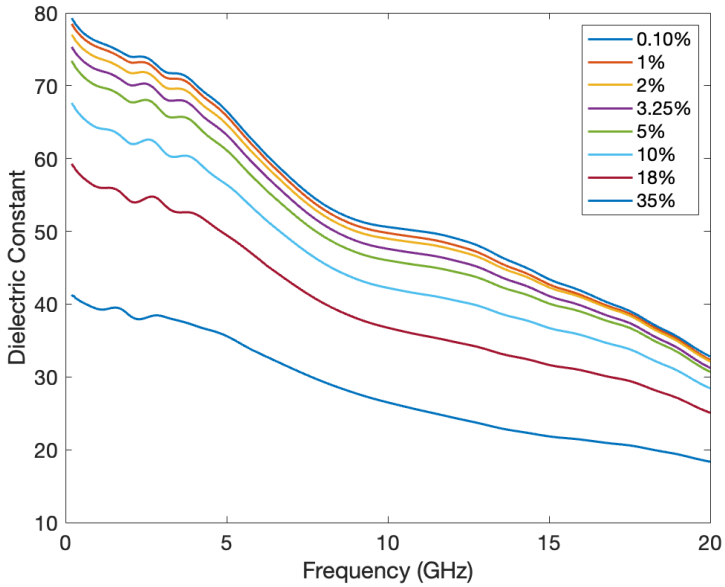
(b) Dielectric loss temperature study.

Figure 3.10: DAK-TL calibration comparison and temperature study.

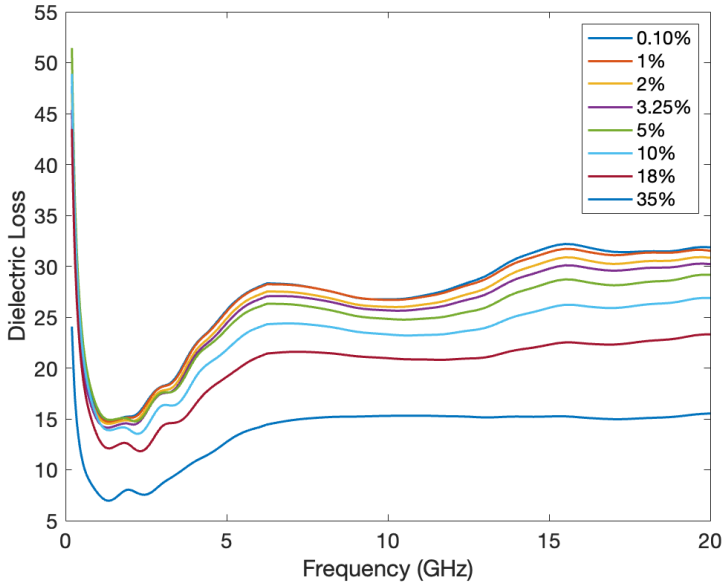
Interestingly, the dielectric properties between the 5% creams at different temperatures made very little difference to the measurements and were nearly unnoticeable as shown in Figure 3.10. Zhu et al. [57] tested the dielectric properties of milk at different temperatures over the frequency range 10 MHz to 4,500 MHz, and there should be quite a difference. The reason why the temperature did not make much of a difference to the measurements may be because the TL controller or probe heats up the sample during testing. Before testing the cream samples, they were on average about 6°C and were about 20°C after testing. Thus, the DAK-TL is not great for studying the temperature-dependent properties of liquids. Furthermore, based on this quick temperature study performed, the calibration that the samples are at about 20°C is valid, similar to the distilled water at room temperature.

3.4.4 Characterization from 200 MHz - 20 GHz

The measured dielectric constant and loss vs frequency of the cream products are plotted in Figure 3.11. The plots seem to agree well with the plots presented in previous literature such as [57] and [66] as a function of frequency. With an increase of fat content, the dielectric constant and dielectric loss both decrease. It can be noted that there appears to be a linear relation between the fat content and dielectric properties, which Zhu et al. [57] pointed out. Using this as a basis for the relationship between the two, the butterfat content may be predicted using a linear regression model, which will be discussed in a later subsection. There are some small oscillations in the dielectric constant and loss at low frequencies and some waviness in the dielectric losses with higher frequencies after 10 GHz, as previously seen with the characterization of the salt and sugar solutions over the same frequency range. Currently, the reason why the dielectric losses are wavy is unclear, but the results show that the trials and the samples are all repeatable. The contribution to the dielectric losses from the ionic conductivity of the sample decreases with an increase of fat concentration due to the fat particles immobilizing the free-flowing ions, as expected according to theory.



(a) Dielectric constant of milk and cream.

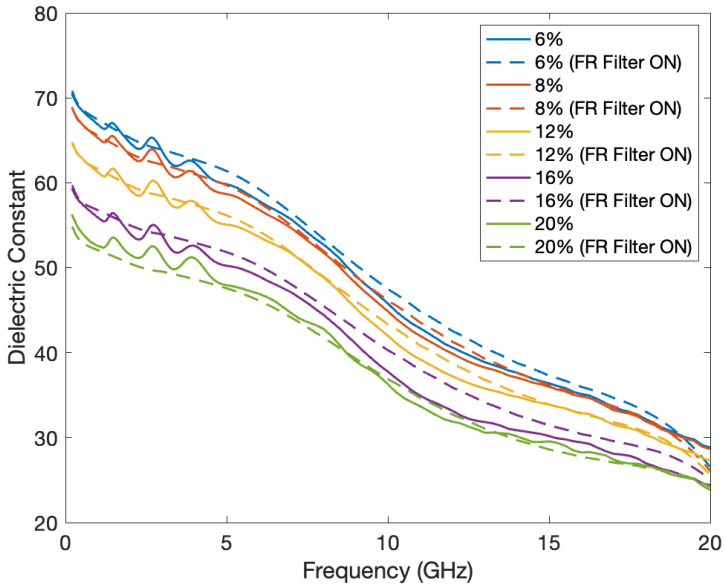


(b) Dielectric loss of milk and cream.

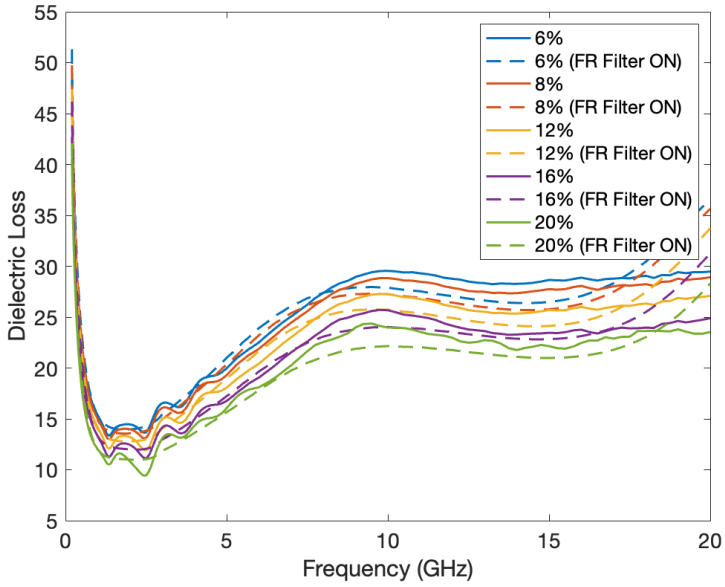
Figure 3.11: Dielectric characterization of milk and cream from 200 MHz to 20 GHz.

To investigate the oscillations at low frequencies, some of the filters in the DAK software were used and compared to the non-filtered data. For example, an attempt was made to smooth out the oscillations using the FR filter. FR stands for Flange Resonance, where the flange is the boundary component of the coaxial sensor that meets the sample being tested (Figure 3.12). Flange resonance can be observed in low loss materials at measured frequencies that have high permittivity such that the loss tangent is less than 0.3 [68], [69].

Upon turning the FR filter on, the dielectric losses changed more drastically than the dielectric constant as seen in Figure 8. Gregory [68] states that flange resonance is a common occurrence in water because of its low losses at low frequencies, and if it is not minimized in the calibration, there will be a greater effect other measurements. In the measurements performed, flange resonance was not noticeable in the distilled water calibration, nor was it visible during the salt, sugar, milk, and cream experiments. Also, milk and cream are quite lossy, especially at low frequencies due to conductive properties of milk. Therefore, the small oscillations seen are most likely due to volumetric resonances from the reflections of the electromagnetic wave at the boundaries of the sample rather than from flange resonances. Understanding exactly how flange resonance arises and how to minimize it is outside the scope of this project and may require a deeper understanding of sensor designs.



(a) Dielectric constant with FR filter.



(b) Dielectric loss with FR filter.

Figure 3.12: Comparison of dielectric constant/loss of cream with FR filter.

3.4.5 Linear Regression Analysis

The most accurate way to effectively model the dielectric constant or loss vs butterfat content would be to use regression analysis. Similarly, to the work presented by Zhu et al. [57], there is a linear relationship between the dielectric properties and the fat content. Thus, by fitting a linear equation with the form

$$\epsilon' = a_1 F + b_1$$
$$\epsilon'' = a_2 F + b_2$$

and inverting them for F, where a_1, b_1, a_2, and b_2 are the coefficients of best fit and F is the fat content, it is possible to obtain a prediction for the fat content of a sample [57]. Using the data collected, the two equations would each give a value for the fat content and can be compared to the milk fat percentage printed on the carton.

Using the scikit-learn package, also commonly referred to as sklearn, in Python, the collected dielectric measurement data of the 5%, 10%, 18%, and 35% creams were split into a 50:50 ratio and an ordinary least squares (OLS) linear regression was performed for the relative permittivities and losses. The OLS technique minimizes the sum of squared error/residual cost function written mathematically as [75]

$$SSE = \sum_{i=1}^{N} (y_i - (\beta_0 + \beta_1 x_i))^2 \tag{3.19}$$

where β_0 and β_1 are the best fit regression coefficients for the intercept and slope respectively. These are calculated using

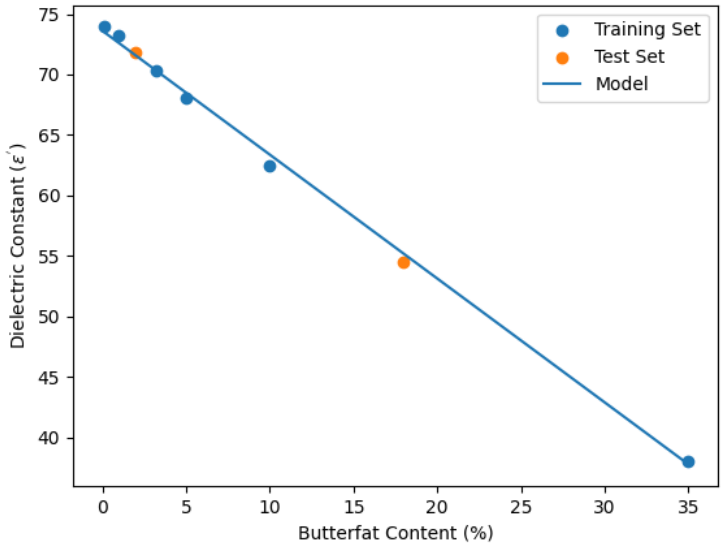
$$\beta_0 = \overline{y} - \beta_1 \overline{x},$$
$$\beta_1 = \frac{\sum_{i=1}^{N} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{N} (x_i - \overline{x})^2}.$$

with the bar representing an average value. A common metric to evaluate the regression model is the root-mean-square error (RMSE) given by [76]

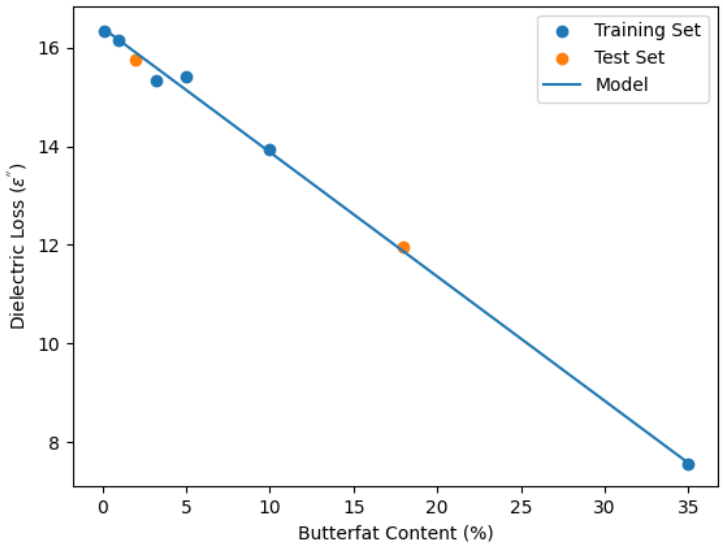
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y})^2} \tag{3.20}$$

where a hat represents a predicted value. The RMSE describes the average absolute error between the actual values and the predicted values given by the regression model.

The results of the linear regressions are plotted at 2450 MHz shown in Figure 3.13. The scikit-learn package is a machine learning library for Python built on scipy and comprises of different classification and regression algorithms. The least squares method is used to find a line of best fit based on the training dataset by minimizing the error or residual between the predicted value and the measured value. After inverting the equation for the fat content of a sample, the random_state of the train_test_split() function was adjusted several times to determine the coefficients that would give the most accurate predictions for the test data and best RMSE value. The random_state value is an integer value that controls the shuffling of data by the random number generator such the output is reproducible. If a random_state value is not set, this would mean a different output would be expected each time.



(a) Dielectric constant regression model.



(b) Dielectric loss regression model.

Figure 3.13: Dielectric constant/loss vs butterfat content at 2450 MHz using a linear regression model.

For the dielectric constant vs butterfat content plot, the equation of best fit obtained using this method, to three decimal places, is given by

$$\epsilon' = -1.030F + 71.342.$$

The RMSE given by the training set and the test set are $5.024e^{-15}$ and 0.630 respectively. As for the dielectric loss vs butterfat content, the obtained equation is given by

$$\epsilon'' = -0.273F + 15.311$$

with RMSE $6.280e^{-16}$ for the training set and 0.389 for the test set. Using these two equations for the test data, which consisted of 5% and 10% for the dielectric constant model and 10% and 18% for the dielectric loss model, the calculated fat results are presented in Table 3.6 and Table 3.7. The results are within decent range but could be improved with a larger set of training data.

Table 3.6: Predicted fat content at 2450 MHz from dielectric constant using test data

BF %	**Predicted BF % from ϵ'**	**Error in Prediction from ϵ'**
5%	4.336%	13.28%
10%	9.447%	5.53%

Table 3.7: Predicted fat content at 2450 MHz from dielectric loss using test data

BF %	**Predicted BF % from ϵ''**	**Error in Prediction from ϵ''**
10%	9.629%	3.71%
18%	16.021%	10.99%

Applying the regression models to the mixed cream data (Table 3.8) showed promising results and had a 4.73% average error based on the dielectric constants and a 5.83% average error based on the dielectric losses. This is consistent with the work conducted by Zhu et al. [57], where the dielectric constant model provided a smaller error than the dielectric loss one, but at 41 MHz instead.

The linear regression model holds well, but some possible sources of error will be discussed. There were small noticeable oscillations in the collected DAK-TL data from Figure 3.11 that may have arisen due to volumetric resonance. Unfortunately, the volumetric resonance issue would be difficult to address because the petri-dish only holds about maximum

Table 3.8: Predicted fat content of mixed cream samples at 2450 MHz

BF %	**Predicted BF % from ϵ'**	**Error in Prediction from ϵ'**	**Predicted BF % from ϵ''**	**Error in Prediction from ϵ''**
6%	6.722%	12.03%	6.108%	1.80%
8%	8.079%	0.99%	8.118%	11.80%
12%	11.985%	0.13%	12.368%	3.07%
16%	17.019%	6.37%	15.237%	4.77%
20%	19.171%	4.15%	21.540%	7.70%

12 mL of liquid. For the most accurate measurements, the classic DAK system should be used, but is unfortunately not available in the microwave lab.

Another issue that may have come into play during the development of the regression model is that the concentration of butterfat in a carton may not be exactly what is printed on the carton. There may be a slight deviance from 5%. It is also possible that the percentage of total solids in the cartons also ranged. To obtain a more accurate regression model, the exact concentrations of butterfat and total solids in the cartons should be known, since the dielectric constant and loss increase with water content for milk. This issue would also be difficult to address because the microwave lab does not have a machine that can be used for the compositional analysis of milk.

Since the determination of the regression model coefficients used a small range of data, 2 points, it potentially introduces a large error. This can be improved by increasing the number of points in both the training and the test sets. Even with the small training set, the fat content predictions are within good range and shows great promise towards determining the amount of butterfat a sample has given its dielectric properties.

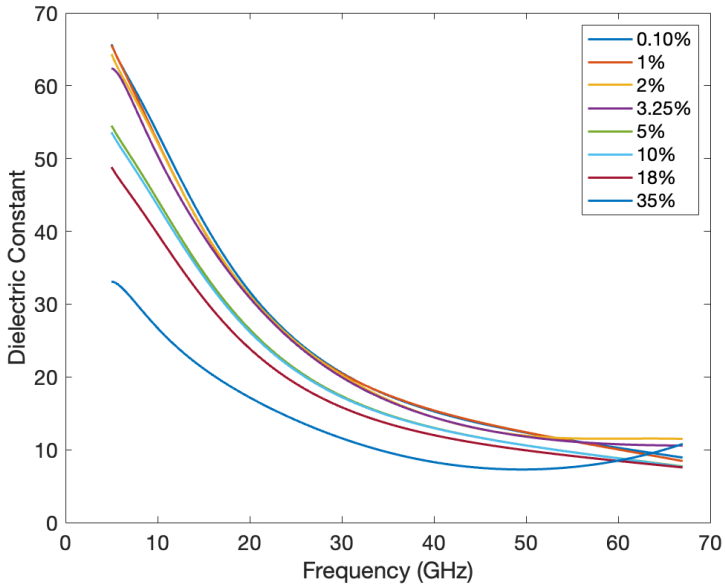
The dielectric properties of 5-35% cream products were measured using the DAK-TL over the frequency range 200 MHz to 20 GHz. Comparing the obtained results with the results found in the literature survey conducted, the trends align well and are expected. The dielectric constant and loss both decreased linearly with an increase in fat content, and the other constituents play a role in lowering the measured properties. These were compared to predictive models of a two-component solution consisting of milk fat and water, and the differences were as expected. Two regression models were made, one based on the dielectric constants and the other based on the dielectric losses of cream, to determine the butterfat content of a given sample. Considering the small training set and test set that was used in the development of the models, they both show predictions within a good range. Based on

the temperature study conducted on the DAK-TL prior to the measurements, it indicated that the instrument is not great for studying the effect of temperature on the dielectric properties of a liquid sample. There are also volumetric resonances that arise with this instrument when testing cream samples at lower frequencies. For future studies, it would be interesting to study the effect of homogenization, seasonality on the dielectric properties and how it could potentially affect the scattering responses received by the radar.

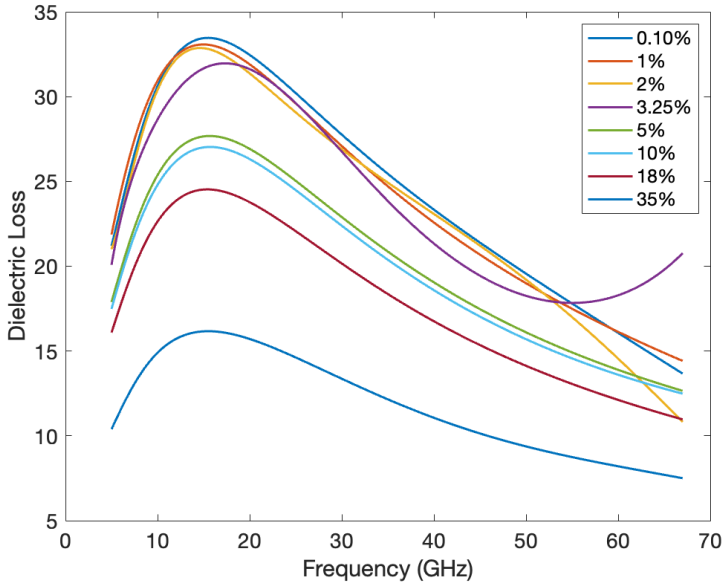
3.4.6 Characterization from 5 GHz - 67 GHz

A handful of work has been done and a thorough study conducted for the milk and cream products on the DAK-TL from 200 MHz to 20 GHz in the previous section. The end goal of this project is to implement a 60 GHz radar for a conveyor belt, yet no liquid characterization work at 60 GHz has been presented so far in this thesis. The same characterization tests conducted from 200 MHz to 20 GHz will be repeated for the higher frequency band of 5 GHz to 67 GHz. The mixture modelling equations will not be repeated as there is no data from current known literature of milk and cream at 60 GHz.

Following the same procedure as previously, the milk and cream tests were performed on the high frequency probe, DAK 1.2E, and characterized from 5 GHz to 67 GHz. The results are shown in Figure 3.14. It is noticeable in the data that the permittivities of 2%, 3.25%, and 35% look off compared to the rest of the samples while the losses of 1%, 2%, and 3.25% look off. These results could be due to not using a fine enough resolution, but 250 MHz steps were used similar to the high end of the sweep for the data up to 20 GHz. Since the products are mostly comprised of water, there is not much happening because the dipolar polarization effect is not as observable compared to around 200 MHz. This means that 250 MHz would be a decent resolution. On the other hand, 5 GHz to 67 GHz is a much wider bandwidth that the DAK-TL was sweeping across compared to the smaller bandwidth of 200 MHz to 20 GHz. The wider the bandwidth, the more susceptible the data is to noise. At the time of testing, there was also an unlevelled local oscillator drive error on the VNA which may have been the root cause of the issues. This issue would prevent frequencies to be mixed properly, but majority of the data appears to be fine.



(a) Dielectric constant of milk and cream.



(b) Dielectric loss of milk and cream.

Figure 3.14: Dielectric characterization of milk and cream from 5 GHz to 67 GHz.

Radar Operating Frequency Selection

After taking a broadband set of dielectric measurements, the first set from the DAK-TL 3.5 probe running from 200 MHz to 20 GHz and the second set from the DAK-TL 1.2 E probe running from 5 GHz to 67 GHz, the operating frequency of the radar needs to be decided on. The advantages of picking a frequency from the first set of measurements, or the lower frequency side, is that the radar is less sensitive to minute displacements, there are larger differences in the dielectric properties, and it is better for long range detection. Some main disadvantages from this are that the radar would require a larger footprint, it'll have lower resolution due to the longer wavelength, and more electromagnetic interferences. For example, if a radar operating at 2.4 GHz was selected, there would be a lot of other devices running at 2.4 GHz such as WiFi and bluetooth that could potentially interfere with the radar measurements. Additionally, the radar would have a much larger footprint and the installation process will not be as simple.

On the other hand, selecting a higher frequency operating radar for this project is much more attractive because of its higher resolution and fewer interferences in the mm-wave band. Higher frequencies are better for short range detection and offer a small and compact form factor. This makes it much easier to install the radar and it is very portable. Additionally, the small size means that there will be optimal coverage from the carton as all of the beams from the antenna array will be covered by the carton. Unfortunately, it does come with the disadvantage that it is very sensitive to minute displacements and has short penetration depth. This mm-wave signal will be attenuated very quickly and does not have long range. To improve the propagation of the signal, an external power amplifier could be introduced in the transmission stage and is mentioned in more detail in the next chapter.

Because of its many appealing advantages, the higher frequency set of measurements were selected. These mm-wave frequencies are not foreign to the food and beverage industry. More details will be discussed in the next chapter about the radar technology and selecting the radar based on performance. The radar selected operates at 60 GHz and was chosen because it was readily available and widely used within the lab group.

Approximation of Outliers using Linear Regression Model

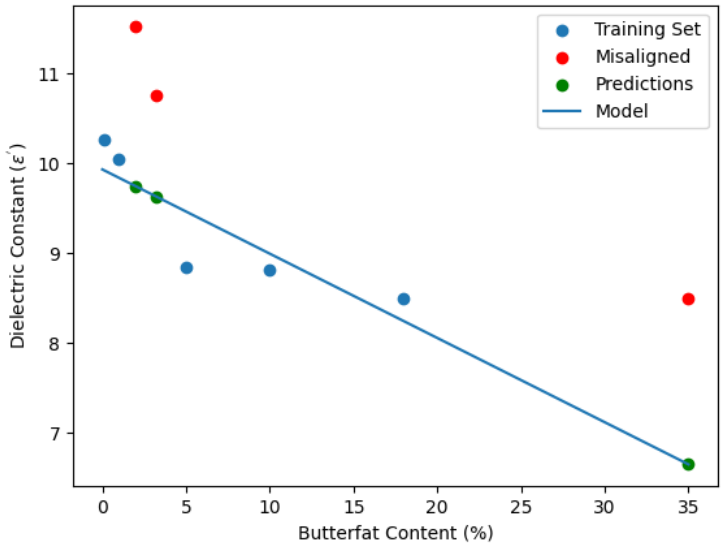
Since the expected behaviour of the permittivities and losses are known, by comparing among each other and through extensive work previously presented, these errors can be addressed through some data processing. There should be a linear relationship between the fat content of a sample and its dielectric constant and loss. As the fat content in a sample

increases, the dielectric constant or loss should decrease accordingly. The data processing done is a simple linear regression, reusing the same strategy as the one created for 2450 MHz but implemented for 60 GHz, the operating frequency selected for the radar. To generate the equations of best fit, the outlier data points were excluded from the dataset. The equations of best fit for the linear models are:

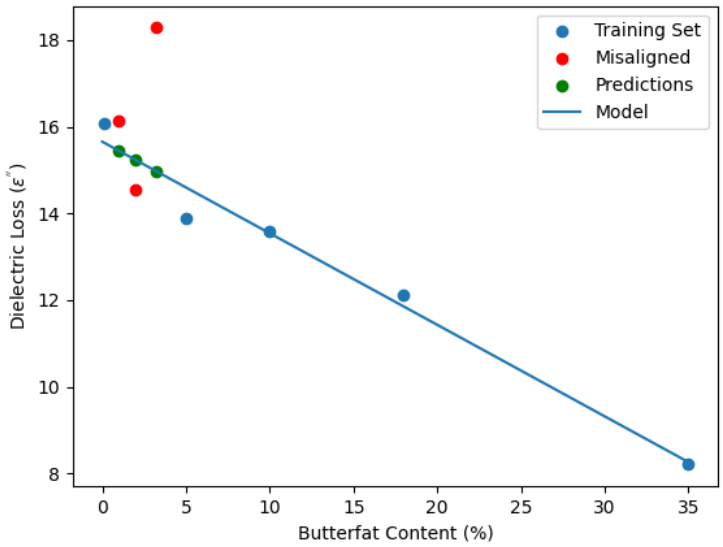
$$\epsilon' = -0.0938F + 9.930,$$
$$\epsilon'' = -0.211F + 15.654.$$

The RMSEs for the dielectric constant and loss models were 0.749 and 0.978 respectively.

A plot of the model and related points are displayed in Figure 3.15. The data points used for the training set are in blue and the misaligned points, or outliers, are in red. It is quite evident that the red points lie quite a distance away from the line of best fit, especially in the permittivity case. In the case of the losses, the differences are not as drastic. To obtain the values of the permittivity or loss at the problematic frequencies, the value was simply read off of the model and the corresponding points highlighted in green. These values are tabulated in Table 3.9 and will be used for the HFSS simulations in the next section.



(a) Dielectric constant regression model.



(b) Dielectric loss regression model.

Figure 3.15: Dielectric constant/loss vs butterfat content at 60 GHz using a linear regression model.

Table 3.9: Predicted dielectric properties at 67 GHz

BF %	**Predicted ϵ'**	**Predicted ϵ''**
1%	–	15.443
2%	9.742	15.232
3.25%	9.625	14.968
35%	6.646	–

This experiment could be redone to find more accurate values of the dielectric properties at each of the problematic frequencies. However, majority of the points were in line with what would be expected at lower frequencies and the general physics background covered in the theoretical section. Due to time constraints, the experiment was not reconducted and the data processing method was used to obtain a rough estimate of the dielectric constant and loss values in order to build the HFSS simulation.

Chapter 4

Radar Work

This chapter will discuss some fundamental radar system theory starting from a system-level approach to radio systems to some of the key parameters of a FMCW radar. The parameters for the Infineon and SOLI radars used are provided. FEM/FEBI and SBR+ simulations in HFSS are presented to showcase the usage of a simple 1 transmitter and 1 receiver radar system for liquid characterization. The setup and simulated results are used to develop further evidence that differentiation is possible based on the obtained dielectric measurements from the previous section. After, experimental results of stationary and moving carton tests will be analyzed and compared to the simulated results. A classification algorithm will be used to sort the data based on their percentages of milk fat concentration.

4.1 Radar System Theory

4.1.1 Radar Fundamentals and Radio System Building Blocks

In general, RADAR stands for Radio Detection and Ranging. Contrary to its primary use of object detection and measuring their distances, the use of electromagnetic radars in this project was for liquid characterization. Based on the built background from the previous section and chapter, the scattering parameters off an object will vary based on their dielectric properties along with many other factors. The Infineon and Google radars selected for this project provide a small low-cost form factor with a potential for simple installation and implementation. More details regarding the radars are given in Section 4.1.4.

A simple continuous wave (CW) radar would radiate a pure sinusoidal wave with zero modulation continuously. Compared to a pulsed radar which is larger and more expensive, there is no minimal detectable range but it would struggle to measure the distances of stationary objects. Another notable difference between a pulsed radar and a CW radar is the fact that the pulsed radar has a high peak power, while the CW radar operates on low-power and is very safe [77]. When measuring a moving object, there is the added Doppler shift effect that would increase or decrease the frequency as it moves towards or away from the radar. Using this change in frequency, the distance of the moving object and its velocity may be calculated. In the case of a stationary object, there will be no change in frequency and calculating range and velocity from the measured parameters of this object would be a challenge.

For the application of differentiating liquids, a CW radar may suffice. That said, modulating the frequency would allow for customizable resolution of the distance and velocity of a measured object [78] especially when implemented for a conveyor belt purpose. This is done by ramping up the frequency of the transmitted sinusoidal wave across the operating bandwidth (BW) known as a chirp through different modulation patterns. There are several different modulation patterns such as sawtooth, triangular, rectangular, and stepped modulations [77]. The radars used in this project use the sawtooth modulation pattern. With the modulation, the received signal has a time delay such that there will be a difference in frequency between the transmitted and reflected signals. These two signals are mixed together and form the resultant beat signal.

There are several advantages to using frequency modulated (FM) signals compared to amplitude modulated signals. One reason is that FM signals are less susceptible to noise and thermals [79] since information is stored in the frequency variation rather than the amplitude. FM signals tend to also occupy larger bandwidth than AM ones, which would improve the sensitivity of the measurements. The exact implementation of the radar on a conveyor belt was not determined, but whether the radar is turned on in bursts or left continuously running will eventually suffer from signal degradation from thermal effects. Some work was done on thermal characterization of the Infineon radar and will be presented in a later section. Optimizing the radar setup for long term use could be a potential project. Since the dielectric properties of the liquid varies with frequency, the chirping could help to detect any differences in liquids over the radar's operating frequency.

Both the Infineon and Google radars operate around 60 GHz are compact and low-cost millimetre-wave (mm-wave) radars. The nature of these mm-wave radars present its own advantages and disadvantages. For example, mm-wave frequencies experience a high attenuation that can lead to high isolation when placed in an array of other mm-wave radars [79]. High attenuation could also mean there will be a lower probability of echo

signals reflected back from other objects. Using an operating frequency of 60 GHz also means there is less electromagnetic noise from its surrounding environment. If a 2.4 GHz radar was selected instead, it could have electromagnetic interference from WiFi and other IoT devices. This could be problematic in a manufacturing facility heavily reliant on IoT technology. The selection of the 60 GHz operating frequency and its benefits were explained in the previous chapter. Both the Infineon and Google radars are readily accessible within the lab group.

The ability of mm-wave radars to detect movements on the millimetre order could be useful for detecting slight variations in the milk carton itself as well as any possible foreign liquids or objects. However, this itself presents a multitude of challenges that cannot be easily accounted for. It will be difficult to ensure repeatability among tests taken on different dates because of the variability in a setup's surroundings. The orientation and geometry of the object will also affect the measurements. Calibration will be difficult and a large number of data samples would need to be carefully collected in order to develop an adequate ML algorithm. Implementing a final product will take time and careful optimization. Because the radar signals are highly nonlinear and unpredictable, the training set for the algorithm requires knowledge of all the possible radar signals for a carton to develop the golden sample.

It is important to note that electromagnetic waves do not propagate well underwater due to their latitudinal behaviour and the dispersive properties of water. It is a known fact that as the frequency of the wave increase, the penetration depth into any material decreases. Thus, when characterizing water or milk with electromagnetic radars, a lack of sensitivity in the radar may arise. This would be due to milk and water being a lossy dielectric medium that would cause the transmitted signal to be attenuated. To address this issue, a possible solution could be to introduce a dedicated amplification stage in the signal transmission using a PA. This may be done by connecting a PA to the transmitter of one of the FMCW radars used in this project.

One example of a PA that could be implemented with this system is a gallium-nitride (GaN) high electron mobility transistor (HEMT) PA. Many power amplifiers designed at higher frequencies use GaN HEMT technology because it provides higher efficiency, better output power, and improved power density compared to conventional silicon technologies [80]. Although this type of technology has become increasingly popular for 5G base stations [80], [81] and applications requiring high voltage and high-power electronics [82], it may be beneficial to use them with the radar systems found in this project. A downside to this is that the GaN PAs are prone to self-heating because of their high-power density and high electron mobility which would cause the overall reliability and lifetime of the system would degrade if temperatures surpassed 200°C [81], [83]. With this approach, the penetration

depth of the transmission signal would increase and better sensitivity of the radar should be attained so that slight changes in liquid measurements are detectable. However, this is merely an example and is much outside the scope of the project. The design of the interface between the radar and PA may take a considerable amount of time as the connection from each pin to the radar would need to be impedance matched. Aside from the self-heating in the GaN HEMT PA, excessive overheating in the FMCW radar poses a challenge otherwise signal degradation would arise. Some work done with examining the self-heating effect in these PAs and thermal results of the Infineon radar can be found in Appendix G. It could be considered as an additional improvement for the radars in future work.

Consider a simple radio system comprising of a transmitter and a receiver (Figure 4.1). Both include several building blocks such as a bandpass filter (BPF), an amplifier, mixer, local oscillator (LO), lowpass filter (LPF), and a digital signal processor. In the transmitter, there is a digital to analog converter (DAC) that converts the digital signal into an analog one with a power amplifier (PA) to boost the transmitted signal. On the other hand, the receiver consists of a low noise amplifier (LNA) that amplifies the received signal before passing into the analog to digital converter (ADC). The mixer is used in conjunction with the LO to down-convert a RF signal to a baseband signal in a receiver or up-convert a baseband signal to a RF signal in a transmitter. This project does not explore the design aspects of these RF components, but designing and optimizing a system for this project's end goal of non-invasive characterization could be a good future work.

To elaborate on the process and the mixer function, it is a 3-port device that takes in two input signals, either or the RF signal or intermediate frequency (IF) signal and the LO signal, and multiplies them together to create an output signal. Mixers are either non-linear or time-varying devices. That is, mathematically, if the general form of the input signals are

$$\begin{aligned} x_1 &= a_1(t) \cdot \cos(\omega_1 t + \phi_1) \\ x_2 &= a_2(t) \cdot \cos(\omega_2 t + \phi_2) \end{aligned}$$

with $a(t)$ being the amplitude, $\omega t + \phi$ is the total phase, and ϕ is the excess phase, then the resulting output signal will be

$$\begin{aligned} x_{out} &= x_1 \times x_2 \\ &= a_1(t)\cos(\omega_1 t + \phi_1) \cdot a_2(t)\cos(\omega_2 t + \phi_2) \\ &= \frac{a_1(t)a_2(t)}{2}\Big[\cos\big((\omega_1 t + \phi_1) - (\omega_2 t + \phi_2)\big) + \cos\big((\omega_1 t + \phi_1) + (\omega_2 t + \phi_2)\big)\Big] \\ &= \frac{a_1(t)a_2(t)}{2}\Big[\cos\big((\omega_1 - \omega_2)t + (\phi_1 - \phi_2)\big) + \cos\big((\omega_1 + \omega_2)t + (\phi_1 + \phi_2)\big)\Big]. \end{aligned}$$

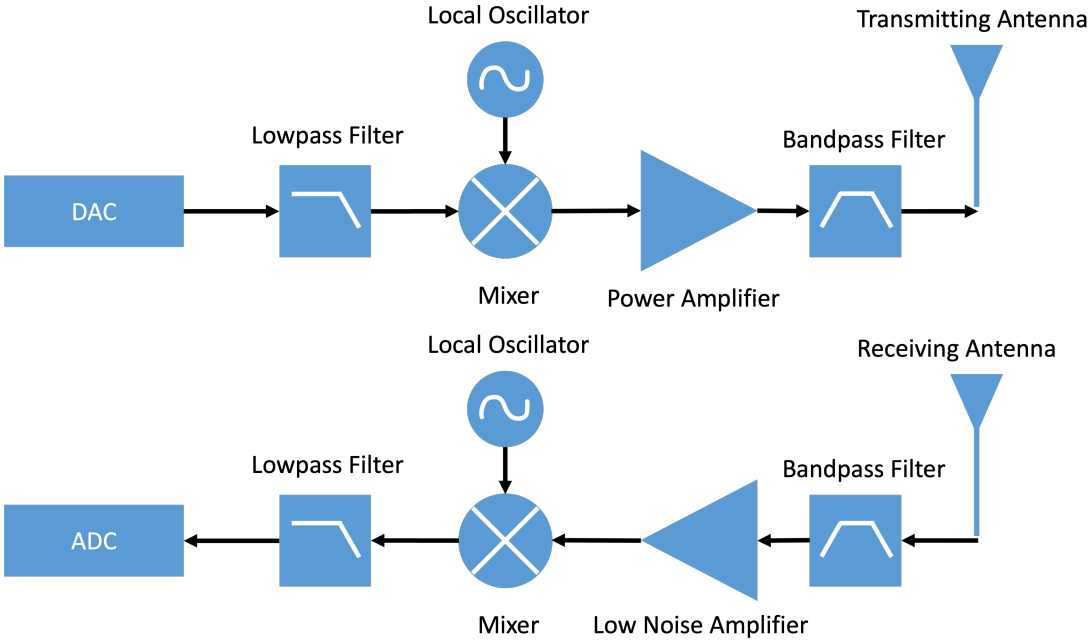


Figure 4.1: Transmitter and receiver radio system. Adapted from [84].

From the product to sum trigonometric identity used above, there are additional frequency components, an addition and a subtraction term, $\omega_1 - \omega_2$ and $\omega_1 + \omega_2$. Alternatively, this could be shown using the complex exponential form of cosine. Usually there is a filter that is placed after the mixer to filter out the unwanted frequency term term which depends on whether the mixer is being used to downconvert or upconvert a signal. For example, if the mixer was converting a RF signal into a DC one, $\omega_{IF} = \omega_{RF} - \omega_{LO}$ would be kept while the $\omega_{RF} + \omega_{LO}$ term is filtered out. This calculation was notable because it serves as a foundation for determining a IF frequency based on the transmitted signal and received signal in a radar system. This IF frequency is then used to determine the range resolution of the radar and is limited by the ADC sampling rate [85]. Typically,these make up the basic front-end of any radio system.

Figure 4.2 illustrates a simple radar architecture built off of the simple transmitter and receiver radio system displayed in Figure 4.1. Instead of utilizing local oscillators, some of the generated signal gets coupled into the mixer on the receiver path and acts as the oscillator. This generates the IF signal which then gets filtered using a lowpass filter and into the ADC. The voltage controlled oscillator (VCO) is used to control the high-frequency oscillations and can sometimes can integrated into a phase-locked loop (PLL). The mixing

of the coupled transmitted signal and the received signal generates the beat signal briefly mentioned earlier.

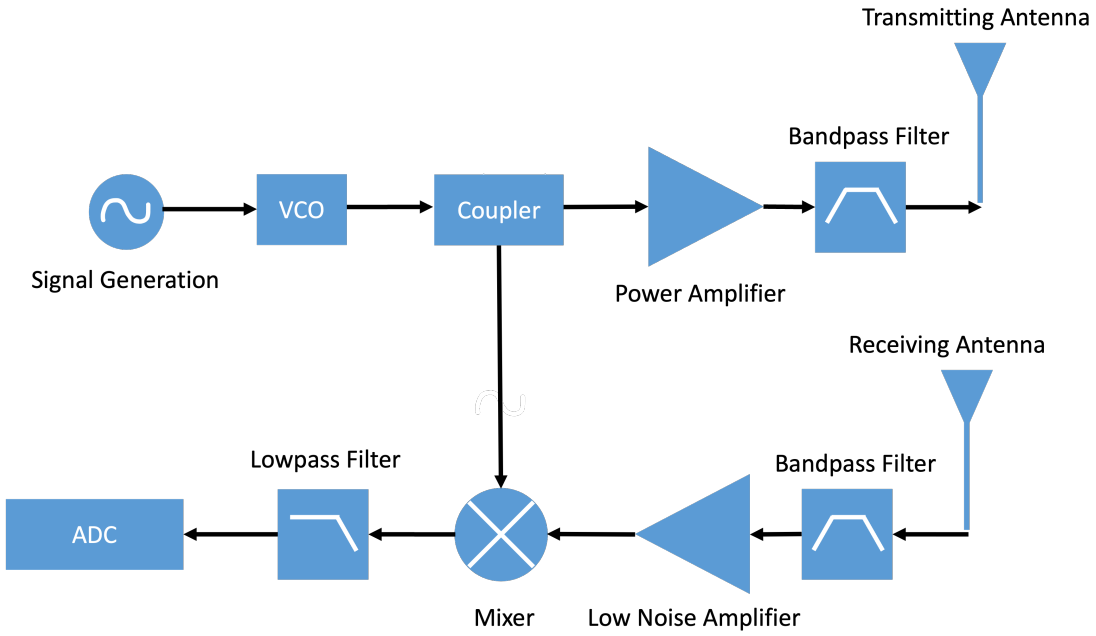


Figure 4.2: Simple radar architecture. Adapted from [78].

4.1.2 Radar Processing

There are some key settings to configure when running the radar. While this project has a larger experimental emphasis when it comes to radars, it could be beneficial to understand the math behind the signal processing. Derivations of the radar equation and beat signal can be found in [78]. The following will be discussed in some more detail:

1. Intermediate frequency (IF) signal
2. Range resolution
3. Number of chirps per frame
4. Number of samples per chirp

5. Sample bin number

The intermediate frequency or IF signal has already been briefly discussed in the prior subsection. Instead of using a mixer to down-convert or up-convert a signal, radars use a mixer to mix the transmitting and receiving signals together to create an IF signal. If the two transmitting and receiving signals are examined with frequency vs time on a single stationary object, it becomes more noticeable that the received signal is the same as the transmitted signal but with a delay [85]. This frequency tone can be described as

$$f_{IF} = \frac{S2d}{c} \tag{4.1}$$

where S is the slope of the chirp, d is the distance of the object, and c is the speed of light. Notice the 2 arises because the transmitted signal needs to travel twice the distance before being received. If there are multiple objects in front of the radar, there will be multiple reflected chirps detected by the receiver and there will be multiple tones in the IF signal.

It could be challenging if two or more objects are situated very close together such that the tones in the IF frequency spectrum appear as a single peak. To resolve this and differentiate the tones, the length of IF signal can be increased. This increases the bandwidth which then improves the resolution of the radar. This is known as the range resolution. As with wideband RF devices, a larger bandwidth would mean that the device is more susceptible to noise and more sensitive to changes in the environment. A short derivation of equation for range resolution can be found in [85], but the final equation only depends on the bandwidth used for the chirp,

$$d_{res} = \frac{c}{2B}. \tag{4.2}$$

After capturing the data and digitizing it, it gets post processed by the radar's micro-controller unit (MCU). The ADC sampling rate, i.e. the number of samples per second, is given by

$$F_s \geq \frac{S2d_{max}}{c}. \tag{4.3}$$

The sampling rate limits the IF bandwidth and desired maximum difference [85]. To pick a good sampling rate, there needs to be a tradeoff between resolution and achieving enough bandwidth based on Nyquist's theorem.

A radar matrix is built with sample bins corresponding to the processed chirp data over time. Taking a Fast Fourier Transform (FFT) on each row gives data about the

range of the object while taking a FFT of each column gives information about its velocity [86]. These are known as the range-FFT and doppler-FFT respectively. When taking measurements with the radar, another important parameter that needs to be specified is the frame rate which determines how long a single frame is. Next, the number of chirps per frame needs to be selected. In this project, 32 chirps per frame was used with each chirp taking 64 samples. When conducting the data processing and plotting the raw data captured in terms of its amplitude, the range bin is held constant so the entire row can be extracted from the radar matrix. By keeping the range bin constant, the radar estimates the distance an object is away from the radar. The range bin changes the plot depending on the range resolution specified in the parameters. Fast time refers to the high frequency sampling of chirped data points within a pulse, while slow time refers to the larger time scale sampling between pulses. This is illustrated in Figure 4.3.

The radar equation or radar range equation describes how much power is received at the receiver from an object's scattering a known distance away. Equation 4.4 [87], [88] relates the received power at a distance R to the transmitted power (P_t), radar cross section, and the frequency the radar is operating at. The first factor represents the power density of the incident or transmitted power. The second factor represents the scattered power density that returns to the radar where σ is the two-dimensional radar cross section of the target or echo area. The final factor represents the effective aperture of the receiving antenna. G_t and G_r are the gain of the transmitting antenna and the receiving antenna respectively. To account for any losses in the system, considering the radar equation in the decibel scale may be simpler. Consider $P_l = L$ as the lost power to the system, then this gets subtracted from the total received power. Thus, in Watts it would appear in the denominator.

$$P_r = \frac{P_t G_t G_r}{4\pi R^2} \frac{\sigma}{4\pi R^2} \left[\frac{\lambda^2}{4\pi} G_r\right] \frac{1}{L}. \tag{4.4}$$

From Equation 4.4, it can be concluded that as the object is moved further away, the received power decreases by a factor of R^4.

The radar transmits chirps periodically between 1-10 kHz, referred to as the radar repetition frequency (RRF). The device then receives the reflected signals at a radar repetition interval (RRI) defined as the inverse of the RRF [89]:

$$RRI = \frac{1}{RRF}. \tag{4.5}$$

The transmitted and received signals can be modelled as a function of two different time scales, t for fast time, and T for slow time [89]:

$$S_{TX}(t, T) = u(t - T)e^{-j2\pi f_c t}, \tag{4.6}$$

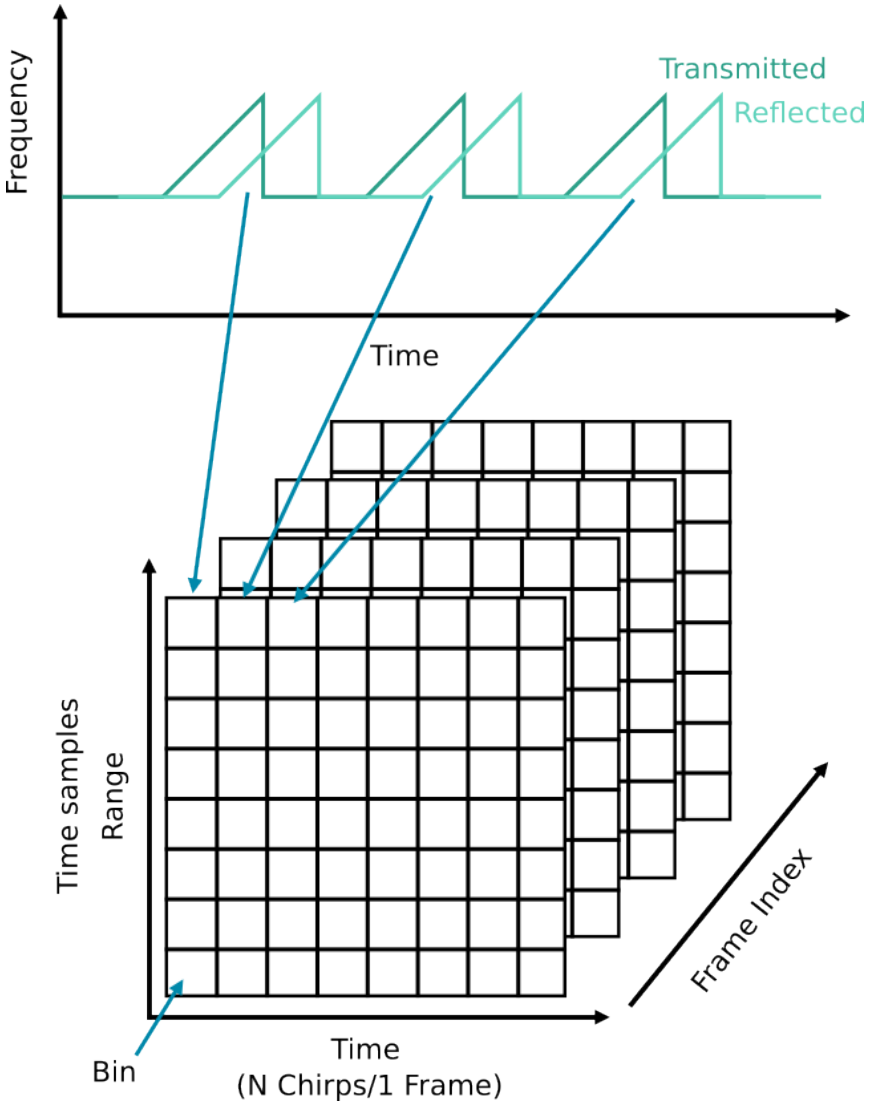


Figure 4.3: Radar chirp and data cube. Adapted from [86].

$$S_{RX}(t,T) = u(t-\tau-T)\sigma e^{-j2\pi f_c(t-\tau)} \tag{4.7}$$

for $T = 0, RRI, 2RRI, ...$ In the above equations, $u(t)$ represents the complex envelope over one period of the modulation scheme used and f_c is the centre operating frequency. In Equation 4.7, σ represents the two-dimensional cross section of the target or echo area, seen previously, and τ is the total time it takes for the transmitted signal to reach the target and its scattered signal to reach the receiver. Noticeably, received signal is just

delayed in fast time by the total travel time defined as

$$\tau = \frac{2d}{c} \tag{4.8}$$

where d is the distance between the radar and the target and c is the speed of light in vacuum.

The radar data collection process is outlined as a block diagram in Figure 4.4 and was adapted from [89]. The analog signal pre-processing happens using the radar front-end module. After the analog data collection, the on-board MCU handles all the basic digital data pre-processing. More complex digital signal processing is conducted through the development kit supplied with the radar or through a custom script. More details of the signal processing can be found in [89]. The radar will spit out processed data which will then be used to create the ML model tailored to a specific application.

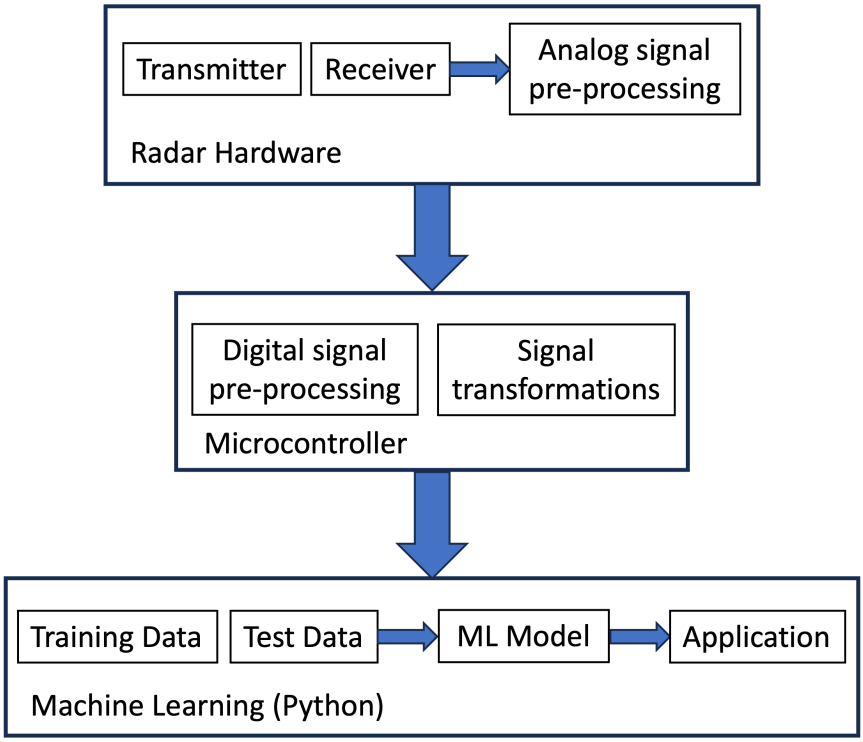


Figure 4.4: Concept workflow describing data collection and processing. Adapted from [89].

4.1.3 Dielectric Characterization using FMCW Radars

To understand the trends gathered by the FMCW radars during the data collection, an intuitive and mathematical understanding of the physical processes occurring should be developed. The simulations presented will feature a single dielectric slab with the properties gathered from the liquid characterization section. When testing with the radars, these simulations are less realistic since the milk needs to be contained in a package. The transmitted signal travels through the air to the cardboard of the carton and then through to the milk. Although the FMCW signal is frequency modulated by definition, the same physics continues to apply as if it were an unmodulated continuous wave. This is a multiple dielectric slab problem.

Approaching this from a single slab first, the reflection coefficient at a single boundary from normal incidence is given by [90]

$$\Gamma = \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1}. \tag{4.9}$$

Here, η is defined as the intrinsic impedance and is defined as $\eta = \sqrt{\frac{\mu}{\epsilon}}$. Subscript 2 denotes the second medium and subscript 1 denotes the first medium. By assuming a constant magnetic permeability for both mediums such that $\mu_1 = \mu_2 = \mu_0$ and the first medium is air/vacuum, the above equation can be expanded and simplified into

$$\Gamma = \frac{1 - \sqrt{\epsilon_r}}{1 + \sqrt{\epsilon_r}} \tag{4.10}$$

with ϵ_r being the relative permittivity of the second medium. By substituting in dummy values for the permittivity, it is clear that the reflection is greater with a larger permittivity. Γ is -1 with a short circuit, meaning the reflected wave is 180 degrees out of phase with the incident wave. Likewise, Γ is 1 with an open circuit and the reflected wave is in phase with the incident wave. Thus, if Γ was a negative value, this would represent a 180 degree phase shift from the incident wave. The smaller the magnitude of the reflection coefficient, more of the incident wave is transmitted through to the second medium.

Now, consider the milk in its packaging. The incident signal now needs to transmit through the packaging before reaching the product. There must be a decent amount of transmission through the packaging in order to have an adequate reading. Simula et al. [91] reported that the dielectric properties of printing paper are between 2-4 at 20 Hz to 1 MHz for both the permittivity and loss. Assuming that the dielectric constant and loss of the cardboard used is similar to the values for paper, when the values are substituted

into Equation 4.10, then the reflection seen at the radar is rather small. Most of the signal passes through to the milk content inside the carton. The second boundary between the milk to cardboard can be negligible because majority if not all of the transmitted signal will be attenuated before then.

In the HFSS simulations where a dielectric slab with a thickness of a guided wavelength is used, the second boundary from milk to air or free space becomes more important. Similar to the experimental case with a carton, there are now two boundaries. The reflection seen at the radar can be approximated as a recursion or geometric series [90] as follows:

$$\Gamma_{radar} = \Gamma_0 + \Gamma_1 e^{-2\gamma_1 d_1} + \Gamma_2 e^{-2(\gamma_1 d_1 + \gamma_2 d_2)} + ... + \Gamma_N e^{-j2(\gamma_1 d_1 + ... + \gamma_N d_N)} \tag{4.11}$$

where Γ represents a reflection coefficient, γ is the propagation factor of the wave propagation through a lossy medium, and d is the distance of the medium from one boundary to another. $\gamma = \alpha + j\beta$ where α is the attenuation constant and β is the phase constant. Figure 4.5 illustrates the general multiple dielectric slab problem that Equation tries to model. Γ_0 is the reflection coefficient at the first boundary or zeroth boundary, Γ_1 is the reflection coefficient at the second boundary and so forth. d describes the distance of the medium because it is important for the exponential terms in the equation, or the propagation factors. The phase constant and attenuation constants are given by [34]

$$\beta = \frac{2\pi\sqrt{\epsilon_r}}{\lambda} \tag{4.12}$$

and

$$\alpha = \frac{\pi\sqrt{\epsilon_r}}{\lambda} \tan\delta \tag{4.13}$$

The physical intuition of the reflection coefficient increasing with an increase of relative permittivity can be attributed to a sample's absorptive properties. In the study conducted by Yavari et al., low-fat milk scattered more light because it had lower solids and less absorption of the visible light into the liquid. Whole milk had a smaller light scattering response because of the increase of total solids in the liquid.

4.1.4 Radar Hardware

In particular, the two FMCW radars that have been chosen for this project: the Infineon XENSIV™ 60 GHz DEMO BGT60TR13C (58-63 GHz bandwidth; 1 transmitter, 3

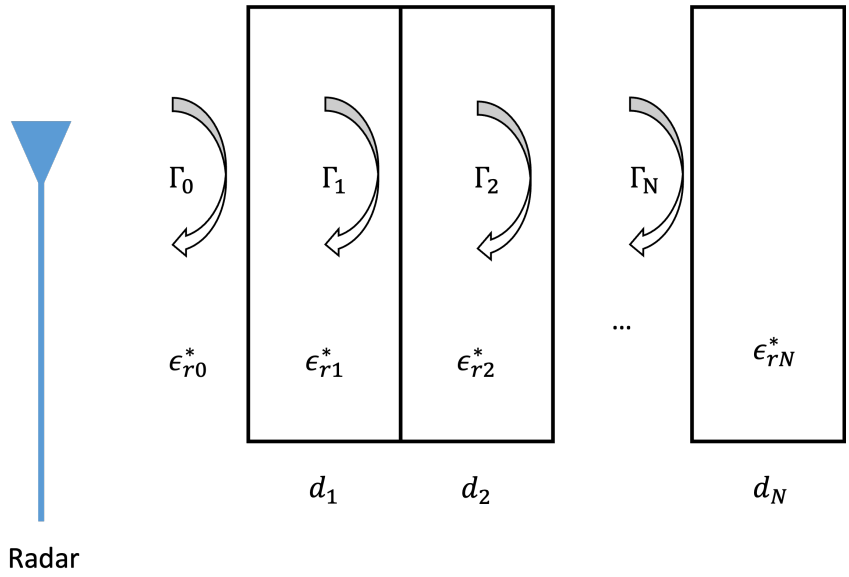


Figure 4.5: Illustration of multiple dielectric slab problem.

receivers) and the Google ATAP Project SOLI radar sensor (57-64 GHz bandwidth; 2 transmitters, 4 receivers). Both of the radars have the same Silicon-Germanium (SiGe) transceiver chip developed by Infineon [92], but feature different bandwidths and a different number of transmitters and receivers. The devices offer low noise, low output power, and low power consumption. A figure with the SOLI radar and its antennas labelled is shown in Figure 4.6. More details about the Infineon hardware are provided in Appendix F.

The radar parameters used for this project are tabulated in Tables 4.1 and 4.2. These values were selected based on the aforementioned equations for the range resolution and derived maximum range given a sampling rate. The number of chirps per frame and samples per chirp values were chosen based on previous work within the lab group. More information regarding processing and chirp parameters can be found in Texas Instrument's introduction to FMCW radars in [85].

4.2 Full-Wave Electromagnetic Simulations

HFSS (High Frequency Simulation Solver) is a full-wave 3D electromagnetic simulator by ANSYS. The software includes several solvers, but is mainly a finite element method (FEM) simulator capable of far-field and near-field analysis as well as being able to simulate

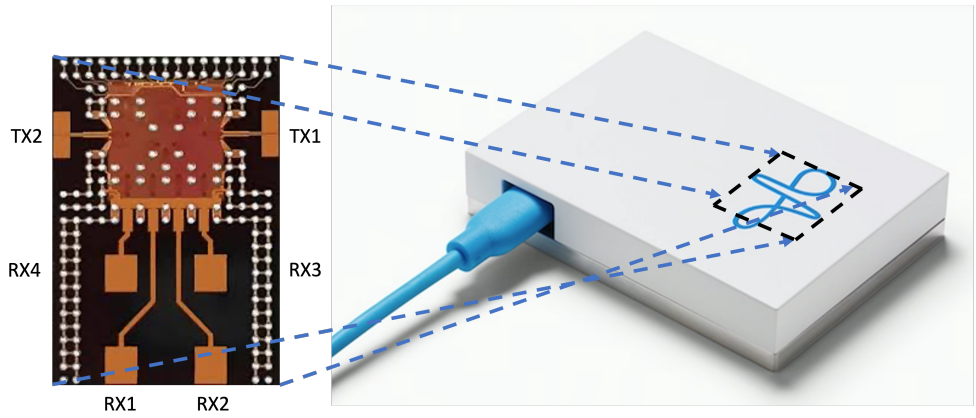


Figure 4.6: Google ATAP SOLI radar chip with labelled antennas (left) and packaging (right). Adapted from [93] and [94].

Table 4.1: Infineon radar parameters

Parameter	**Value**
Range Resolution	0.15625 m
Number of Chirps per Frame	32
Number of Samples per Chirp	64
Sampling Rate	1 MHz
Output Power	31 dBm
Frequency Range	58 - 63 GHz
Centre Frequency	60.5 GHz

various materials and geometries. Depending on the setup of the problem, it may be more efficient to run the FEM analysis with an integral equation (IE) based method. Or rather, if the problem is large and requires extensive GPU resources or moving elements, then the shooting and bouncing rays (SBR) solver would be most efficient.

4.2.1 FEM/FEBI Simulation

The FEM based method in HFSS divides up a geometry or model up into tetrahedral shapes. The solver then finds a solution to Maxwell's equations over each of the tetrahedras with the surrounding boundary conditions. If the specified S-parameter error tolerances are not met, the mesh size, or the size of the tetrahedrals, are reduced until the error in

Table 4.2: SOLI radar parameters

Parameter	**Value**
Number of Transmitters	2
Number of Receivers	4
Range Resolution	0.02 m
Number of Chirps per Frame	32
Number of Samples per Chirp	64
Sampling Rate	1.79 MHz
Output Power	24 dBm
Frequency Range	57 - 64 GHz
Centre Frequency	60.5 GHz
Bandwidth	7 GHz

the tolerances are met. Once this condition is met, the solution is said to have converged. This divide and conquer strategy is useful in handling complex geometries, but because it is a volume based mesh solution, it could be inefficient in handling radiation and scattering problems [95]. In the case of an open radiation problem, if the material used is uniform, then it may be more efficient to use the integral equation based method as it uses a surface mesh instead.

For example, if the FEM method was used to compute the scattering parameters of a dielectric with two antennas, it would be much slower than if the IE method was used. This is because an airbox would have to be defined that surrounds all three objects, and since the FEM method is volume based, it would compute over the volume of empty space in the airbox. This problem would be handled much more efficiently with the IE solution method because it allows a discontinuous solution space [96], truncating the free space in the FEM airbox. Using a hybrid approach combining both the FEM and IE solution methods, three airboxes can be defined, enclosing each object tightly. There is no theoretical minimum distance that the airbox needs to be be from the object [95] and it could be an arbitrary shape conforming to the object. The FEM method would be used for the dielectric and the two antennas inside of the airbox, while the FEBI method would be used on the surface of the airbox. If the dielectric was uniform, it could also be modelled as a dielectric IE-region [96]. This significantly reduces the amount of computation time required. More details can be found in [95] and [96] about the computational differences between the FEM method and the hybrid FEBI/FEM-IE solution.

The approach just described will be used to simulate a simple radar system with 2 waveguide antennas and a thin dielectric whose properties are based off the 60 GHz mea-

surements from the previous section. Since milk is very lossy, only a small portion of the milk carton needs to be simulated in order to approximately see the differences. Otherwise, the solution will run into RAM and memory issues and will become highly inefficient. This is where the SBR+ method comes into play in the next subsection, as it can more efficiently solve for large geometries. Since waveguide antennas will be used, the driven modal solution type is selected. Here, the solution type supports modal propagation of energy where the modes could be transverse electric (TE), transverse magnetic (TM), or transverse electromagnetic (TEM) [97]. Note that waveguides do not support TEM modes.

The waveguide must be designed for a specific operating frequency, because the operating frequency depends on the cross-section of the waveguide. Frequencies below this cut-off are attenuated by the waveguide. Frequencies above this cut-off are minimally attenuated. There could be various waveguide modes, but the design of the waveguide should support only one active mode. The operating frequency is chosen between the cutoff frequency of the mode with lowest cutoff frequency and the cutoff frequency of the mode with the next lowest cutoff frequency. In rectangular waveguides, the mode with the lowest cutoff frequency is TE10.

To calculate the cutoff frequency of the TEmn or TMmn modes, the following expression can be used [90]:

$$f_c = \frac{c}{2}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}, \tag{4.14}$$

where m and n are selected based on the mode, b is the height of the waveguide, and a is the width of the waveguide. In this case, TE10 is the desired operating mode, also called the dominant mode, and the equation simplifies to

$$f_c = \frac{c}{2a}. \tag{4.15}$$

Then to calculate the lower and upper operating frequencies, a frequency can be picked between $f_1 = 1.25f_c$ and $f_2 = 1.89f_c$. Note that the length of the waveguide does not affect the cutoff frequency.

For a waveguide operating at 60 GHz, the cutoff frequency can be written as

$$f_o = 1.5f_c \Rightarrow f_c = \frac{f_o}{1.5}$$

where f_o is the operating frequency. Substituting this into the cutoff frequency for the TE10 mode gives

$$\frac{f_o}{1.5} = \frac{c}{2a}.$$

Rearranging for the width of the waveguide yields

$$a = \frac{1.5c}{2f_o} = \frac{1.5(3 \times 10^8)}{2(60 \times 10^9)} = 3.75 \times 10^{-3} m = 3.75 mm.$$

The height of the waveguide can be taken to be equal to the wavelength to stabilize the results. In this case, the actual cutoff frequency is 40 GHz and the operating range is 50 GHz to 75.6 GHz. For a more efficient simulation, the waveguides were made from perfect electrically conducting (PEC) sheets and a wave port was defined as the inner face of the PEC cap backing the waveguide. The PEC sheets should not make a noticeable difference in performance compared to 3D PEC walls, and were selected for computational efficiency. A FEBI airbox was defined to be $\lambda/50$, or 1 mm in this case, away from the non-radiating sides of the waveguide, and $\lambda/10$ for the radiating portions. The simulated frequencies were swept from 59 GHz to 61 GHz. The first antenna was made and assigned as the transmitter and the magnitude of S_{11} was measured as a reference. The antenna was duplicated and assigned as the receiver and placed 1 wavelength away from the transmitter. A 1W excitation was applied to the transmitter and a 0W excitation was applied to the receiver, both with matching integration lines. The PEC cap on the waveguide was used to ensure that the would be a perfect E field perpendicular to the surface of the cap [97]. There was negligible effect on the S_{11} of the transmitter from adding the receiver, indicating that the two antennas are uncoupled. The antenna setup can be found in Figure 4.7.

Instead of simulating the full sized carton of milk as previously mentioned, it is sufficient to simulate the S_{11} and S_{21} responses of a thin dielectric slab defined with the measurement data from the previous section for the desired frequency range. There was very little difference between defining the relative permittivity and dielectric loss only at 60 GHz and defining it as a list for various frequencies withing the simulation bandwidth. A 5 cm x 5 cm square was created with distilled water first and placed 2 cm away from the simple radar system. 2 cm was selected because this would be the approximate distance that the radar would be placed from the cartons on the conveyor belt. The dimensions of the square were selected such that the width and height were much larger than the wavelength at 60 GHz. The thickness was selected to be equal to one guided wavelength, i.e. the wavelength in the dielectric,

$$\lambda_g = \frac{c_0}{f\sqrt{\epsilon_r}} \tag{4.16}$$

where c_0 is the speed of light in a vacuum, f is the desired centre frequency of operation, and ϵ_r is the relative permittivity at that frequency. Using the distilled water target data

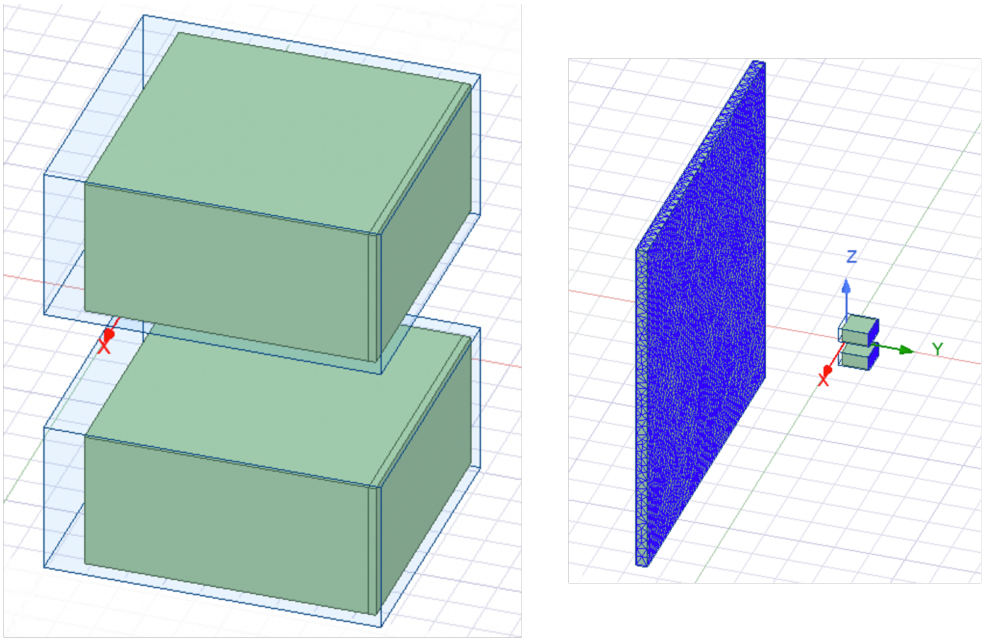


Figure 4.7: Simplified waveguide radar system in HFSS.

from the DAK-TL database at 60 GHz, the wavelength in distilled water is

$$\lambda_g = \frac{3 \times 10^8}{60 \times 10^9 \sqrt{11.1462}} = 1.498 \times 10^{-3} m = 1.498 mm.$$

This calculation process was repeated for the 4 milks being tested. The relative permittivity was taken from the DAK-TL measurement data from the previous section at 60 GHz. The calculated values are displayed in Table 4.3. The thickness of the dielectric slab was changed for each product, and the material was changed to the measured permittivity and loss tangent at 60 GHz accordingly. There are a few noticeable resonances in the simulated results, but more importantly the reflections are differentiable.

The simulated results give an idea of what to expect as the electromagnetic waves reflect off the boundary, alongside the attenuation throughout the material at one wavelength. Increasing the thickness of the dielectric slab did not change the simulated response by very much and the values were still roughly the same. It was determined that 1 wavelength was suitable for this simulation, and if the slab is too thick the simulator will run into insufficient memory problems.

Table 4.3: Calculated wavelength inside dielectric based on measured relative permittivity

Item	**Relative Permittivity,** ϵ_r	**Guided Wavelength,** λ_g
Distilled Water	11.1462	1.498 mm
0.1%	10.7485	1.525 mm
1%	10.0855	1.574 mm
2%	9.742	1.602 mm
3.25%	9.625	1.612 mm

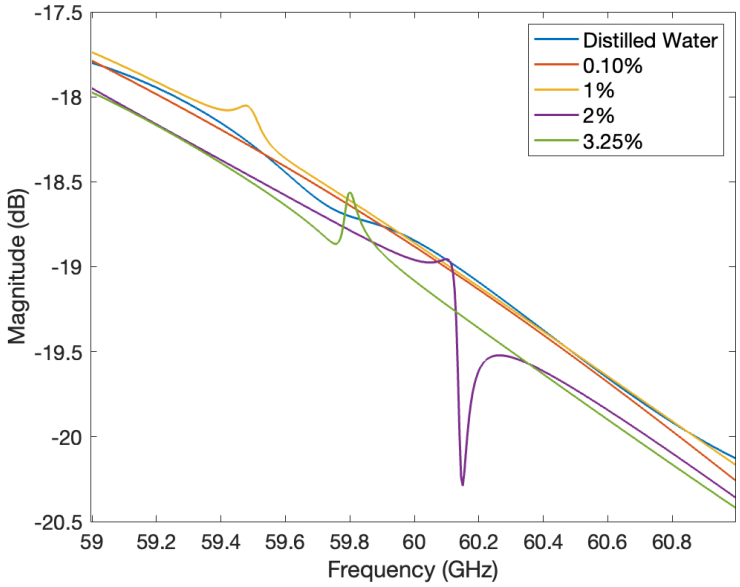


Figure 4.8: Magnitude of reflections of various milk products at 60 GHz in HFSS.

4.2.2 SBR+ Simulation

The shooting and bouncing rays solution method emits a cluster of rays from a source and then calculates the scattering field based on the behaviour of the rays [98]. Ray-tracing techniques are based on physical optics and are useful for higher electromagnetic frequencies/smaller wavelengths as the waves behave similarly to light [99]. Thus, in using this solution type, the following assumption is made: the propagation of radio waves are

straight lines that can be refracted, reflected, diffracted, and scattered [99]. Since each ray has to be accounted and solved for, this solution type is much more computationally heavy. Using a denser amount of rays would improve accuracy, but also slow down the computation time. However, because this is the only solution type that HFSS allows to use a GPU, the computation time is drastically reduced.

To simulate a full-sized 1L carton both stationary and moving, it would be more efficient to use the SBR+ method in HFSS. The reason is that the SBR+ method is able to make use of the GPU on a computer as mentioned and only simulates the electromagnetic fields radiated by the radar as rays, which simplifies the problem of solving the problem several times over finite elements. A vertically polarizing parametric beam with a 30 degree vertical beam width, a 60 degree horizontal beam width, and a footprint size of 0.1 mm was created in HFSS and placed at the origin. A rectangular box of 19.5 cm by 7.5 cm by 7.5 cm was created and placed along the x-axis with the closest edge to the origin being placed 2 cm away. These dimensions were measured from a 1L carton and excluded the triangular top part of the carton. The SBR+ solution setup mimics the Infineon radar settings (Table 4.1) as closely as possible and transmitted 1W of power. The model setup is shown in Figure 4.9.

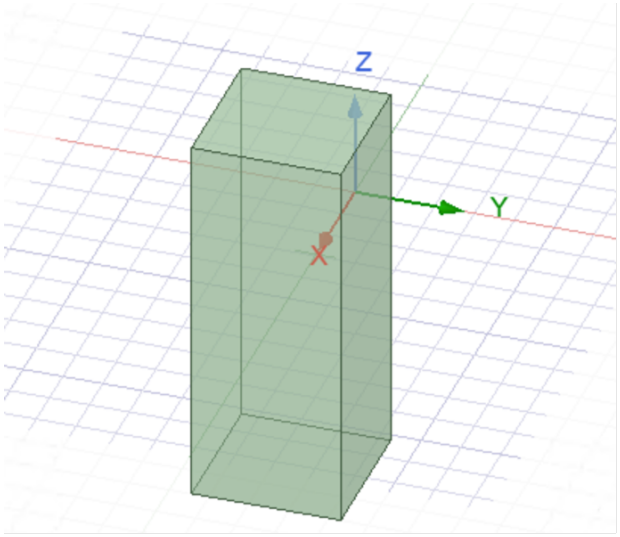
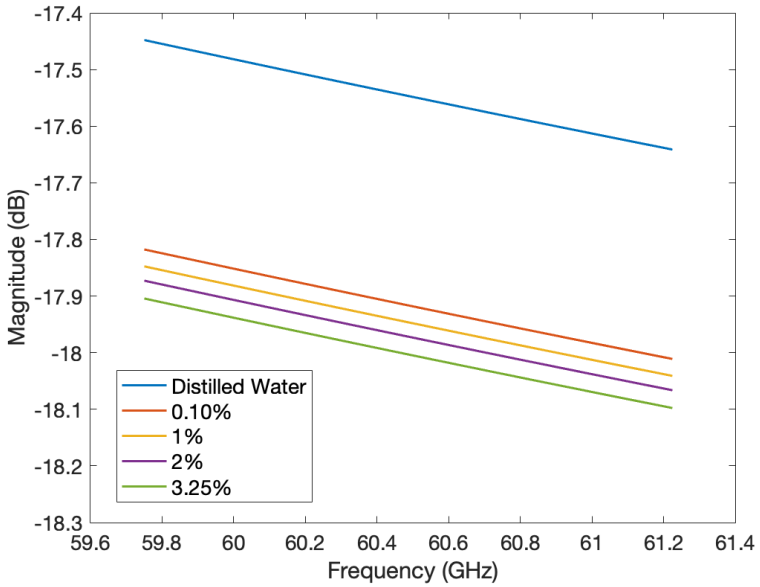


Figure 4.9: Carton model in HFSS SBR+.

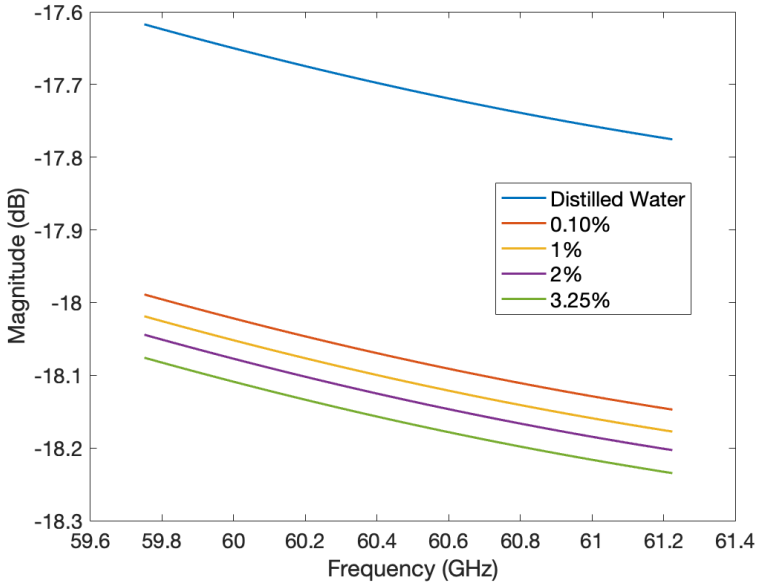
Although inaccurate and unrealistic, the cartons themselves were made with 0 thickness such that it is just a material with the defined dielectric properties floating space. A reason for this was that there was no material property in HFSS for cardboard. All the material

properties in HFSS are at DC, which would render more inaccuracy in the simulation if these values were used anyways. This is to see whether or not the radar can pick up differences in the material itself and not due to the material of the carton. An effort to use distilled water, fresh water, and salt water from the list of HFSS material properties was made, but the attempt was unsuccessful. The radar was barely be able to tell the difference between the three types of water at 60 GHz. This demonstrates the importance of having dielectric data for milk at 60 GHz.

A second instance of the radar was set up as the receiving antenna with 0W. The MoveIT ACT Extension was used to add a 20 cm line parallel to the y-axis. The carton was set to travel along this path with a speed of 0.143 m/s, which was the speed of the conveyor belt during the conveyor belt testing. There is no particular reason why this speed was selected, but this experiment could be repeated for other speeds as part of a future work to determine the limitations of the radar. The calculation of this linear speed is described more in the moving carton setup later. The S-parameter measurements plotted and presented are all taken when the carton was directly in front of the radar. In the case of the stationary test, the same line was drawn so that a parametric variable was generated and swept across, but no object was assigned to move along it. A comparison of the stationary and moving responses are shown in Figure 4.10. From the simulated data, it appears that the magnitude of reflection from a moving carton is approximately the same as when it is stationary.



(a) Stationary case.



(b) Moving case.

Figure 4.10: Magnitude of reflections of various milk products at 60 GHz in HFSS SBR+.

Both plots agree with Equation 4.10 in that reflections are greater with larger permittivities. Distilled water has the highest permittivity of the samples and has the highest amount of reflection. 3.25% milk has the lowest amount of reflections and more attenuation because it has the lowest permittivity. Theoretically as seen in Equation 4.4, the reflections should be the same between both the stationary and moving cartons as the received power is not changing since they are both the same distance away. The slight difference between the stationary and moving cartons may be due to the time steps not being fine enough in the moving simulation such that the radar is directly and perfectly in front of the radar. The differences between the two, however, are small enough to be negigible. Additionally, the trend of the signals tend downwards as frequency increases because as frequency increases the signal attenuates more as it propagates through space. Less of the transmitted signal is making its way back to the receiver.

A small packaging study was also conducted and presented in Appendix E. The conclusion of the packaging study presented shows that the HFSS SBR+ method is not great for studying reflections off multiple stacked dielectric slabs.

4.3 Machine Learning

The final piece of background information needed before diving into the experimental setup and results is the ML algorithm. The ML algorithm is a key factor in determining the accuracy of the radar when it comes to differentiating the milk samples, because the scattering picked up by the radar signals are nonlinear it is difficult to generate a physics based model for received signals. Without a ML algorithm, there would only be raw magnitudes sampled by the radar and no way to tell what signal is what. The OLS linear regression was used in Section 3.4.5, which is better for dealing with continuous predictions involving a value while the techniques presented here will be better for dealing with discrete categories. The strategy is similar to the linear regression analysis presented to predict butterfat percentages given a dielectric constant model or dielectric loss model, but this time to sort the measured radar signals into different milk fat percentages. Classification is selected for the radar and carton analysis because the number of discrete milk products used is not large enough to provide a well-rounded guess on the milk fat content in the cartons. For this, better resolution on the radar may need to be known and there would need to be a wider selection of discrete products.

The following classification techniques will be briefly discussed and compared: support vector machine (SVM), random forest (RF), and k-nearest neighbours (KNN). All of the implementations can be found in Python's scikit-learn or sklearn package. Understanding

these algorithms in depth are outside the scope of the project. These are 3 commonly used ML algorithms that are used for example, but the scope of optimizing the ML in a future work should not be limited to just these 3.

Support Vector Machine

One way to classify the collected signals corresponding to their milk fat percentages is using the SVM algorithm. It is a traditional supervised ML method known for its simplicity and computational efficiency to handle small datasets where the number of features is much greater than the number of samples [100]. In particular, the linear support vector classifier (SVC) method from the sklearn package was used. In the final implementation for industrial purposes, the kernel type may change to better fit the problem at hand.

In the SVC method, a decision boundary is created by the algorithm where the margin between each class of data is maximized. That is, the algorithm solves the following convex optimization problem with a constraint [101]

$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i \tag{4.17}$$

subject by

$$y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i \qquad i = 0, ..., n, \quad \zeta_i \geq 0 \tag{4.18}$$

where x is a feature vector, y is the class label vector, w is a weight vector, C is an inverse regularization parameter, ζ is the distance a sample is from their correct margin boundary, n is the number of training samples, b is the independent intercept term or bias term, and ϕ is the identity function [101], [102]. Figure 4.11 illustrates an example of a simple SVM case using two classes.

The characteristics of a dataset can be considered as features. These features are given a class label when sorting the data. The classifier draws a hyperplane in an N-dimensional space during the training phase given by Equation 4.18, which is some distance away from the decision boundary. The positioning of these hyperplanes maximize the margin boundaries between classes which corresponds to minimizing the weight vector in Equation 4.17. By minimizing the weighting vector, the weight of each feature has a smaller impact on the decision boundary. At the same time, to model any potential errors in the dataset, a slack variable (ζ) is introduced. This way, the algorithm is more robust in handling outliers. During the testing phase, a dataset will be classified depending on its position in space with respect to the decision boundary.

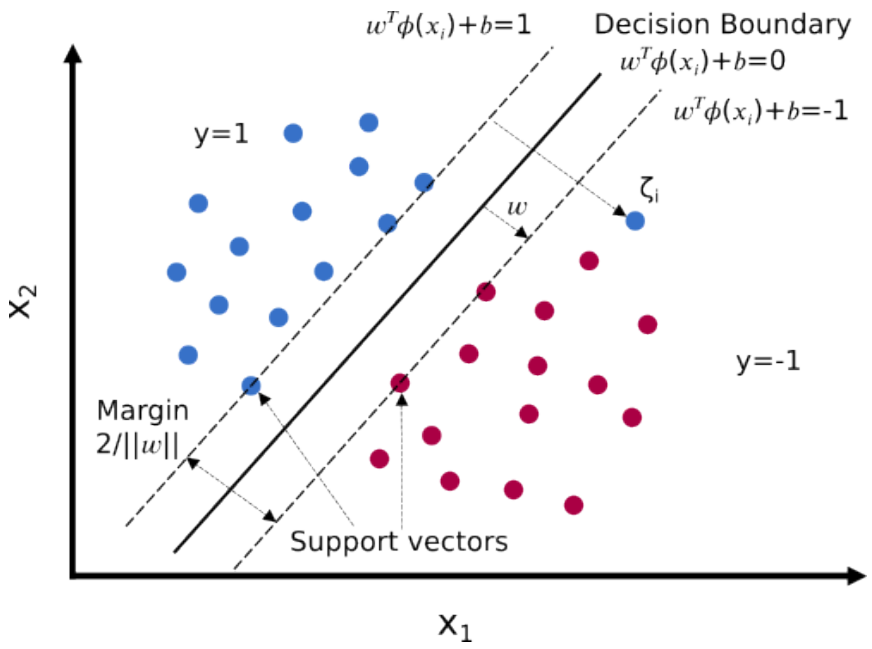


Figure 4.11: Illustration of support vector machine.

Random Forest

The RF algorithm is an ensemble learning technique that depends on multiple decision trees and each tree contributes a vote towards the classification of a dataset [103]. The prediction of a dataset is then made up of a majority vote among the individual trees. Because of its reliance on deep trees and individual overfitting, this algorithm is computationally expensive [104]. A tree structure consists of leaves and branches where the leaves represent the class labels and the branches represent the different features associated to that class [105]. This method is less analytical the the SVM and the KNN algorithm, because each individual tree makes a prediction based on flowchart structure, for example, will it rain: yes or no. In this case, consider an unknown milk product that needs to be classified into 4 different categories. This would act as the root and it would branch out into the different features. If the dataset matches particular subset of features, then the path branches towards a leaf that specifies the category or class label. Because the RF algorithm accounts for many of these decision trees, it is great for handling datasets with a large number of features and is highly scalable. However, this algorithm is also slower and computationally expensive because of the number of trees needed.

K-Nearest Neighbours

The KNN algorithm is an instance based learning technique which does not depend on a training phase [103]. Because of the lack of training phase, the algorithm is computationally intensive when it comes to large datasets because of their predictions on the go, leading to poor scalability. Datasets are classified based on their closest neighbour in a feature space, or k-nearest neighbour where k represents a positive integer. The value k represents the number of nearest neighbours that the dataset sees and will be classified based on a majority vote.

The biggest challenge in using the KNN algorithm is the determination of the value k for a model. As k increases, the predictions offered by the algorithm become more stable because of the larger number of close neighbours the dataset is being compared to. The value $k = 1$ will always yield the most accurate values if the dataset is very easily distinguishable, but is also the least stable because the dataset is only being compared to one near neighbour in the training set. There are a few ways that the algorithm determines a near neighbour, but a common method is based on Euclidean distance specified by Equation 4.19. The majority vote is done by estimating the conditional probability of an outcome given some input in Equation 4.20 where $I(y_i = j)$ is an indicator variable.

$$d(x, y) = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2} \tag{4.19}$$

$$Pr(Y = j|X = x_0) = \frac{1}{K} \sum_{i \in N_0} I(y_i = j) \tag{4.20}$$

$$I(y_i = j) = \begin{cases} 0, & \text{if } y_i \neq j \\ 1, & \text{if } y_i = j \end{cases}$$

The resolution of the radar and whether it can differentiate between 1.80%, 1.90%, 2.00%, etc. is unclear right now. The milk fat content of each carton is taken as what was printed on the carton, as no instrument is current accessible to give the exact milk fat percentages. A summary of the presented ML classification algorithms can be found in Table 4.4 based on [103].

Table 4.4: Summary of machine learning techniques

Machine Learning Technique	**Approach**	**Advantages**	**Disadvantages**
Support vector machine	Draws an optimal decision boundary maximizing margin between classes	Simple implementation and interpretation, computationally efficient for small sample sizes, handles high dimensionality well, less prone to overfitting	Computationally expensive for large datasets
Random forest	Multiple decision trees where prediction happens on a majority vote	Handles high dimensionality, less prone to overfitting, high scalability, high accuracy	Slower classification with large datasets, computationally expensive, harder to interpret
K-Nearest neighbours	Predictions based on nearest neighbours in feature space	Simple implementation and interpretation, no training period, works well for small to moderate sized datasets	Computationally expensive classification phase, high computational complexity, scales poorly

4.4 Proof of Concept

Dielectric characterization and simulations have demonstrated that the milk is differentiable based on milk fat concentration using precise instrumentation or a simple radar. The next step in the process is using the radar system for dielectric characterization.

4.4.1 Stationary Testing

Both the Infineon radar and the SOLI radar were used for the stationary testing part of the proof of concept. A considerable amount of time has been put in collecting data using both radars to determine which is best suited for this application. In the end, it was determined that the project would proceed with the SOLI radar because of its higher sensitivity.

Infineon Radar

The Infineon radar features 1 transmitter and 3 receivers and a 4 GHz bandwidth. A lengthy study was performed with the radar, including thermally characterizing the device with an infrared camera which will be presented in Appendix F. For the stationary carton measurements, the radar was mounted to a handmade wooden fixture shown in Figure 4.12. The wooden fixture consisted of a 90 degree angle that would help align the corners of the carton for consistent positioning. To test for repeatability, the wooden fixture with the mounted radar was removed and placed back by aligning the corner of the fixture with the corner of the carton. It was later realized that the design of this experiment was not ideal because the side of the carton was not perfectly flat and would have a bulge. When aligning the flat surface against the bump, the radar could have been placed at slightly different angles each time. The reason the radar was moved instead of the carton itself was because 2L cartons were measured as shown in Figure 4.12. The cartons were relatively heavy compared to the wooden fixture and kept knocking it out of place. Each carton was measured 3 times and averaged according to Equation 4.21 with N being the number of trials. Repeatability among the 3 trials was kept as close as possible and outliers were discarded.

$$v_{AVG} = \frac{1}{N}\sum_{i}^{N} v_i \qquad 1 \geq i \geq N \tag{4.21}$$

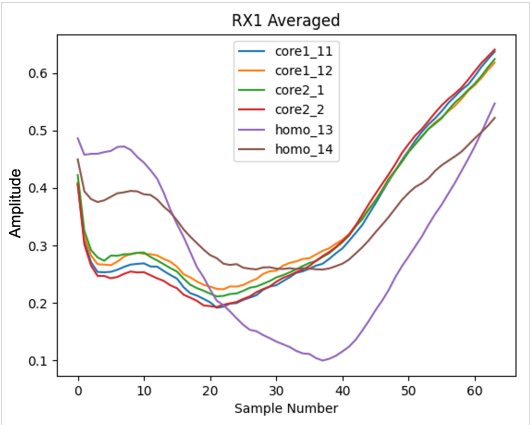
17 2L cartons were measured, 5 regular products (core) and 12 microfiltered products. Results are displayed in Figures 4.13 and 4.14. The Python plots were generated with a Python script that was provided to plot the data obtained by the radar. The naming convention is as follows in Table 4.5 while the number following the underscore represents the sample number in the order of testing.



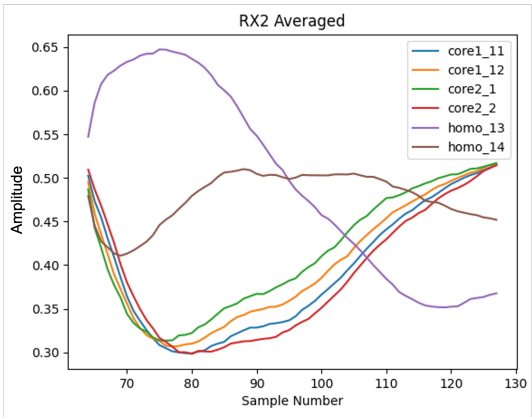
Figure 4.12: Infineon radar and setup with wooden fixture.

Table 4.5: Legend labels defined for data collection using Infineon radar

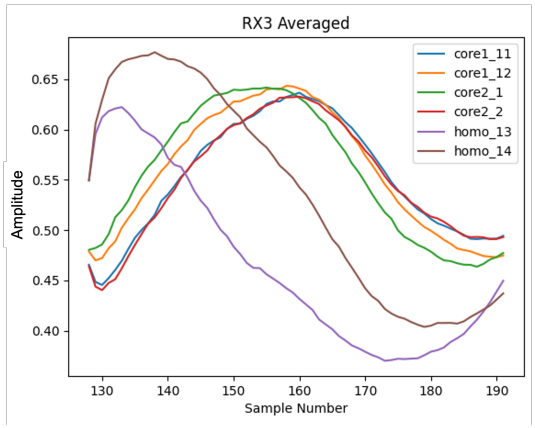
Label	**Product**	**Type**
core1	1%	Regular/Core
core2	2%	Regular/Core
homo	3.25%	Regular/Core
ntlh	Nutrilait 3.25%	Microfiltered
tt1	TruTaste 1%	Microfiltered
tt2	TruTaste 2%	Microfiltered
tth	TruTaste 3.25%	Microfiltered
tts	TruTaste 0.10%	Microfiltered



(a) RX1

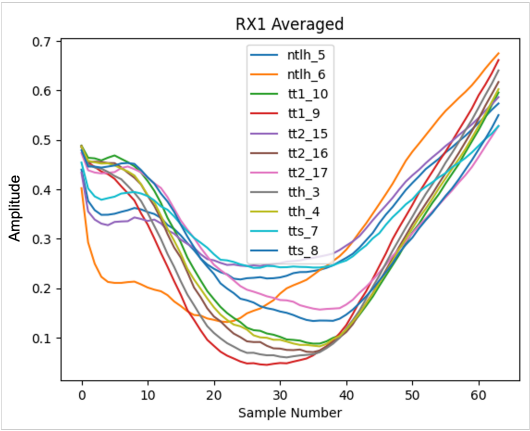


(b) RX2

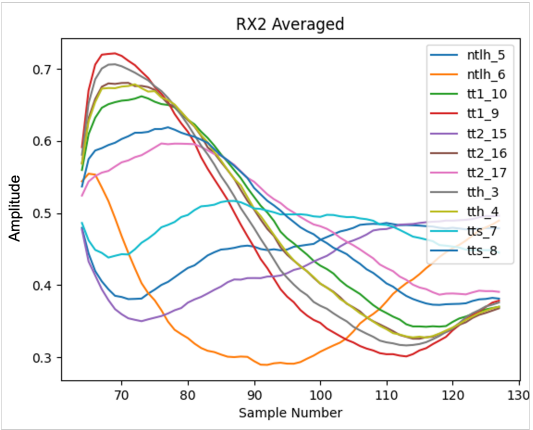


(c) RX3

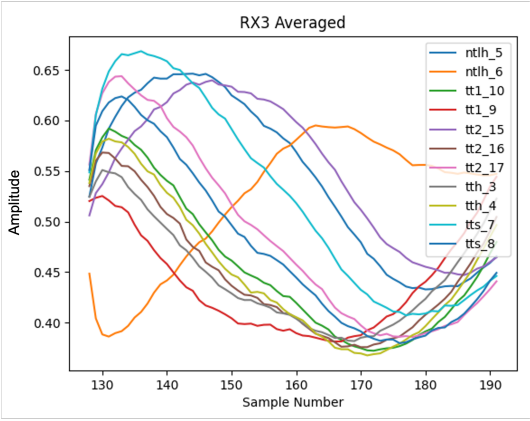
Figure 4.13: Averaged magnitude of reflected responses of regular/core milk.



(a) RX1



(b) RX2



(c) RX3

Figure 4.14: Averaged magnitude of reflected responses of microfiltered milk.

In Figures 4.13 and 4.14, the top row displays the measured reflections at the 3 receivers of the core/regular milk products while the bottom row displays the microfiltered products. Examining the plots, the first most noticeable aspect of these graphs are the spread and inconsistency of the results. It is difficult to identify a consistent pattern that would correspond to a particular milk fat concentration as all the responses seem jumbled up. It may be possible that a ML algorithm would be able to classify and sort these signals in better detail, but this was not attempted. A good portion of the signals were close together, which may be difficult for the algorithm to pick up on. For example, in the top row, the responses appear to be nearly overlapping aside from the 3.25% outlier data. In the bottom row, the spread is larger in the microfiltered products most likely due to the thickness of packaging. The radar is on millimetre order, and should not be able to detect the processing differences between a regular and premium product. The cartons used for premium products are slightly thicker than the thickness of the cartons used for the regular ones. This is to improve shelf-life on the premium products. Thus, if an algorithm were to be implemented, it would need to be able to take into consideration this thickness and differentiate the same milk fat concentration at the end of the day.

Another possible issue that affected the measurements was condensation forming on the cartons after it was taken out of the fridge. The cartons were wiped dry before testing, but the condensation was recurring and more condensation formation was unpreventable. If the measurements were performed in the fridge, simulating the temperature of the production area rather than at room temperature, the effect of condensation would be minimal.

After several more experiments, including adding foreign objects and liquids to the milk samples and testing other food products as well as beverages, it was decided that the Infineon radar required higher sensitivity. It did not have the resolution to detect these small changes very well, especially since milk is mostly comprised of water which is a lossy material. To address this issue, an idea of attaching a power amplifier to the transmitter was proposed the ongoing work section. Unfortunately, due to time constraints and limited resources, the topic of power amplifiers for the radar remains an early work-in-progress.

SOLI Radar

After experiencing difficulties in differentiating the milk cartons consistently with the Infineon radar, it was determined that higher sensitivity was needed. Since the SOLI has a wider bandwidth and utilizes 2 transmitters and 4 receivers, it was strong candidate. The additional receiver helps to improve the signal-to-noise ratio (SNR) of the measurements and add an an extra dimension to the data. The improvement of the SNR can be thought of as modelling the noise as resistances in parallel, which decrease with more

parallel branches. When examining the data, it is important to look at every receiving channel to avoid missing key characteristics that can be picked up by each of them, such as the angling of the carton. Initially with the SOLI, the same problems persisted as seen in the Infineon radar. To address these issues, any movement and variation in packaging was carefully minimized. For example, instead of moving the fixture that the radar was firmly attached to, the carton itself was moved instead and placed back carefully.

Two major tests were performed here, where the first test was performed with a single glass jar and the second with the original packaging. In this subsection, measurements of milk fat concentrations of 1%, 2%, and 3.25% are reported using the SOLI. Due to supply chain shortages, 0.10% was unavailable at the time for purchase. The testing was done by initially measuring the milk in their respective cartons and then measuring the milk in a 1L glass jar with flat edges. All the glass jar measurements were done with the same jar to eliminate any variations in the jar itself. 6 1L cartons were tested and 2 cartons per type of milk. This was to see if there is consistency in the signal responses from 2 different cartons of the same milk and what would happen if they got transported to a glass jar. Because the radar is extremely sensitive to displacement, it is difficult to get a repeatable reading. 10 trials were taken for each measurement to help account for this.

For the milk carton measurements, the bottom of the SOLI was 11.5 cm off the table and the top of the SOLI was 18.5 cm off the table and 1 cm away from the carton. The setup is displayed in Figure 4.15 and the repeatability results for a carton of 3.25% milk is shown in Figure 4.16. Extra care was taken to ensure that the environment was kept as spacious as possible to avoid scatterings from the surroundings. Pieces of cardboard were used to help align the corners of the carton when performing the repeatability test and for consistency with the other carton measurements. In Figure 4.16, each curve represents a trial from the repeatability test for a 3.25% Sealtest milk carton.

From Figure 4.16, it can be seen that there is good consistency in the shape of the measurements but with slight shifts of the plots. Although the measurements were performed as carefully as possible, there may have been slight variations in the carton placement in terms of angle and the distance from the radar. The fact that the signals are more spread out in RX2 compared to the other receivers may suggest that the angle of the carton is slightly off. Since 10 trials were taken of the same carton, it should be a large enough sample size to determine an averaged signal, especially if the shape of the measurements are roughly consistent.

The test was repeated for the 1% and 2% milk cartons while taking 10 trials each for repeatability. Each had their own unique shape that was preserved among the trials, showing high consistency. The results were averaged over each receiver and displayed in



Figure 4.15: Milk carton positioned in front of SOLI for stationary test.

Figure 4.17 with their unique shapes preserved. Judging from the plots, it would most likely not be possible to differentiate the cartons based on a single receiver alone since there could be some overlap in some of the receivers but not in others. For instance, 1% Sample 1 and 2% Sample 2 look quite similar in RX1, RX2, and RX3, but different in RX4. Meanwhile, 1% Sample 2 looks similar to 3.25% Sample 1 and 3.25% Sample 2 in all of the receivers but with a bit more differences in RX2. Sensitivity of the radar could always be improved by using an array of them, a different module, attaching a power amplifier, and many other ways. Noticeably, 3.25% seems to be most consistent among all the cartons doing this stationary test.

In the case each of the cartons had some differences between them, for example if the bulge was larger in one carton than the other for the same milk type, the obtained results would have differences. To eliminate these differences, using the same container or packaging would be ideal to eliminate any of these variances. Glass was chosen because it has a low dielectric constant and is transparent. A single 1L glass jar was reused for each of the samples to eliminate any differences in the glass jar itself. For the jar measurements, the bottom of the SOLI was 5 cm off the table and the top of the SOLI was 12 cm off the table. Placing the radar at the same height as previously for the stationary carton measurements would have been too high and measuring the rim of the jar and air. The setup is displayed in Figure 4.18. After measurements, the milk was poured back into the

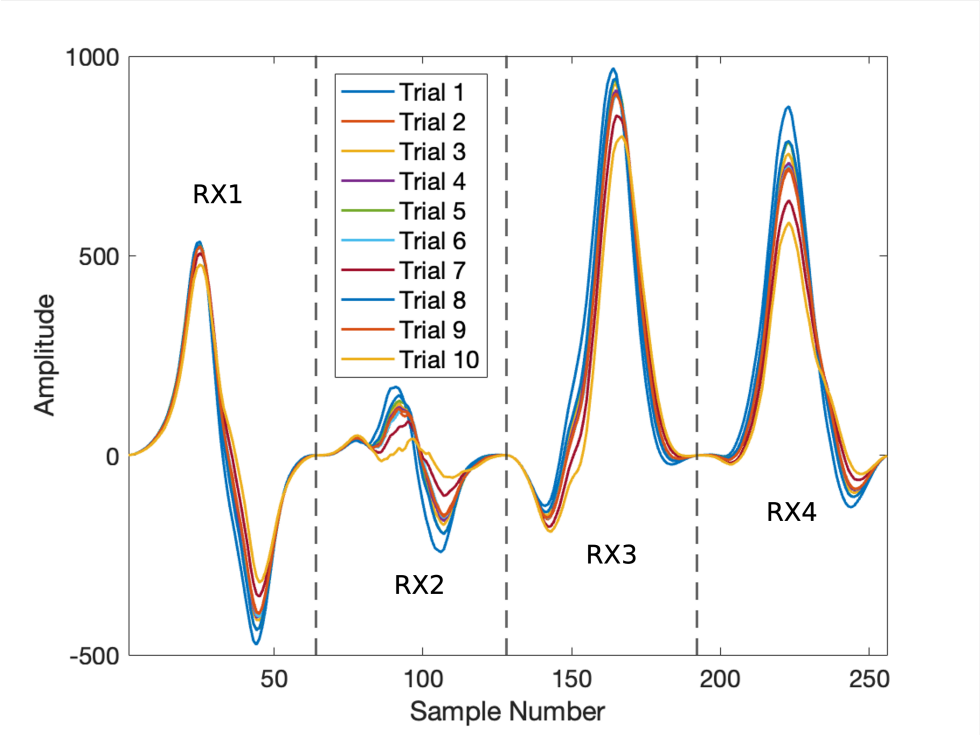


Figure 4.16: Magnitude of reflected responses from a 3.25% milk carton using SOLI radar.

carton and the jar was cleaned thoroughly for the next sample of milk. The results for a 3.25% milk sample are shown in Figure 4.19.

Examining Figure 4.19, it is clear that the trials are more consistent than the stationary carton test. The jar seems to help with the repeatability compared to the carton of milk, illustrating the differentiability of the different milk fat concentrations better. It can be noted that the jar was also placed closer to the radar than the carton in the same setup because the jar was wider than the carton. This could also contribute to the consistency of the data as there is less room for positioning variations. Moreover, the repeatability on the 3.25% milk in both the carton and jar seemed to be better than the other milks. This may be due to the higher concentration of fat content that could contribute to signal propagation and scattering.

Repeating the same measurement process for 1% and 2% milk, the data obtained remained consistent among each other. The plots all had same shape and little spread among the different trials. Perhaps this experiment should have been performed as one of the initial proof of concept experiments as a considerable amount of time has been spent working

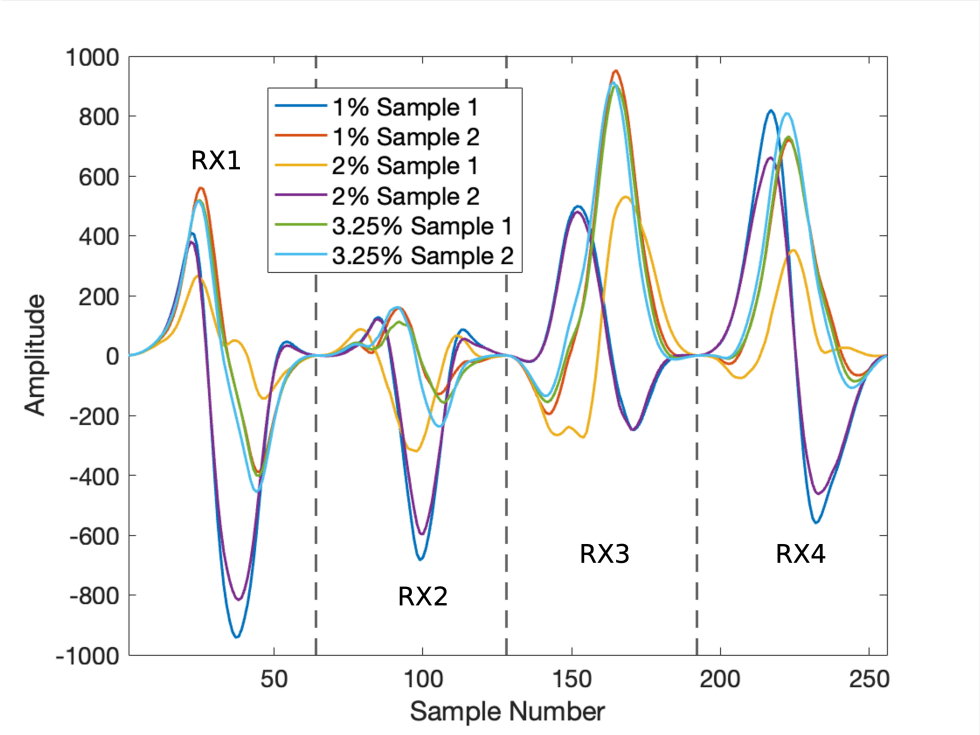


Figure 4.17: Averaged magnitude of reflected responses using SOLI radar.



Figure 4.18: Preparation of glass jar experiment.

with the Infineon radar and the SOLI radar to take clearly differentiable measurements of the different milks. Had this experiment been performed first, perhaps a more in-depth

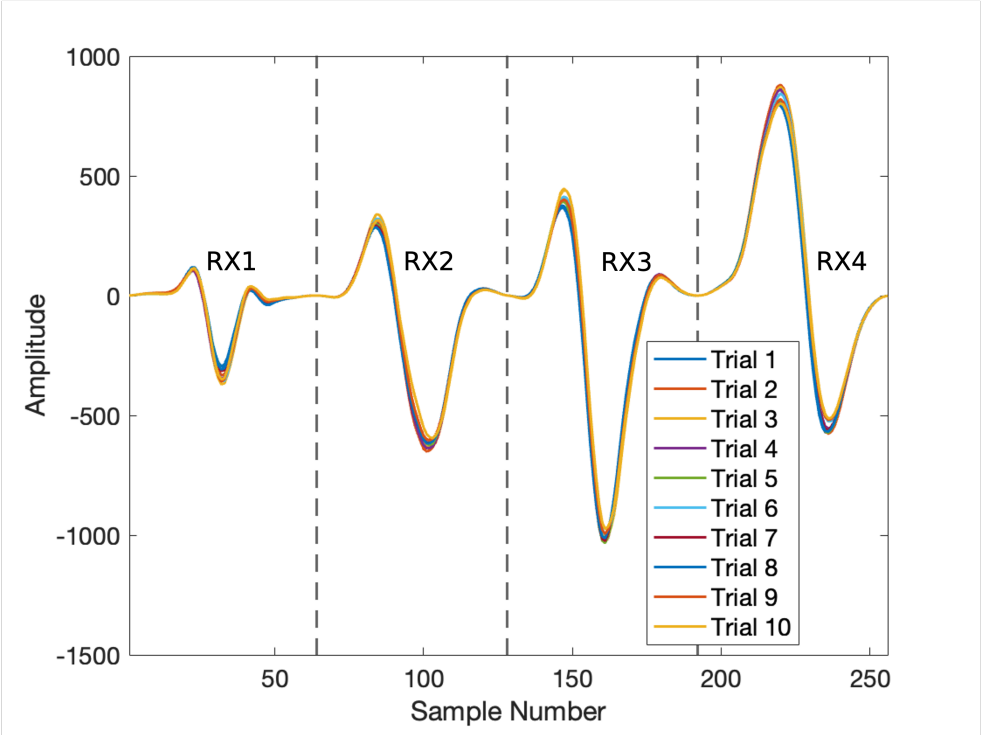


Figure 4.19: Magnitude of reflected responses from a 3.25% milk in a glass jar using SOLI radar.

carton study could have been performed knowing that the milk was indeed differentiable based on the magnitude of their scattering responses. The in-depth study could reveal more information about how the shape, material, angle, and size of the packaging would affect the responses.

The averaged scattering responses from the 6 cartons after 10 trials each are plotted in Figure 4.20. Unfortunately, the scattering responses are not a gradient or linear function of the concentration of milk fat of a sample. However, this is to be expected as the scattering is a highly nonlinear process and is not easily controllable or predictable. From the data collected so far, each of the milks have a unique and repeatable response as seen in Figure 4.19. It is not clear why milks of the same concentration do not overlap each other and each have their own unique amplitude. The cartons selected had manufacturing times right after the other on a production line so the milk itself should not have varied. The differences may be attributed to, again, the randomness in the environment which would be difficult to calibrate out unless the training data for the ML algorithm is large enough.

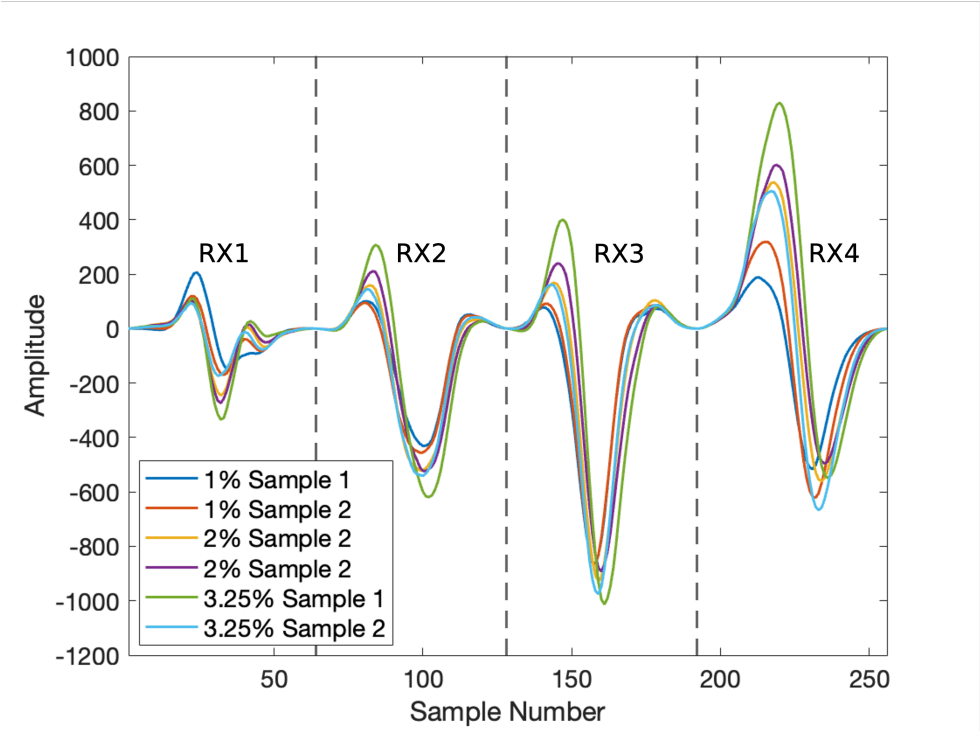


Figure 4.20: Averaged magnitude of reflected responses with glass jar using SOLI radar.

The data collected up until now supports the fact that the milk fat concentrations can be differentiable based on their reflected scattering responses while the carton is stationary. The next step in the proof of concept is to measure the cartons while they are moving on a conveyor belt. This should be possible as well, from the simulated results using HFSS SBR+ though the results should be different than the stationary ones because of the changing environment.

Stationary Test Machine Learning Discussion

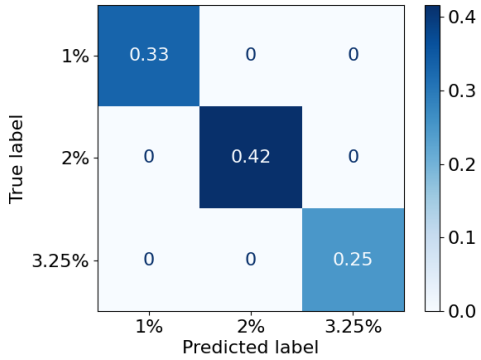
A comparison between the three classification algorithms can be tested and compared using the collected data from the SOLI. This is important to determine the accuracy of the radar in detecting changes in the different milk products before proceeding to the moving carton measurements, which will contain more errors and uncertainty by nature. The accuracy of the ML results obtained will also act as a solid comparison for the moving carton predictions and possible upper bound. That is, if the stationary carton results yield

some accuracy, the moving carton predictions should yield an accuracy that is the same or lower. Since the moving cartons are bound to have more error in them, the same accuracy cannot be expected unless the setup is optimized.

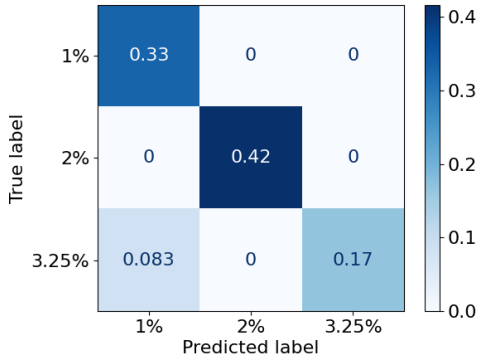
Since no 0.1% milk was tested, this problem will utilize 3 classes (1%, 2%, and 3.25%). There were 64 sample points per chirp and all four receivers on the SOLI were utilized, so each dataset contained 256 points. There were a total of 6 cartons, 2 per class, which were tested for 10 trials each. The train_test_split() function was used to split the 60 datasets up using the 80/20 rule. 80% of the datasets were selected for training and 20% of the datasets were selected for testing randomly. This resulted in 48 datasets for training and 12 datasets for testing. The predictability of this randomness was set by the random state value, which was kept to zero. More details of the code used can be found in the Appendix. The results from the classifier are tabulated in Table 4.6 and the confusion matrices are displayed in Figure 4.21.

Table 4.6: Machine learning predictions for stationary carton data

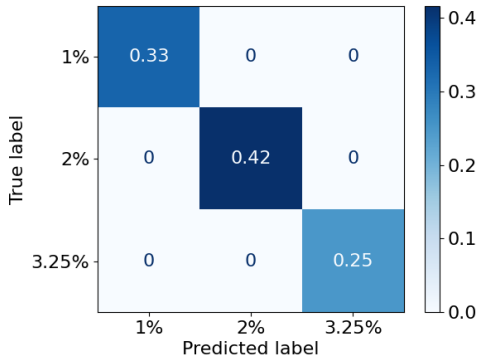
Test Sample	**Label**	**SVM Prediction**	**RF Prediction**	**KNN Prediction**
1	2%	2%	2%	2%
2	2%	2%	2%	2%
3	3.25%	3.25%	3.25%	3.25%
4	2%	2%	2%	2%
5	1%	1%	1%	1%
6	1%	1%	1%	1%
7	2%	2%	2%	2%
8	3.25%	3.25%	1%	3.25%
9	3.25%	3.25%	3.25%	3.25%
10	2%	2%	2%	2%
11	1%	1%	1%	1%
12	1%	1%	1%	1%
Accuracy	–	**100%**	**91.7%**	**100%**



(a) Support vector classifier, accuracy = 100%.



(b) Random forest classifier, accuracy = 100%.



(c) K-nearest neighbours, accuracy = 100%.

Figure 4.21: Confusion matrices for stationary carton data.

From the results of the stationary ML results, the SVM and KNN classifiers both demonstrated a perfect 100% accuracy while the RF classifier demonstrated 91.67% accuracy. The RF classifier predicted a 3.25% carton as a 1% one. For the KNN classifier, a value of $k = 1$ was used because the measured data was relatively distinguishable. The same training set and testing set were used in all 3 cases with the same random state. Regardless of the algorithm used, the accuracies are comparable. The radar signals collected in the measurement were quite different from one another, the reduction in total number of classes, and the larger 10 trial per sample helped improve the accuracy of the models being used. Results of the glass jar experiment can be found in Appendix D.

4.4.2 Conveyor Belt Testing

Based on the collected data from the previous two subsections, the SOLI was chosen for the conveyor belt proof of concept. The radar was fixed to a wooden fixture that was mounted to the side of the conveyor belt and the surface of the radar was situated about 2 cm away from the cartons (Figure 4.22). The positioning of the height of the SOLI was kept the same as in the glass jar experiment. By placing the radar lower than in the stationary experiment, this would minimize scattering effects caused by the curvature in the carton's bulge. Three cartons were measured per type of milk, with 12 cartons in total. Each carton was fixed to the belt with tape and measured 5 times and the best 3 trials were taken for the particular carton to be presented. The first carton was positioned 12.5 cm away from the centre of the SOLI, while the second and third cartons were spaced 16 cm between each other and from the first.

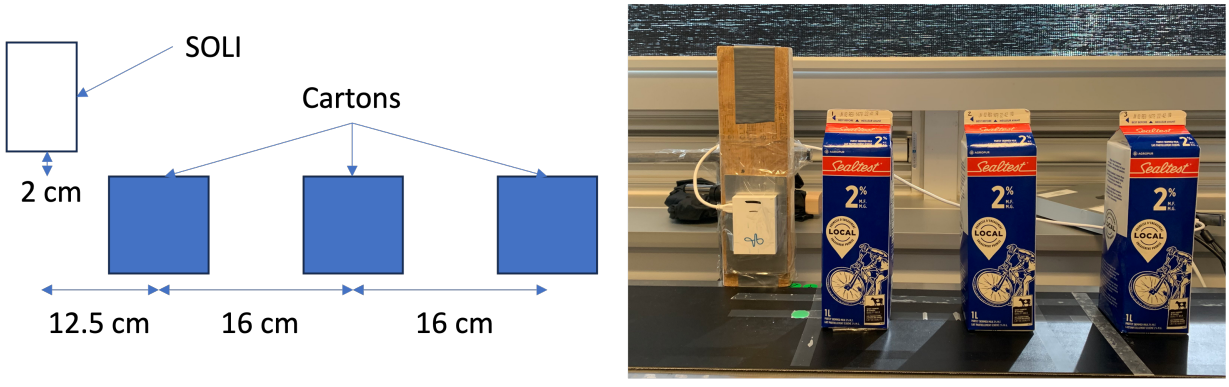


Figure 4.22: Positioning of cartons on the conveyor belt.

The linear conveyor belt speed was set to 0.143 m/s and the cartons were searched

through manually in the dataset using timestamps. As mentioned, there is no particular reason why this speed was selected and the experiment could be repeated for different speeds to determine the limitations of the radar. It is possible that if the conveyor is moving too quickly, the radar sampling time may not appropriately capture the carton signals in the position we want as it is being conveyed. Generally, the conveyor belt speed is fixed but the speed can vary from model to model. The conveyor used was a 59 x 7.8 inch conveyor belt provided by Vevor and it has a few speeds available. The 20 rpm (0.143 m/s) speed was selected. To convert this to a linear speed, the speed of the belt was calculated based on measuring the speed of an object moving from point A to point B. In this case, it was a round trip on the conveyor belt where the length of the conveyor belt is 59 inches or 1.5 m. For a round trip, 1.5 m is multiplied by 2 and the round trip time at 20 rpm was measured to be 21 seconds. Thus, performing a simple linear speed calculation, one can obtain 0.143 m/s.

Each estimated timestamp was calculated based on a simple time calculation using the speed of the conveyor belt and the distance it needs to travel in order to reach the centre of the radar. The first carton was expected to appear centred in front of the SOLI radar at 874 ms, while carton 2 at 2000 ms and carton 3 at 3111 ms. Before running the test, a single carton of each type was centred in front of the SOLI. This stationary result was taken as a "calibration" curve, that is the closest moving carton response to this stationary carton will be taken and presented. Although the datasets could be fine tuned better and more carefully, manually preparations in Excel of the data were sufficient for a quick proof of concept. The radar could be programmed to take a some frames every few seconds, which may have been a more efficient method of sorting and collecting the data. This feature was attempted but not fully implemented because the results would still need to be searched for in the dataset regardless. Additionally, there were some chunks of missing timestamps in the generated .CSV files, about 0.043 seconds every so often, which may be due to sampling time. The closest response timestamps are presented, but better ones may not have been recorded in the .CSV file.

The response curves for 0.10% milk, or skim milk, are displayed in Figure 4.23 with all the receivers shown on one axis. The dashed lines represent the stationary curves that are used to pick out the associated moving carton response. Recall only the best 3 trials of 5 are selected. The reason for this inconsistency could be due to the missing timestamps in the .CSV file which could contain the best fitted signal to the data or some slight displacements of the carton over the span of testing. When the conveyor belt starts or stops, the cartons jerk and wobble in place, getting shifted. This was minimized as carefully as possible by placing the radar lower on the wooden fixture and by taping the cartons to the conveyor belt. Luckily, the amount of wobbling experienced may not be as severe as on an actual

production line thanks to guard rails holding the cartons in place.

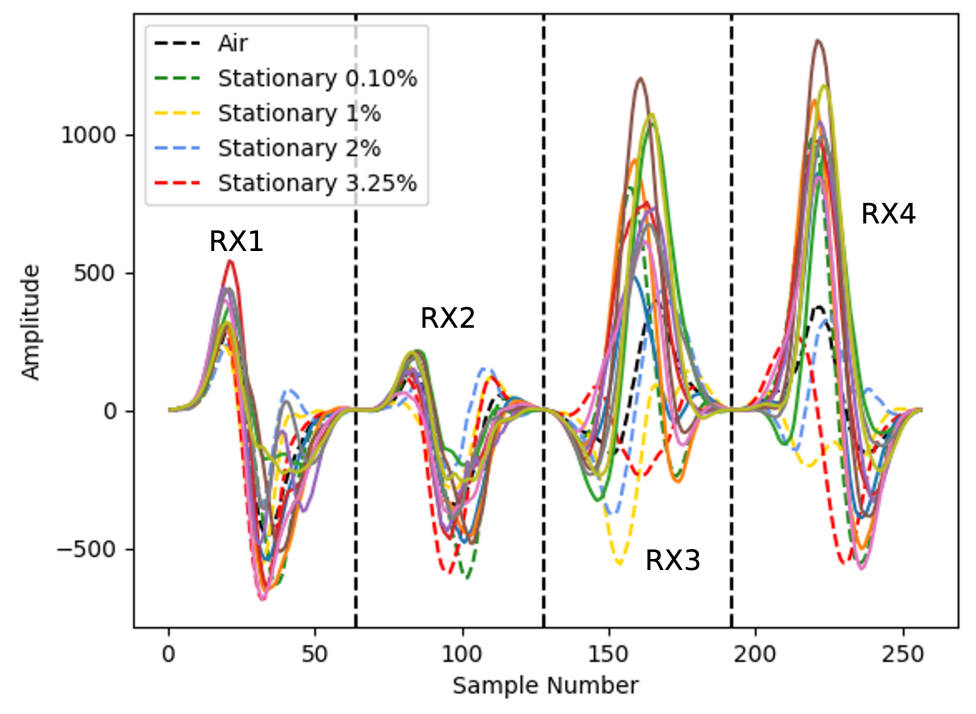


Figure 4.23: Responses of 0.10% milk cartons on a conveyor belt.

The experiment was repeated for 1%, 2%, and 3.25% and the results are displayed in Figure 4.24. Visibly from the plot, there are differences between the signals of the different types of milk. This is interesting because the signals can nearly be manually classified by eye, which means the ML algorithm should not have any difficulty in classifying each of the signals. Another point to mention is that, similar to the stationary testing, all of the frames need to be examined and used in the classification algorithm. A general linear or gradient trend as predicted from theory is difficult to identify here. Looking at receivers 3 and 4 on the second half of the x-axis shows that there is larger spread in the data compared to receivers 1 and 2 on the first half of the x-axis. This may be due to the positioning of the antennas in Figure 4.6. Since receivers 1 and 2 are further from the transmitting antennas, there is a longer geometric distance for the transmitted and reflected signal to travel. The additional distance attenuates the signal more, causing receivers 1 and 2 to read a lower magnitude than receivers 3 and 4. Taking into account all the receivers provides more information about the carton and the contents inside it.

Examining Figure 4.24 in some more detail, it is noticeable that the differences between

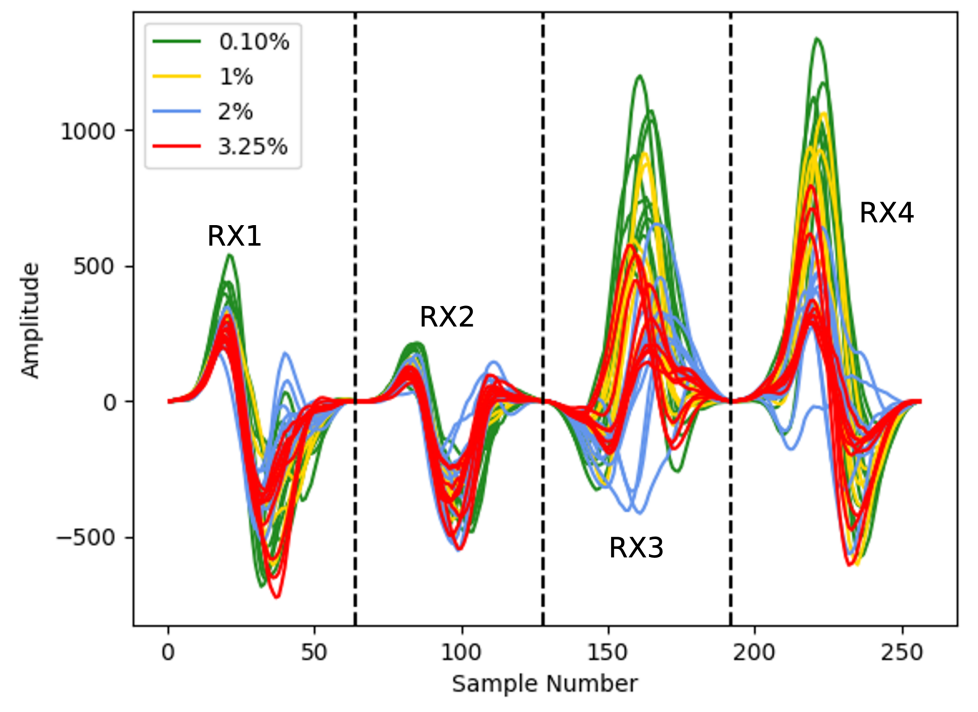


Figure 4.24: Responses of different milk cartons on a conveyor belt.

the signals are much larger for the moving cartons than for the stationary cartons. This may be due to the displacements that occur when starting and stopping the conveyor belt. Though, this was minimized as much as possible. Another possible reason could be due to weight, as a 3.25% carton would weigh slightly more than a 0.1% carton as it contains a high amount of fat concentration. From general physics knowledge, this is because the viscosity of the milk is linearly dependent on the amount of milk fat [106]. Thus, when measuring the moving cartons of milk, although there are more components at play, the viscosity of the liquid itself could reveal a lot of information about the product. Since the cartons selected for this measurement were manufactured at around the same time, the weight of the cartons for each type of milk should not vary drastically. However, there may be a possibility of a heavier 0.10% carton and a lighter 3.25% carton. Whether or not the radar would be able to still differentiate them and if the weight or viscosity is indeed a factor affecting the measurements would be a topic for a future study.

The data collection shows that it may be possible to train the ML algorithm using stationary measurements as a golden sample. The radar could be continuously running similar to the demonstration presented, however the details of efficiency should be left to

a much later stage in an actual implementation. The recorded responses from the moving carton can be capture and compared to the stationary carton response. However, this may not be the most effective method to differentiate signals. Another strategy could be to exclude the stationary cartons and develop a golden sample measurement purely on a multitude of data points which would take a while to collect. In this case, it would be simpler to implement the radar on a production line and collect an extremely large dataset. This would require a proper timing mechanism for the radar captures on the conveyor belt. Pulsing the device on and off may not work well because there are times when the filler is down. Due to time constraints and the tedious measurement process, the experiment was not repeated.

Conveyor Belt Test Machine Learning Discussion

A similar ML algorithm comparison was performed for the moving carton data. The accuracy of the algorithms for this set of data are expected to be lower than the accuracies presented for the stationary carton predictions. The conveyor belt measurements are a lot more prone to errors that have already been described. These issues in the testing could be improved in a future experiment. However, because the data collection and validation were extremely tedious without automation, the automation of this experiment for consistency and standardization are recommended to be improved upon first.

From the conveyor belt testing, there were 9 measured trials for each milk product from 0.10% to 3.25% resulting in a total of 36 datasets. This introduces an additional class that the stationary testing did not have, which adds an additional layer of complexity to the problem that was not account for before. All 256 data points across the receiving channels were utilized. The comparison of the predicted results are tabulated in Table 4.7 and the confusion matrices are populated with the predictions in Figure 4.25.

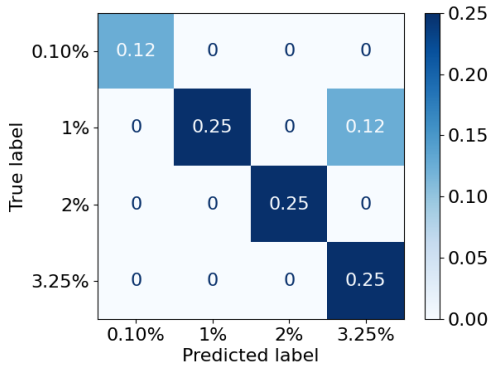
The accuracy of the predictions are expected to be improved with a larger number of datasets to train the classifier along with an optimized and automated test setup. It can be seen from the table and the matrix that a 1% carton was misclassified as a 3.25% carton in SVM, which yielded 87.5% accuracy. Examining Figure 4.24, there is some noticeable overlap between the 1% and 3.25% cartons, which may have confused the classifier. Fortunately, the other 1% and 3.25% cartons were classified correctly as this was the only misclassified set in the batch. The RF and KNN classifiers obtained 100% and 75% respectively. It seems that the KNN classifier struggles more in this experiment compared to before and the value $k = 1$ was taken again because there are very clear differences in the data. The RF classifier's accuracy of 100% does not necessarily mean that this is accurate,

Table 4.7: Support vector machine algorithm predictions

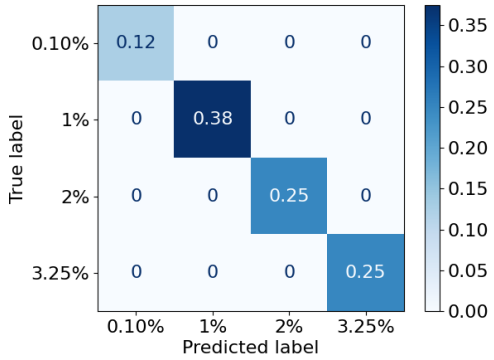
Test Sample	**Label**	**SVM Prediction**	**RF Prediction**	**KNN Prediction**
1	3.25%	3.25%	3.25%	3.25%
2	2%	2%	2%	3.25%
3	1%	1%	1%	1%
4	3.25%	3.25%	3.25%	3.25%
5	2%	2%	2%	3.25%
6	1%	1%	1%	1%
7	1%	3.25%	1%	1%
8	0.10%	0.10%	0.10%	0.10%
Accuracy	–	**87.5%**	**100%**	**75%**

and this only shows that it performed well with this dataset. A larger dataset and more trials would help scale this.

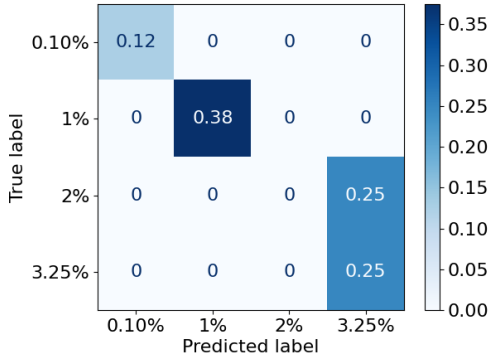
This shows that if a radar was implemented at the side of the conveyor belt, then a classification algorithm can be written to identify the correct product type. Whether the radar will be able to detect slight variations in milk fat content is a question for a future study. Additionally, the current implementation classifies a carton as one of the 4 different milk products and does not introduce a questionable class. Further development of the script will be needed to identify the tolerances of each product before the carton is misclassified or labelled as questionable.



(a) Support vector classifier, accuracy = 100%.



(b) Random forest classifier, accuracy = 100%.



(c) K-nearest neighbours, accuracy = 100%.

Figure 4.25: Support vector machine confusion matrix.

Furthermore, either the SVM algorithm or RF classifier approaches could be selected for future testing as they both maintained high accuracies across the stationary and moving carton testing. SVM could be selected as it is simple to implement and is computationally efficient at handling small datasets with high dimensionality. Meanwhile, the RF classifier could be selected for its better multi-class handling and higher accuracy than SVM in the moving carton measurement. This project did not delve into testing different algorithms for the proof-of-concept, but it may be a good future study to look into. This is the end of the proof of concept and an extensive study has been shown that milk and cream are differentiable starting with their dielectric properties. This would then affect their scattering responses as seen by a radar and a ML algorithm can be used to classify these responses. Further improvements to the experimental setup will be part of a future study.

Chapter 5

Conclusion

5.1 Summary

This thesis presents a simple proof-of-concept for a mm-wave radar system used for the real-time detection of milk fat content in cartons while it is being conveyed. This was done by using the radar for dielectric characterization and examining the unique electromagnetic fingerprints of each milk product. The project was divided up into two main sections.

In the first half of the thesis, an extensive background section describing the physics behind why liquids may be characterized using electromagnetic waves. The Debye equations are described and the theory was applied to salt and sugar solutions. The different constituents of milk are introduced, with the main focus being milk fat. The dielectric theory was applied to milk and cream. The trends were supported throughout literature and through linear regression modelling equations. Theoretical mixing models are explored in Appendix A. After taking the dielectric measurements and verifying that the liquids were differentiable according to theory, the radar operating frequency was selected. The hardware used for testing are also described. Based on the dielectric theory presented, the physics of why liquids may be characterized based on their scattering responses is covered.

The second half of this thesis begins by exploring why a radar was selected and why FMCW would work best. It delves into the reason why the particular radar used was selected as well. The relative permittivity and loss tangents were obtained from measurements using the DAK-TL from the first half of the thesis and used in HFSS simulations of stationary and moving cartons. This was to determine what the radar might see and validate that the different milk products would be differentiable. This half of the thesis focuses

on validating the work experimentally using 2 FMCW radars. The information gathered was further used to narrow down the technology to one FMCW radar after characterizing stationary cartons. The selected radar was then used to characterize the moving cartons on a conveyor belt. From the collected conveyed cartons radar data, a SVM algorithm was implemented and shown to have high accuracy of 87.5%. This was compared with the RF and KNN classifiers.

Overall, numerous other improvements were mentioned throughout this thesis such as an automated timing mechanism for the radar captures on a conveyor belt and expanding testing on other milk fat concentrations to determine the current limitations of the radar. There could also be experimentation with different algorithms for accuracy, but this would require a larger dataset than the one already collected. A large array of radars can be used synchronously to gather more information about the carton in terms of transmission and orientation. The work presented in this thesis successfully demonstrated an in-depth proof-of-concept for non-invasive liquid monitoring on a production line for quality control. This has not been achieved or attempted yet in literature and this type of technology would greatly benefit supply chain companies in food and beverage industries. Although the focus of the work was on milk, the same techniques and physics are applicable to other liquids.

5.2 Suggestions for Future Work

5.2.1 Experimental Improvements

Some recommendations for future testing to determine the limitations of the radar and to provide further improvements in measurement accuracy have been described throughout this thesis. They are summarized below in no particular order:

1. **Automated timing mechanism for the radar captures on a conveyor belt:** The current conveyor belt and radar implementation require the manual starting of both systems. This introduces inaccuracy as both systems may not be started at the exact same time. This could be addressed by introducing a laser that provides the positioning of the carton and to automate the radar capture as soon as the carton is directly centred and in front of the radar. This would help both the repeatability and precision of the conveyor belt measurements.

2. **Incorporating guard rails:** In a more realistic case, the conveyor belt will have guard rails as shown in Figure 1.2 that will help hold the carton in place as it is being

conveyed. These guard rails will have a certain height to them. The stainless steel part of the guard rail may affect the radar measurements by acting as a reflector and the radar placement would need to be adjusted. The guard rail would also prevent the carton from wobbling too much during the radar captures as the conveyor motor starts up, which could improve the accuracy of the measurements.

3. **Expand testing to other milk fat concentrations to determine the limitations of the radar:** The work presented in this thesis classifies milk between 4 classes (0.10% 1%, 2%, and 3.25%). In reality, these gaps are too large and companies would require a finer resolution. It is currently not known whether or not this technology has the sensitivity to pick up the differences between 1.80%, 1.90%, and 2.00% for example. The reason why we could not present this type of data is because a reliable machine that can measure milk fat is not readily accessible. The percentages printed on the carton are taken as is. Much more work would need to be done to determine the resolution and milk samples would need to be verified on an industry specific machine for a proper study.

4. **Extend testing to other types of milk and cream products:** The work presented in this thesis only experiments with regular 0.1%, 1%, 2%, and 3.25% milk. A question may arise is whether or not the filtration process affects the radar readings, for example what are the differences in microfiltered vs ultra high temperature (UHT) treated products vs regular pasteurized milk? Another question that may arise is if this works for such as flavoured milk products, buttermilk, creams, or milk products with a high protein content.

5. **Experimentation with different classification algorithms for higher accuracy and efficiency:** Only one ML algorithm was used in this project and it was the SVM classifier. Other algorithms exist and may potentially present better accuracy with the tradeoff being speed and computational efficiency. Moreover, a larger dataset would be required for proper training and accuracy. To extract and observe gradients in the datasets, advanced signal processing would be needed and would be a whole project of its own.

6. **Use a large array of radars for additional information about a carton:** This project only utilized one radar operating a centre frequency of 60 GHz. The radar on its own provided enough information about the magnitude of the scattered responses, but using a few radars could give information about the orientation of the carton, phase shifts, and the transmission properties of the fluid. The other radars could

be at different operating frequencies or at the same operating frequency. Additional work would have to be done to determine this.

7. **Thermal management of the radar to prevent signal degradation:** Because the radar will be continuously on or pulsing on and off in the final stages of the implementation, the heating of the radar would introduce thermal effects. These thermal effects would lead to signal degradation if left unmanaged. It is important to see how much the radar heats up over several hours and how much the degradation occurs to the signal. After this characterization, thermal resistances or a heat sink could be incorporated into the radar to help maintain a consistent signal quality over time.

8. **Power amplifier implemented in transmission stage to improve measured signal spread:** Because of the radar's low output power, the received radar signals may not be able to give enough information or resolution about the milk. Increasing the transmitted power will improve the sensitivity of the radar by increasing the amount of power that gets reflected back and received by the receiving channels. This would help with the spread of the radar signals across the different types of milk products and help improve the classification accuracy.

9. **Low noise amplifier implemented in receiver path to improve signal-to-noise ratio:** Implementing a low noise amplifier in the receiver path will amplify the received signals above the noise floor, improving the SNR of the system. With a better SNR, the accuracy of the system would be improved.

5.2.2 Infineon Radar Thermal Characterization

To address the second last point in the list of improvements mentioned in Section 5.2.1, some work characterizing the heating of Infineon TR13 radar was conducted. The goal of this experiment was to answer the following questions:

1. How much does the radar heat up over the span of a few days?

2. How much does the radar heat up if it was covered with a 3D printed package?

3. How much do two radars or more heat up when operating together side by side?

4. How does the signal degrade when experiencing thermal effects? How much does the signal degrade by?

This largely remains an ongoing project and task and only the first two items were roughly answered. There were some communication issues with the radars when two or more were operating at the same time and had to be reported to Infineon. This may have something to do with the quality control in manufacturing and fabrication of the radars. Further details and results of the initial experiment can be found in Appendix F.

5.2.3 Sentaurus TCAD Electrothermal Simulation

To address the final point in the list of improvements mentioned in Section 5.2.1, some work in examining the heating of GaN HEMTs are explored through electrothermal simulations in Sentaurus. This remains largely an ongoing exploration. A GaN PA is a potential improvement for the radar performance to increase the sensitivity of the radar by increasing the transmitted power from the output of the radar. Due to its high output power and power density compared to silicon technologies and the access to GaN PAs in the group at around 50-60 GHz, GaN was selected for this exploration. Aside from knowing how the radar thermally performs, it is important to know how the GaN PA thermally performs as well. Although there are many aspects to integrating a GaN PA such as the wirebonds and matching network, the frequency of operation, and others, the thermal aspect is one important parameter. Simulations of a single finger based on [81] were conducted and extended to multifinger GaN HEMTs. Further details of this topic and the simulation results can be found in Appendix G.

References

[1] A. Yavari, D. Georgakopoulos, H. Agrawal, H. Korala, P. P. Jayaraman, and J. K. Milovac, "Internet of Things Milk Spectrum Profiling for Industry 4.0 Dairy and Milk Manufacturing," *2020 International Conference on Information Networking*, pp. 342–347, 2020.

[2] K. Joshi *et al.*, *"Milk Quality Monitoring Using Electromagnetic Wave Sensors" in Sensors for Everyday Life*, vol. 23, pp. 205–227. Springer, 2017.

[3] H. Yeo, R. Minami, K. Rodriguez, G. Shaker, and A. Quigley, "Exploring tangible interactions with radar sensing," *Association for Computing Machinary*, 2018.

[4] M. S. Venkatesh and V. Raghaven, "An Overview of Dielectric Properties Measuring Techniques," *Canadian Biosystems Engineering/Le Genie des biosystems au Canada*, vol. 47, pp. 15–30, 2005.

[5] Statista, "Cow milk production worldwide from 2015 to 2023." [Online]. Available: https://www.statista.com/statistics/263952/production-of-milk-worldwide/.

[6] Statista, "Volume of milk produced in Canada from 2010 to 2022." [Online]. Available: https://www.statista.com/statistics/707340/volume-of-milk-produced-in-canada/.

[7] Food and A. O. of the United Nations (2023), "Milk production, 1961 to 2022." [Online]. Available: https://ourworldindata.org/grapher/milk-production-tonnes.

[8] Food and A. O. of the United Nations, "Dairy Market Review - Overview of global dairy market developments in 2020." [Online]. Available: https://openknowledge.fao.org/server/api/core/bitstreams/11e38125-eeb1-49ba-bbb2-b4bad3ef31cd/content.

[9] R. C. Chandan, *Dairy Processing & Quality Assurance*. John Wiley & Sons, Inc., 2008.

[10] F. Martin-Gatton College of Agriculture and Environment, "Butterfat Percentage or Milk Yield: Which Is More Important for a Dairy Farm's Milk Income?." [Online]. Available: https://afs.ca.uky.edu/content/butterfat-percentage-or-milk-yield-which-is-more-important-for-a-dairy-farms-milk-income.

[11] F. Martin-Gatton College of Agriculture and Environment, ""'Tis the Season" for Improving Butterfat Content of Milk." [Online]. Available: https://afs.ca.uky.edu/content/

[12] FOSS, "FTIR analysis." [Online]. Available: https://www.fossanalytics.com/en/news-articles/technologies/a-short-intro-to-ftir-analysis.

[13] Labexchange, "Foss WineScan Flex." [Online]. Available: https://shop.labexchange.com/en/foss-winescan-flex.html.

[14] CEM, "Smart Trac II Rapid Analysis System." [Online]. Available: https://www.dairyfoods.com/ext/resources/Digital_Brochures/DF-SMART-Trac-II.pdf?1520948731=.

[15] "CEM Smart Trac II Microwave Moisture and Fat Analyzer." [Online]. Available: https://lissci.com/product/cem-smart-trac-ii-microwave-moisture-and-fat-analyzer/.

[16] QuantWave, "Case II: milk testing." [Online]. Available: https://quantwavetech.com/case-ii-milk-testing.

[17] C. S. Manikandababu, P. S. Prakash, V. Vaishnavi, and S. Siddharth, "Empowering Dairy Farmers IoT Milk Quality Monitoring Devices," *2024 International Conference on Advances in Computing, Communication and Applied Informatic*, 2024.

[18] A. E. Omer, J. Chong, L. Chen, and G. Shaker, "Characterization and Sensing of Milk Fat

[19] P. G. Kumar, A. Alagammai, B. S. Madhumitha, and B. Ishwariya, "IoT Based Milk Monitoring System for the Detection of Milk Adulteration," *ESP Journal of Engineering & Technology Advancements*, vol. 2, pp. 6–9, April 2022.

[20] G. Kalaiarasi, K. Dinakaran, and J. Ashok, "IOT Based Milk Analyzer using Arduino with Wi-Fi Module," *2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications*, 2022.

[21] K. Pugazhenthi, A. Sengamalam, and B. Ganesan, "Milk Quality Monitoring System using IoT," *Proceedings of the International Conference on Sustainable Computing and Smart Systems*, 2023.

[22] A. S. Gaikwad, M. S. Ganeshmurthy, and R. Periyasamy, "Development of IoT Based Sensor System for Detection and Identification of Milk Adulterants," *2024 8th International Conference on Biomedical Engineering and Applications*, 2024.

[23] L. Colaruotolo, E. Peters, and M. G. Corradini, "Novel Luminescent Techniques in Aid of Food Quality, Product Development, and Food Processing," *Current Opinion in Food Science*, vol. 42, June 2021.

[24] Infineon, "GBT60TR13C XENSIV 60GHz radar sensor for advanced sensing." [Online]. Available: https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-iot/60ghz-radar/bgt60tr13c/.

[25] P. Instrumentation, "Non-contact FMCW level measurement technology for the food and beverage industry." [Online]. Available: https://www.piprocessinstrumentation.com/measurement/article/21289042/non-contact-fmcw-level-measurement-technology-for-the-food-and-beverage-industry.

[26] Krohne, "Frequency-Modulated Continuous-Wave (FMCW) Radar Level Measurement Systems." [Online]. Available: https://root.krohne.com/fileadmin/content/files-2/PDF-Upload/Whitepapers/WP$_R$$adar - Level - Measurement - Systems_e$$n.pdf$.

[27] U. of Waterloo, "Food quality testing using millimeter waves." US and CA Patent. [Online]. Available: https://uwaterloo.ca/research/sites/default/files/uploads/documents/10177-10223-food-quality-testing-sept-2023.pdf.

[28] A. E. Omer, S. Gigoyan, G. Shaker, and S. Safavi-Naeini, "Whispering-Gallery-Mode Microwave Sensing Platform for Oil Quality Control Applications," *IEEE Internet of Things Journal*, vol. 9, March 2022.

[29] G. Shaker, "Using Radars for Food Quality Control and Authentication." [Online video]. Available: https://www.youtube.com/watch?v=XjqGfVXADDw.

[30] A. E. Omer, G. Shaker, and S. Safavi-Neaeini, "Non-invasive Glucose Monitoring at mm-Wave Frequencies," *University of Waterloo Open Journals*, 2018.

[31] A. E. Omer, G. Shaker, S. Safavi-Naeini, K. Murray, and R. Hughson, "Glucose levels detection using mm-wave radar," *IEEE Sensors Letters*, vol. 2, no. 3, pp. 1–4, 2018.

[32] D. Lee, K. Smith, C. Csech, and G. Shaker, "Glucose Concentration Estimation using Electromagnetic Waves," *2018 18th International Symposium on Antenna Technology and Applied Electromagnetics*, 2018.

[33] S. Liu, G. Shaker, S. Safavi-Naeini, and J. M. Chong, "Low-cost Gas Sensors utilizing mm-Wave Radars," *2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, 2017.

[34] A. M. Hegazy, "Remote Material Characterization using mmWave FMCW Radar with Complex Baseband," 2020.

[35] J. Chen, X. Wang, Y. Xu, Y. Peng, W. Wang, J. Xiang, and Q. Xu, "Convolution neural network and 77 GHz millimeter wave radar based intelligent liquid classification system," *Journal of Information and Intelligence*, vol. 1, pp. 352–363, November 2023.

[36] R. G. Geyer, "Dielectric Characterization and Reference Materials," 1990.

[37] D. J. Griffiths, *Introduction to Electrodynamics*. Pearson Education Inc., 4 ed., 2013.

[38] G. G. Raju, *Dielectrics in Electric Fields*. Boca Raton, RL, USA: Taylor & Francis Group, 2 ed., 2017.

[39] M. J. Kadhim, A. K. Abdullah, I. A. Al-Ajaj, and A. S. Khalil, "Dielectric Properties of Epoxy/Al2O3 Nanocomposites," *International Journal of Application or Innovation in Engineering & Management*, vol. 3, January 2014.

[40] J. D. Jackson, *Classical Electrodynamics*. John Wiley & Sons Inc., 3 ed., August 1998.

[41] H. Frohlich, *Theory of Dielectrics: Dielectric Constant and Dielectric Loss*. Clarendon Press, 2 ed., 1987.

[42] F. T. Ulaby and D. G. Long, *Microwave Radar and Radiometric Remote Sensing*. Ann Arbor, MI, USA: University of Michigan Press, 2014.

[43] D. C. Elton, *Understanding the dielectric properties of Water*. PhD thesis, Stony Brook University, Stony Brook, NY, USA, December 2016.

[44] C. Meng, "REFLECTION AND DIELECTRIC MEASUREMENT FOR SALINITY OF WATER USING MICROSTRIP LOOP ANTENNA AND DIELECTRIC PROBE," *International Journal of GEOMATE*, 2016.

[45] D. H. Gadani, V. A. Rana, S. P. Bhatnagar, A. N. Prajapati, and A. D. Vyas, "Effect of salinity on the dielectric properties of water," *Indian Journal of Pure & Applied Physics*, vol. 50, pp. 405–410, June 2012.

[46] K. Fuchs and U. Kaatza, "Molecular Dynamics of Carbohydrate Aqueous Solutions. Dielectric Relaxation as a Function of Glucose and Fructose Concentration," *Journal of Physical Chemistry B*, vol. 105, pp. 2036–2042, 2012.

[47] V. S. Bagotsky, *Fundamentals of Electrochemistry*. Hoboken, NJ, USA: John Wiley & Sons Inc., 2 ed., 2006.

[48] J. E. Anderson, "The Debye-Falkenhagen effect: experimental fact ory fiction?," *Journal of Non-Crystalline Solids*, vol. 172–174, pp. 1190–1194, 1994.

[49] A. Chandra and B. Bagchi, "Frequency dependence of ionic conductivity of electrolyte solutions," *Chinese Journal of Chemical Physics*, vol. 112, 2000.

[50] A. E. Omer, S. Gigoyan, G. Shaker, and S. Safavi-Naeini, "WGM-Based Sensing of Characterized Glucose-Aqueous Solutions at mm-Waves," *IEEE Access*, vol. 8, pp. 38809–38825, 2020.

[51] C. Meng, M. Fareq, S. A. B., M. Afendi, Z. Z., S. F. Khor, L. Z., N. F. M. N., W. H. Tan, N. S. M. Noorpi, N. M. Mukhtar, and M. Othman, "Development of Low Cost Microwave Detection System for Salinity and Sugar Detection," *International Journal of Mechanical & Mechatronics Engineering*, vol. 14, no. 5, 2014.

[52] E. J. Murphy and S. O. Morgan, "The Dielectric Properties of Insulating Materials, III Alternating and Direct Current Conductivity," *Bell System Technical Journal*, vol. 18, pp. 502–537, 1939.

[53] A. Stogryn, "Equations for Calculating the Dielectric Constant of Saline Water (Correspondence)," *IEEE Transactions on Microwave Theory and Techniques*, vol. 19, pp. 733–736, August 1971.

[54] L. Klein and C. Swift, "An improved model for the dielectric constant of sea water at microwave frequencies," *IEEE Transactions on Antennas and Propagation*, vol. 25, pp. 104–111, January 1977.

[55] X. Y. et al., "The Physical Model of Dielectric Function and Radiofrequency for NaCl Solution and Application of the Seawater," November 2004.

[56] Z. Zhu and W. Guo, "Recent developments on rapid detection of main constituents in milk: a review," *Critical Reviews in Food Science and Nutrition*, vol. 61, no. 2, pp. 312–324, 2021.

[57] X. Zhu, W. Guo, and Z. Liang, "Determination of the Fat Content in Cow's Milk Based on Dieletric Properties," *Food Bioprocess Technology*, vol. 8, pp. 1485–1494, 2015.

[58] T. Kudra, V. Raghaven, C. Akyel, R. Bosisio, and F. van de Voort, "Electromagnetic Properties of Milk and its Constituents at 2.45 GHz," *Journal of Microwave Power and Electromagnetic Energy*, vol. 27, no. 4, pp. 199–204, 1992.

[59] H. D. Goff, A. Hill, and M. A. Ferrer, "Dairy Science and Technology eBook."

[60] D. Agranovich, I. Renhart, P. B. Ishai, G. Katz, D. Bezman, and Y. D. Feldman, "A microwave sensor for the characterization of bovine milk," *Food Control*, December 2015.

[61] J. F. O'Connor and E. C. Synnott, "Seasonal Variation in Dielectric Properties of Butter at 15 MHz and 4°C," *Irish Journal of Food Science and Technology*, vol. 6, no. 1, pp. 49–59, 1982.

[62] M. R. Wright, *An Introduction to Aqueous Electrolyte Solutions*. John Wiley & Sons Ltd, 2007.

[63] V. Orsat and G. V. Raghaven, *Emerging Technologies for Food Processing*, ch. Chapter 21 - Radio-Frequency Processing, pp. 285–398. Academic Press, 2 ed., 2014.

[64] H. Lizhi, K. Toyoda, and I. Ihara, "Dielectric properties of edible oils and fatty acids as a function of frequency, temperature, moisture, and composition," *Journal of Food Engineering*, vol. 88, pp. 151–158, 2008.

[65] I. Gouw and I. Vlugter, "Physical Properties of Triglycerides IV. Dielectric Constant," *European Journal of Lipid Science and Technology*, vol. 69, no. 4, pp. 223–226, 1967.

[66] A. C. Nunes, X. Bohigas, and J. Tejada, "Dielectric study of milk for frequencies between 1 and 20 GHz," *Journal of Food Engineering*, vol. 76, pp. 250–255, 2006.

[67] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, Inc., 4 ed., 2012.

[68] A. P. Gregory and R. N. Clarke, "Dielectric Metrology with Coaxial Sensors," *Measurement Science and Technology*, March 2007.

[69] R. Vijay, R. Jain, and K. S. Sharma, "Dielectric Properties of Water in Microwave Frequencies," *International Journal of Engineering Research & Technology*, vol. 3, no. 3, 2015.

[70] K. C. Yaw, *Measurement of Dielectric Material Properties*. Rohde & Schwarz, April 2012. Application Note.

[71] P. B. Ishai, M. S. Talary, A. Caduff, E. Levy, and Y. Feldman, "Electrode polarization in dielectric measurements: a review," *Measurement Science and Technology*, vol. 24, 2013.

[72] SPEAG, "Dielectric Assessment Kit V2.4 Professional Handbook," September 2018.

[73] "De-Embedding and Embedding S-Parameter Networks Using a Vector Network Analyzer." [Online]. Available: https://www.keysight.com/ca/en/assets/7018-06806/application-notes/5980-2784.pdf.

[74] D. Rytting, "Network analyzer error models and calibration methods." [Online]. Available: https://www.rfmentor.com/sites/default/files/NA_Error_Models_and_Cal_Methods.pdf.

[75] M. Complesso, "Linear Regression with Ordinary Least Square (OLS) in Machine Learning with scikit-learn." [Online]. Available: https://www.meccanismocomplesso.org/en/linear-regression-with-ordinary-least-square-ols-in-machine-learning-with-scikit-learn/.

[76] M. learning nuggets, "Linear regression in python with scikit-learn (with examples, code, and notebook)." [Online]. Available: https://www.machinelearningnuggets.com/python-linear-regression/.

[77] radartutorial.eu, "Frequency-Modulated Continuous-Wave Radar (FMCW Radar)." [Online]. Available: https://www.radartutorial.eu/02.basics/Frequency

[78] S. Abu-Sardanah, "Cardio-respiratory Health Monitoring Using a Wearable Radar System," 2023.

[79] M. Alizadeh, "Remote vital signs monitoring using a mm-wave FMCW radar," 2019.

[80] S. Azam, *Microwave Power Devices and Amplifiers for Radars and Communication Systems*. PhD thesis, Linköping University, 2009.

[81] X. Chen, S. Boumaiza, and L. Wei, "Self-Heating and Equivalent Channel Temperature in Short Gate Length GaN HEMTs," *IEEE Transactions on Electron Devices*, vol. 66, September 2019.

[82] K. R. Bagnall, "Device-level Thermal Analysis of GaN-based Electronics," June 2013.

[83] X. Chen, S. Boumaiza, and L. Wei, "Modeling Bias Dependence of Self-Heating in GaN HEMTs Using Two Heat Sources," *IEEE Transactions on Electrons Devices*, vol. 67, August 2020.

[84] S. Boumaiza, "ECE 671 Microwave and RF Engineering." [PowerPoint slides].

[85] S. Rao, "Introduction to mmwave Sensing: FMCW Radars." [Online]. Available: https://www.ti.com/content/dam/videos/external-videos/2/3816841626001/5415528961001.mp4/subassets/mmwaveSensing-FMCW-offlineviewing_0.pdf.

[86] M. M. Vázquez, "Basics of FMCW Radar." [Online]. Available: https://www.renesas.com/us/en/blogs/basics-fmcw-radar.

[87] C. A. Balanis, *Antenna Theory Analysis and Design*. John Wiley & Sons Inc., 3 ed., 2005.

[88] MathWorks, "Radar equation." [Online]. Available: https://www.mathworks.com/help/radar/ug/radar-equation.html.

[89] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar," *ACM Transactions on Graphics*, vol. 35, July 2016.

[90] C. A. Balanis, *Advanced Engineering Electromagnetics*. John Wiley & Sons Inc., 2 ed., 2012.

[91] S. Simula, S. Ikäläinen, and K. Niskanen, "Measurement of the Dielectric Properties of Paper," *Journal of Imaging Science and Technology*, vol. 43, pp. 472–477, September/October 1999.

[92] A. E. Omer, *Non-Invasive Blood Glucose Monitoring Using Electromagnetic Sensors*. PhD thesis, University of Waterloo, 2022.

[93] Infineon, "Project Soli information document." [Online]. Available: https://www.infineon.com/dgdl/Infineon-Google+Soli+FAQ+Document-FAQ-v03_00-EN.pdf?fileId=5546d4625d5945ed015d9845149e04c4.

[94] K. Smith, "The Tiny Radar Chip Revolutionizing Gesture Recognition: Google's ATAP's Project Soli." [Online]. Available: https://www.allaboutcircuits.com/news/radar-chip-revolutionizing-gesture-recognition-google-atap-project-soli/.

[95] D. Edgar, "HFSS 13: Hybrid FE-BI for Efficient Simulation of Radiation and Scattering," 2011. [Online].

[96] "Lecture 8: Introduction to Hybrid FEM IE," 2015. [Online].

[97] "Lecture 3-1: HFSS 3D Design Setup," 2016. [Online].

[98] Y. Ahmed and J. Reed, "A Primer on Ray-Tracing: Shooting and Bouncing Ray Method," *TechRxiv*, December 2023.

[99] Altair, "Shooting and Bouncing Rays (SBR)." [Online]. Available: https://help.altair.com/winprop/topics/winprop/user_guide/propagation_methods/propgation_models_indoor_sbr.htm.

[100] "Support Vector Machines with Python and Scikit-Learn." [Online]. Available: https://github.com/pb111/Support-Vector-Machines-Project.

[101] scikit learn, "Support Vector Machines." [Online]. Available: https://scikit-learn.org/stable/modules/svm.html.

[102] J. Charest, "Implementing a Support Vector Machine using Sequential Minimal Optimization and Python 3.5." [Online]. Available: https://jonchar.net/notebooks/SVM/.

[103] E. Y. Boateng, J. Otoo, and D. A. Abaye, "Basic Tenets of Classification Algorithms K-Nearest-Neighbor, Support Vector Machine, Random Forest and Neural Network: A Review," *Journal of Data Analysis and Information Processing*, vol. 8, November 2020.

[104] scikit learn, "1.11. Ensembles: Gradient boosting, random forests, bagging, voting, stacking." [Online]. Available: https://scikit-learn.org/stable/modules/ensemble.htmlrandom-forests-and-other-randomized-tree-ensembles.

[105] M. in Data Science with edX, "What Is a Decision Tree?." [Online]. Available: https://www.mastersindatascience.org/learning/machine-learning-algorithms/decision-tree/.

[106] A. S. Bakshi and D. E. Smith, "Effect of Fat Content and Temperature on Viscosity in Relation to Pumping Requirements of Fluid Milk Products," *Journal of Dairy Science*, vol. 67, pp. 1157–1160, 1984.

[107] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics Volume 8: Electrodynamics of Continuous Media*. Pergamon Press, 1960.

[108] R. E. Mudgett, D. I. C. Wang, and S. A. Goldblith, "Prediction of Dielectric Properties in Oil-Water and Alcohol-Water Mixtures at 3,000 MHz, 25°C Based on Pure Component Properties," *Journal of Food Science*, vol. 39, pp. 632–635, 1974.

[109] R. Simpkin, "Derivation of Lichtenecker's Logarithmic Mixture Formula From Maxwell's Equations," *IEEE Transactions on Microwave Theory and Techniques*, vol. 58, pp. 545–550, March 2010.

[110] U. Erle, M. Regier, H. Persch, and H. Schubert, "Dielectric Properties of Emulsions and Suspensions: Mixture Equations and Measurement Comparisons," *Journal of Microwave Power and Electromagnetic Energy*, vol. 35, no. 3, pp. 185–190, 2000.

[111] B. Panikuttira, F. A. Payne, N. O'Shea, J. T. Tobin, D. J. O'Callaghan, and C. P. O'Donnell, "Investigation of an in-line prototype fluorescence and infrared backscatter sensor to monitor rennet-induced coagulation of milk milk and at different protein concentrations," *International Journal of Food Science & Technology*, vol. 55, pp. 175–182, 2020.

[112] J. Kuzmik, P. Javorka, A. Alam, M. Marso, M. Heuken, and P. Kordos, "Investigation of self-heating effects in AlGaN/GaN HEMTs," *International Symposium on Electron Devices for Microwave and Optoelectronic Applications*, 2001.

[113] Infineon, *Radar Baseboard MCU7 XENSIV 60 GHz radar system platform*. Infineon, February 2023. [Online]. Available:

https://www.infineon.com/dgdl/Infineon-AN599_Radar_Baseboard_MCU7-ApplicationNotesv02_70EN.pdf?fileId=5546d4627550f45401755594a51a4d27.

[114] Teledyne FLIR, "FLIR ONE Edge Pro." [Online]. Available: https://www.flir.ca/products/flir-one-edge-pro/.

[115] Synopsys, *Synopsys Sentaurus Device User Guide*. Synopsys, June 2015.

APPENDICES

Appendix A

Modelling Behaviour using Mixture Equations

The dielectric properties of a material are purely experimental and modelling them is a complicated topic of study. There are several different mixing models based on different assumptions that can be used to predict the dielectric properties of a material. Determining the model that fits best would be an important part in predicting the dielectric properties of milk and validating experimental behaviour because they can provide a basis for analysis. For the purposes of this project, the Fricke-Mudgett equation, the Landau and Lifshitz equation, and the Lichtenecker and Rother equation will be briefly examined and then applied in the methodology section to validate experimental data collected. These three equations and many others were used in [110] to predict the dielectric properties of water-in-oil and oil-in-water emulsions at 2.45 GHz. It was found that out of the models used, the Fricke-Mudgett equation was the best for describing the oil-in-water emulsions while the Lichtenecker and Rother equation was the best for describing the water-in-oil emulsions.

Fricke-Mudgett Equation

The Fricke-Mudgett equation is given by

$$\epsilon_m = \epsilon_c \frac{\epsilon_d(1 + a\phi_d) + \epsilon_c a(1 - \phi_d)}{\epsilon_c(a + \phi_d) + \epsilon_d(1 - \phi_d)} \tag{A.1}$$

where ϵ_d is the dielectric constant or loss of the disperse phase, ϵ_c is the dielectric constant or loss of the continuous phase, ϕ_d is the volume fraction of the dispersed particles,

and a is the shape coefficient of the dispersed particles [110], [108]. The shape equation takes on one of two values, either $a = 1$ for long cylinders or $a = 2$ for suspended spheres [110]. Fricke derived this equation based on the complex conductivity equations derived by Maxwell for solid spheres and Rayleigh for long cylinders [108]. As mentioned, the Fricke-Mudgett equation best predicted the dielectric properties of oil-in-water emulsions in [110] due to the condition that the dielectric constant of the continuous phase must be greater than that of the dispersed phase. Furthermore, this equation assumes that the suspended particles are relatively inert and not affected by electrochemical interactions in the solution [108], which gives rise to the alternative name noninteractive Fricke model. There is an interactive Fricke model which combines the Debye equations that can be used to describe alcohol-in-water mixtures mentioned in [108].

Landau and Lifshitz Equation

The Landau and Lifshitz Equation is given by

$$\epsilon_m = \epsilon_1 + 3\frac{\epsilon_1(\epsilon_2 - \epsilon_1)\phi_2}{2\epsilon_1 + \epsilon_2} \tag{A.2}$$

where ϵ_1 is the dielectric constant or loss of the medium, ϵ_2 is the dielectric constant or loss of the dispersed phase, and ϕ_2 is the volume fraction of the dispersed phase [110], [107]. The model assumes a finely dispersed solution with spherical suspended particles of low concentration, such that the inhomogeneity of the overall solution can be neglected. Additionally, the model is best used when the dielectric constant of the medium and the dielectric constant of the dispersed phase are similar. Due to these assumptions, the applications of the Landau and Lifshitz model are limited in the food industry [110].

Lichtenecker and Rother Equation

The Lichtenecker and Rother equation is given by

$$\ln \epsilon_m = \phi_2 \ln \epsilon_2 + (1 - \phi_2) \ln \epsilon_1 \tag{A.3}$$

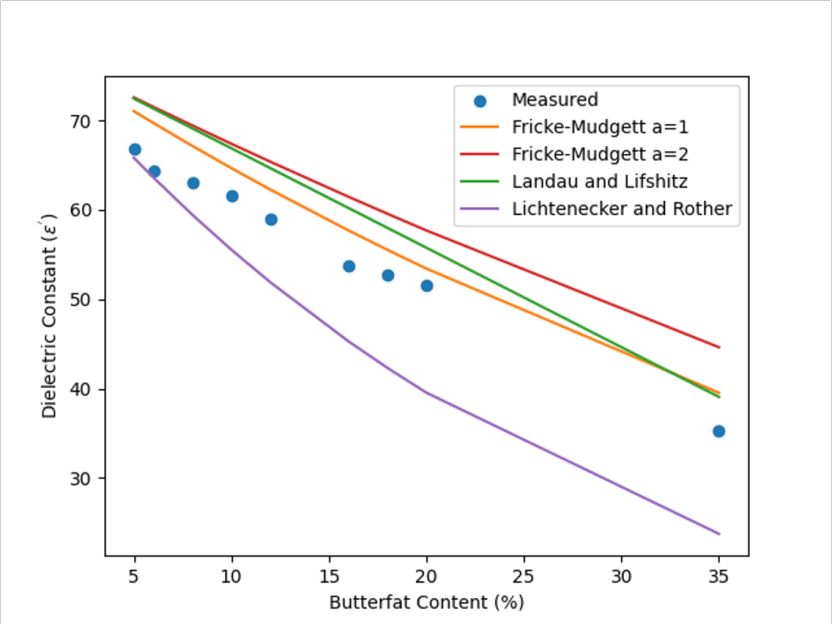
where the variables ϵ_m, ϵ_1, ϵ_2, and ϕ_2 are the same as the variables from the Landau and Lifshitz equation [110]. The equation requires that the relative permittivity is continuous, monotonic, and differentiable for ϵ_1, ϵ_2, and ϕ_2 [110]. Because the derivation of Lichtenecker-type mixing equations has a more mathematical approach than a physical one, there have been some criticisms of it for lacking rigour [109].

A summary of the 3 equations have been compiled in Table A.1 based on the above descriptions from [110], [108]-[107].

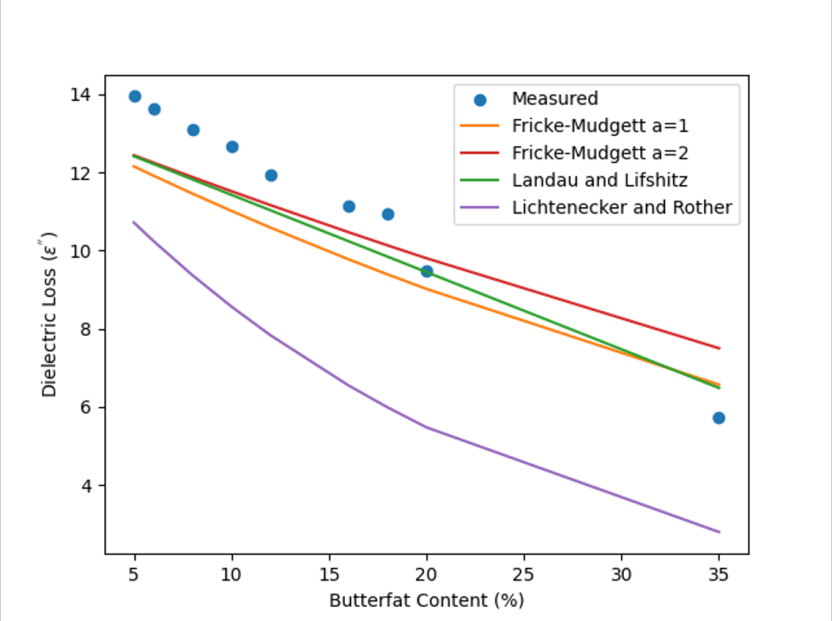
Table A.1: Summary of mixture equations

Model Name	**Equation**	**Conditions**
Fricke Mudgett / Noninteractive Fricke	$\epsilon_m = \epsilon_c \frac{\epsilon_d(1+a\phi_d)+\epsilon_c a(1-\phi_d)}{\epsilon_c(a+\phi_d)+\epsilon_d(1-\phi_d)}$	$a = 1$ (long cyclinders) $a = 2$ (spheres) $\epsilon_c \gg \epsilon_d$
Landau and Lifshitz	$\epsilon_m = \epsilon_1 + 3\frac{\epsilon_1(\epsilon_2-\epsilon_1)\phi_2}{2\epsilon_1+\epsilon_2}$	ϕ_2 low ϵ_1, ϵ_2 similar in value
Lichtenecker and Rother	$\ln \epsilon_m = \phi_2 \ln \epsilon_2 + (1-\phi_2) \ln \epsilon_1$	ϵ_m continuous, monotonic, and differentiable for ϵ_1, ϵ_2, ϕ_2

Using the Fricke-Mudgett, Landau-Lifshitz, and Lichtenecker-Rother mixing equations described above, an attempt to model milk and cream at a constant temperature and frequency was made. The dielectric constant and loss vs butterfat content are shown in Figure A.1 at 2450 MHz. This frequency was chosen because [58] measured the dielectric properties of pure fat and distilled water at 2450 MHz. These values were used when computing the values for the predictive mixing models. Erle et al. [110] also used 2450 MHz for their comparison between the mixing models and their measured values. From [58], the measured dielectric constant of fat at 2450 MHz is 2.613 and the measured dielectric loss is 0.153. Meanwhile, the measured dielectric constant and loss of water at 2450 MHz are 78.00 and 13.4 respectively. The value of the volume fraction of the dispersed phase was taken to be the concentration of butterfat in a sample as printed on its carton.



(a) Dielectric constants using mixing models.



(b) Dielectric losses using mixing models.

Figure A.1: Modelling the dielectric constant/loss vs butterfat content at 2450 MHz.

From the plotted measurements in Figure A.1, the Fricke-Mudgett $a = 1$ model appears to best fit the measured data for both the dielectric constant and dielectric loss. The values are not exact, but this is to be expected because milk and cream have other constituents that would lower the dielectric constant. As for the dielectric loss, all the models begin to converge to a similar value for 0% butterfat, but the measured values of the dielectric loss are higher. This could be due to the treatment of milk as a two-component solution of emulsified fat in water. At 0% fat, the solution is not truly 100% water, as there are other components such as proteins, salts, lactose, etc. that are also in the solution that would decrease the dielectric constant and increase the dielectric losses. At lower concentrations, the effect on the dielectric constant and losses from the other constituents of milk becomes more apparent, especially in the case of the dielectric losses.

For both plots, the Lichtenecker-Rother model is very different from the measured values. It was reported in [110], as previously mentioned, that the Lichtenecker-Rother model was best used for water-in-oil emulsions. Since milk is treated as an oil-in-water emulsion, it is expected to see a large deviation. In general, the Fricke-Mudgett models, especially $a = 1$, fit the measured data well considering that there were no fitted parameters used. This result is similar to the result obtained by Erle et al. [110] regarding oil-in-water emulsions. The reason why the Fricke-Mudgett $a = 1$ model fitted better than the Fricke-Mudgett $a = 2$ model was not explained in [110], but may be due to the fact triglycerides are made up of long chains of hydrocarbon molecules.

Appendix B

Python Code for Linear Regression Modelling

Code used in Section 3.4.5.

```
# import the data for a given frequency
f = 2450
findex = df.loc[df["fmhz"] == f]

# import the regular cream data for the frequency given above
cream5dc = findex["5_dc"].values
cream5dl = findex["5_dl"].values

cream10dc = findex["10_dc"].values
cream10dl = findex["10_dl"].values

cream18dc = findex["18_dc"].values
cream18dl = findex["18_dl"].values

cream35dc = findex["35_dc"].values
cream35dl = findex["35_dl"].values

# import the mixed cream data for the frequency given above
cream6dc = findex["6_dc"].values
cream6dl = findex["6_dl"].values
```

```
cream8dc = findex["8_dc"].values
cream8dl = findex["8_dl"].values

cream10dc = findex["10_dc"].values
cream10dl = findex["10_dl"].values

cream12dc = findex["12_dc"].values
cream12dl = findex["12_dl"].values

cream16dc = findex["16_dc"].values
cream16dl = findex["16_dl"].values

cream20dc = findex["20_dc"].values
cream20dl = findex["20_dl"].values

# prepare data sets to plot
fat = np.array([5.00, 10.00, 18.00, 35.00]).reshape(-1, 1)
dc = np.concatenate([cream5dc, cream10dc, cream18dc, cream35dc])
dl = np.concatenate([cream5dl, cream10dl, cream18dl, cream35dl])

# split the data randomly for training and testing
sample_size = 0.5
randomdc = 4
randomdl = 8
fat_train, fat_test, dc_train, dc_test = train_test_split(fat, dc,
    test_size=sample_size, random_state=randomdc)
fat_train2, fat_test2, dl_train, dl_test = train_test_split(fat, dl,
    test_size=sample_size, random_state=randomdl)

dcmodel = LinearRegression().fit(fat_train, dc_train)

dctrainpredict = []
dctestpredict = []
for i in range(0, len(dc_train)):
    dctrainpredict.append(dcmodel.coef_*fat_train[i] + dcmodel.intercept_)
for i in range(0, len(dc_test)):
    dctestpredict.append(dcmodel.coef_*fat_test[i] + dcmodel.intercept_)
```

```
print("Dielectric Constant Equation: ", dcmodel.coef_ , "x + ",
    dcmodel.intercept_)
print("R2 (Training Set): ", dcmodel.score(fat_train, dc_train))
print("RMSE (Training Set): ", np.sqrt(mean_squared_error(dc_train,
    dctrainpredict)))
print("R2 (Test Set): ", dcmodel.score(fat_test, dc_test))
print("RMSE (Test Set): ", np.sqrt(mean_squared_error(dc_test,
    dctestpredict)))

dlmodel = LinearRegression().fit(fat_train2, dl_train)

dltrainpredict = []
dltestpredict = []
for i in range(0, len(dl_train)):
    dltrainpredict.append(dlmodel.coef_*fat_train2[i] + dlmodel.intercept_)
for i in range(0, len(dl_test)):
    dltestpredict.append(dlmodel.coef_*fat_test2[i] + dlmodel.intercept_)

print("Dielectric Loss Equation: ", dlmodel.coef_ , "x + ",
    dlmodel.intercept_)
print("R2 (Training Set): ", dlmodel.score(fat_train2, dl_train))
print("RMSE (Training Set): ", np.sqrt(mean_squared_error(dl_train,
    dltrainpredict)))
print("R2 (Test Set): ", dlmodel.score(fat_test2, dl_test))
print("RMSE (Test Set): ", np.sqrt(mean_squared_error(dl_test,
    dltestpredict)))

# perform fat content prediction based on dc_test data and
    compare with fat_test
print("Fat Content Prediction using Test Data")
fatpredict = []
fatpredict2 = []
print("From Dielectric Constant")
for i in range(0, len(dc_test)):
    fatpredict.append((dc_test[i] - dcmodel.intercept_)/dcmodel.coef_)
    print("Measured Fat: ", fat_test[i], "Predicted Fat from DC: ",
        fatpredict[i])
```

```
    i += 1
print("From Dielectric Loss")
for i in range(0, len(dl_test)):
    fatpredict2.append((dl_test[i] - dlmodel.intercept_) / dlmodel.coef_)
    print("Measured Fat: ", fat_test2[i], "Predicted Fat from DC: ",
        fatpredict2[i])
    i += 1

# predict fat based on the new set of collected data
print("Fat Content Prediction using Mixed Cream Data")
fatmix = np.array([6.00, 8.00, 12.00, 16.00, 20.00])
mixdc = np.concatenate([cream6dc, cream8dc, cream12dc, cream16dc,
    cream20dc])
mixdl = np.concatenate([cream6dl, cream8dl, cream12dl, cream16dl,
    cream20dl])

fatmixpredict =[]
fatmixpredict2 = []
for i in range(0, len(mixdc)):
    fatmixpredict.append((mixdc[i] - dcmodel.intercept_)/dcmodel.coef_)
    fatmixpredict2.append((mixdl[i] - dlmodel.intercept_) / dlmodel.coef_)
    print("Measured Fat: ", fatmix[i], "Predicted Fat from DC: ",
        fatmixpredict[i], "Predicted Fat from DL: ",
        fatmixpredict2[i])
    i += 1

# plot the models, training sets, and test sets
fat_ = np.linspace(0, 35, 100)

plot1 = plt.figure(1)
plt.scatter(fat_train, dc_train, label="Training Set")
plt.scatter(fat_test, dc_test, label="Test Set")
plt.scatter(fatmix, mixdc, label="New Data")
plt.plot(fat_, dcmodel.coef_*fat_ + dcmodel.intercept_, label="Model")

handles, labels = plt.gca().get_legend_handles_labels()
order = [1, 2, 0, 3]
plt.legend([handles[i] for i in order], [labels[i] for i in order])
```

```
plt.xlabel("Butterfat Content (%)")
plt.ylabel("Dielectric Constant ($\epsilon^{'}$)")

plot2 = plt.figure(2)
plt.scatter(fat_train2, dl_train, label="Training Set")
plt.scatter(fat_test2, dl_test, label="Test Set")
plt.scatter(fatmix, mixdl, label="New Data")
plt.plot(fat_, dlmodel.coef_*fat_ + dlmodel.intercept_, label="Model")

handles, labels = plt.gca().get_legend_handles_labels()
order = [1, 2, 0, 3]
plt.legend([handles[i] for i in order], [labels[i] for i in order])

plt.xlabel("Butterfat Content (%)")
plt.ylabel("Dielectric Loss ($\epsilon^{''}$)")

plt.show()
```

Appendix C

Python Code for SVM Classification

Code used in Section 4.3.

```
# prepare the datasets to train SVM algorithm
milk = [skim1, skim2, skim4, skim4, skim5, skim6, skim7, skim8, skim9,
        milk1p1, milk1p2, milk1p3, milk1p4, milk1p5, milk1p6, milk1p7,
            milk1p8, milk1p9,
        milk2p1, milk2p2, milk2p3, milk2p4, milk2p5, milk2p6, milk2p7,
            milk2p8, milk2p9,
        homo1, homo2, homo3, homo4, homo5, homo6, homo7, homo8, homo9]
fat = ["0.10%", "0.10%", "0.10%", "0.10%", "0.10%", "0.10%", "0.10%",
            "0.10%", "0.10%",
       "1%", "1%", "1%", "1%", "1%", "1%", "1%", "1%", "1%",
       "2%", "2%", "2%", "2%", "2%", "2%", "2%", "2%", "2%",
       "3.25%", "3.25%", "3.25%", "3.25%", "3.25%", "3.25%", "3.25%",
            "3.25%", "3.25%"]

# split the data randomly for training and testing
sample_size = 0.2
random = 0
milk_train, milk_test, fat_train, fat_test = train_test_split(milk, fat,
    test_size=sample_size, random_state=random)

# train the SVM algorithm
clf = SVC(kernel='linear')
```

```
clf = clf.fit(milk_train, fat_train)
clf_predict = clf.predict(milk_test)
print("SVM Prediction: ", clf_predict)
print("Actual: ", fat_test)

accuracy = clf.score(milk_test, fat_test)
print("Accuracy: ", accuracy)

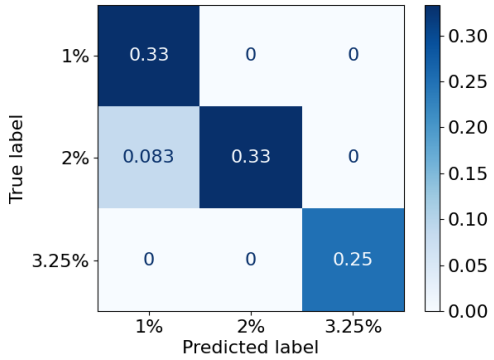
cm = confusion_matrix(fat_test, clf_predict)
disp = ConfusionMatrixDisplay(confusion_matrix=cm,
    display_labels=clf.classes_)
disp.plot()

plt.show()
```

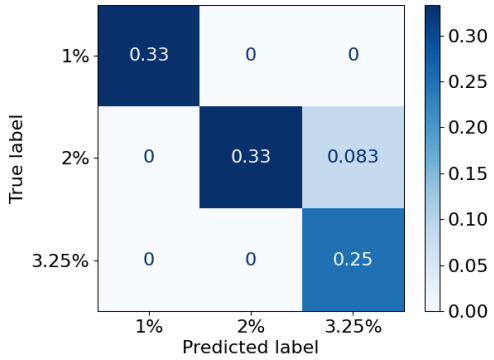
Appendix D

Machine Learning Applied to Glass Jar Experiment

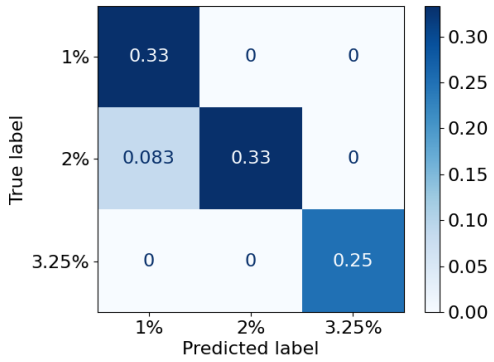
Support vector classifier, random forest classifier, and k-nearest neighbour algorithms are applied to the measurement data from the glass jar experiment in Section 4.4.1. Each algorithm yielded 91.67% accuracy.



(a) Support vector classifier, accuracy = 91.67%.



(b) Random forest classifier, accuracy = 91.67%.



(c) K-nearest neighbours, accuracy = 91.67%.

Figure D.1: Confusion matrices for stationary glass jar data.

Appendix E

HFSS SBR+ Packaging Study

A brief packaging study was conducted to see the effects of the packaging on the scattering response picked up by the simple 1 transmitter and 1 receiver system from Section 4.2.2. The same simulation setup was used, but with 2 mm thick carton packaging. The two types of packaging selected are cardboard and glass. The reason SBR+ method is selected for this packaging study instead of the FEM/FEBI method is for computational speed and efficiency. For the reference, 7.5 cm by 7.5 cm by 19.5 cm block was created and assigned a cardboard material using a relative permittivity of 3.4 and permeability of 1 and dielectric loss tangent of 0. The relative permittivity was taken randomly between 3 to 4, which was the approximate dielectric constant of printing paper from [91]. The simulated magnitude of response of the cardboard block is shown in Figure E.1. The material was then changed to the default glass material provided in the HFSS material library which has a relative permittivity of 5.5, permeability of 1, and dielectric loss tangent of 0. This is shown in Figure E.2.

It should be noted that the relative permittivity values used for the packaging study is valid for low frequencies. For the most accurate values at 60 GHz, they would need to be characterized using the DAK-TL and imported into the material library. The liquids from Section 4.2.2 are simulated with the 2 mm thick cardboard and glass packaging. They are shown in Figure E.3 and Figure E.4 respectively.

From the simulated results, it can be seen that the reflections picked up by the radar are different from the reference blocks. This indicates that the liquid dielectric is contributing to the reflections as predicted from the multiple dielectric slab equation. However, in both cases, the different milk products and distilled water cannot be differentiated. This was thought to be due to the 1W of power that was used, and this was increased to 5W to

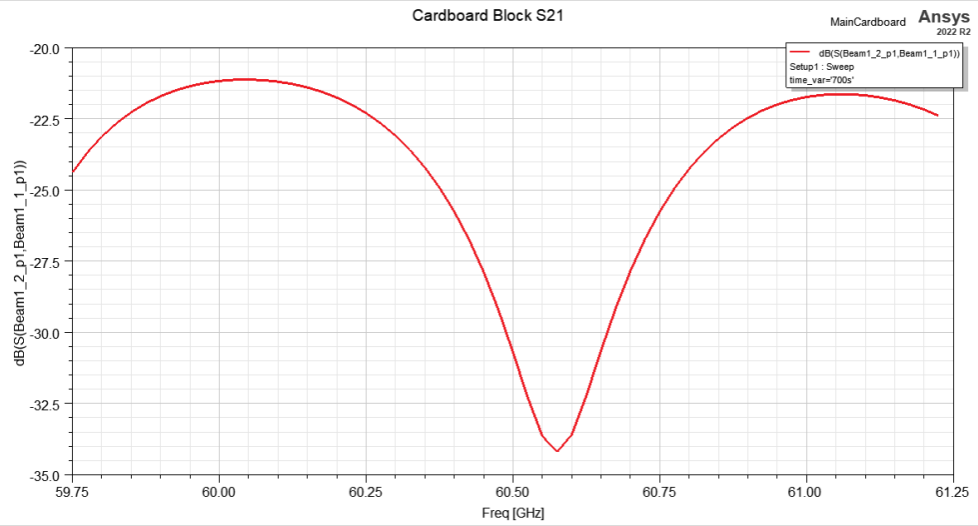


Figure E.1: Magnitude of reflection from cardboard block in HFSS SBR+.

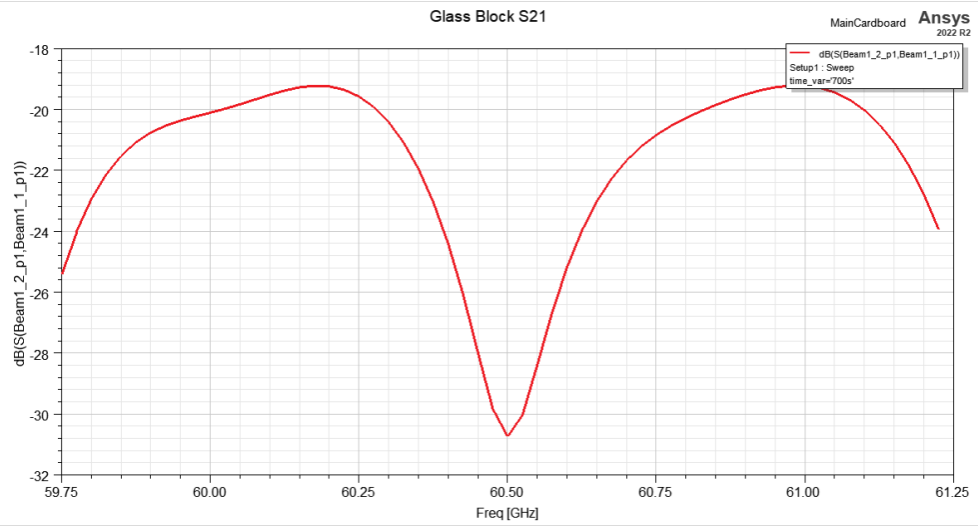


Figure E.2: Magnitude of reflection from glass block in HFSS SBR+.

determine whether or not differences would appear. It turned out to not make a difference in the simulated results. It may be concluded that the SBR+ method is not great for studying this sort of interaction. The FEBI method may work better, but it was more difficult to see trends in the simulated data compared to the SBR+ method, as described in Section 4.2, and would take some more time to simulate.

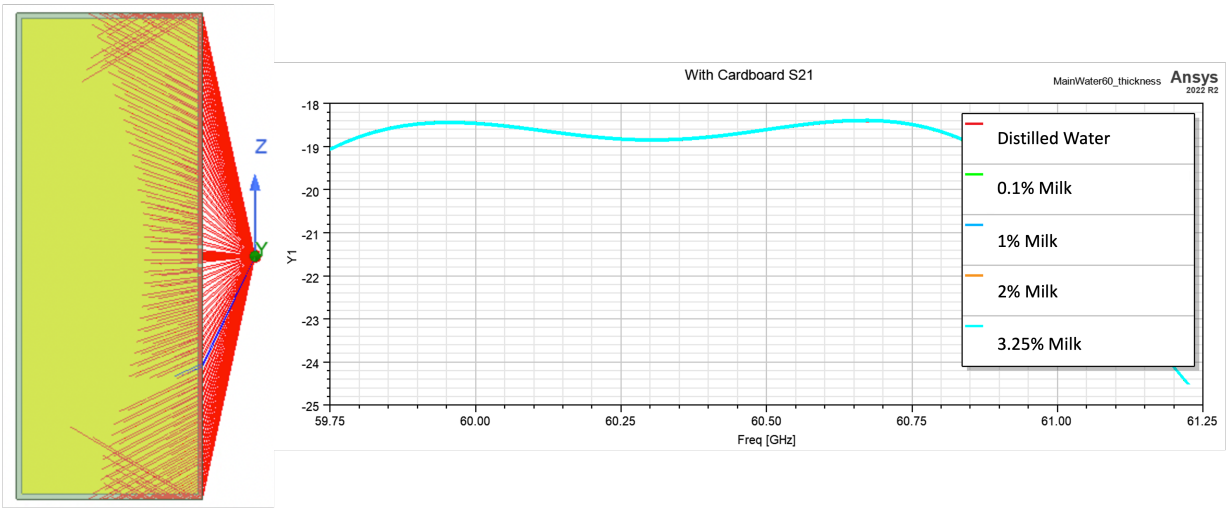


Figure E.3: Magnitude of reflection with 2 mm thick cardboard packaging in HFSS SBR+.

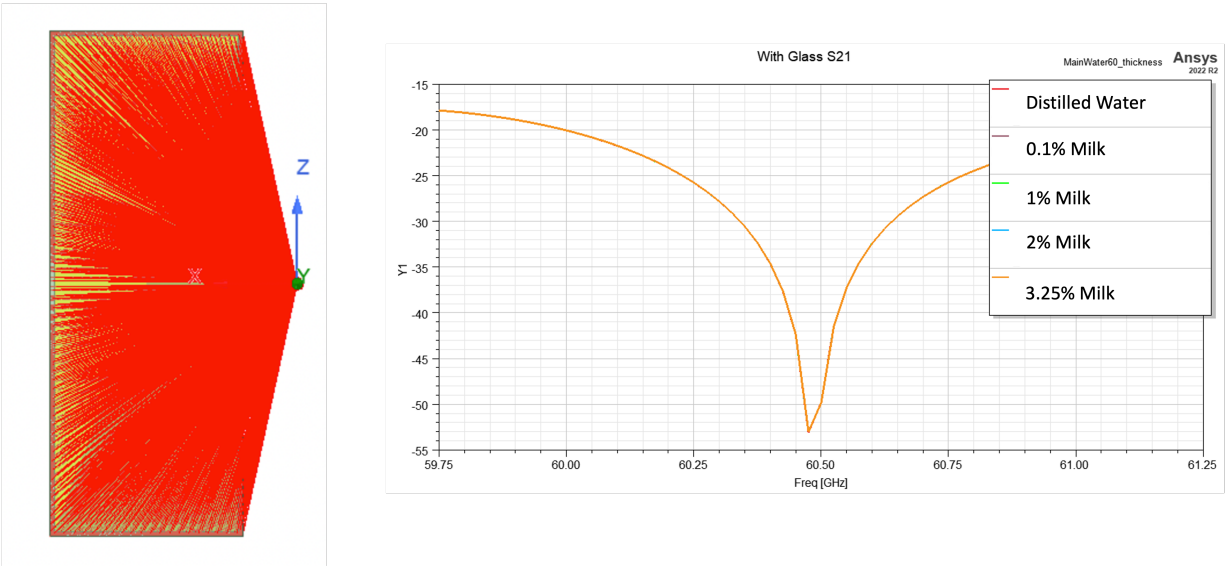


Figure E.4: Magnitude of reflection with 2 mm thick glass packaging in HFSS SBR+.

Appendix F

Infineon Radar Thermal Characterization

In this application of installing a radar on a conveyor belt for monitoring, or any other monitoring applications, it is possible that the signal degrades as the device heats up overtime. Thermally characterizing the radar while it is under normal operation is important. Some work has been done on thermally characterizing the Infineon radar throughout its continuous operation over the span of 2 days and the task was short lived due to communication issues with the radar. At the time of writing, this idea of thermal characterization has been short lived, and most of the work will be passed on as a future project. The Infineon radar was selected for this task because of its wider usage within the lab group as the data collected could be used for more applications.

A figure of the Infineon radar baseboard and antenna as it is plugged in and running after 2 minutes are shown in Figure F.1. The antenna and sensor board are a separate piece that mounts onto the radar baseboard. The baseboard contains a MCU that handles all the signal processing. As an active component, the MCU is expected to be the hottest region in a thermal map. The MCU operates as a small computer on an integrated chip. It contains a CPU, memory, and input/output peripherals. It is a 32-bit system with a clock frequency of 300 MHz and SRAM of 384 kilobytes [113]. The antenna on the other hand, is passive and does the heat generated in the antenna is mainly driven through from the rest of the radar. In the heat map of the radar when it is plugged in but not running, there the sensor board is yellowish instead of red. This is because the sensor supply is deactivated and only the MCU is running.

An image of the setup is displayed in Figure F.2. For this measurement campaign, two

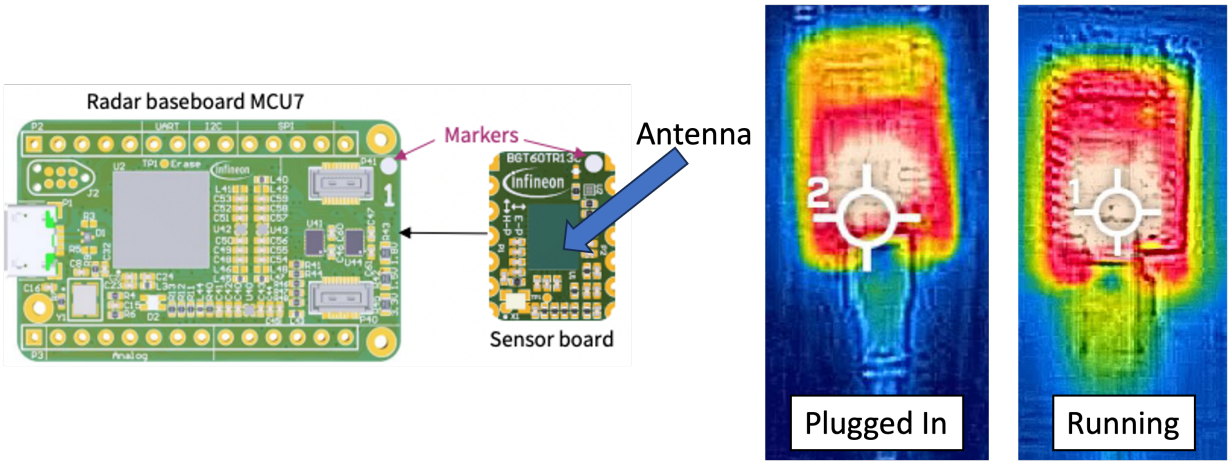


Figure F.1: 60 GHz radar system platform and thermal characterization. Adapted from [113].

radars were connected to a laptop running Windows 10 using micro-USB cables. One of the radars was positioned inside a 3D printed packaging and the other was exposed to air. All the thermal measurements were done using a FLIR ONE Edge Pro which has a thermal resolution of 160 pixels x 120 pixels (or 19,200 pixels in total) [114]. The thermal sensor's pixel size is 12 micrometres with a 90 mK thermal sensitivity [114]. A script allowing both radars to run at the same time was used to compare the effects of the packaging on the radar heating while in the same environment. Results of the radar in packaging after two minutes and 55 hours is shown in Figure F.3.



Figure F.2: Radar thermal characterization setup.

There wasn't a huge notable difference between the radar in packaging the radar without

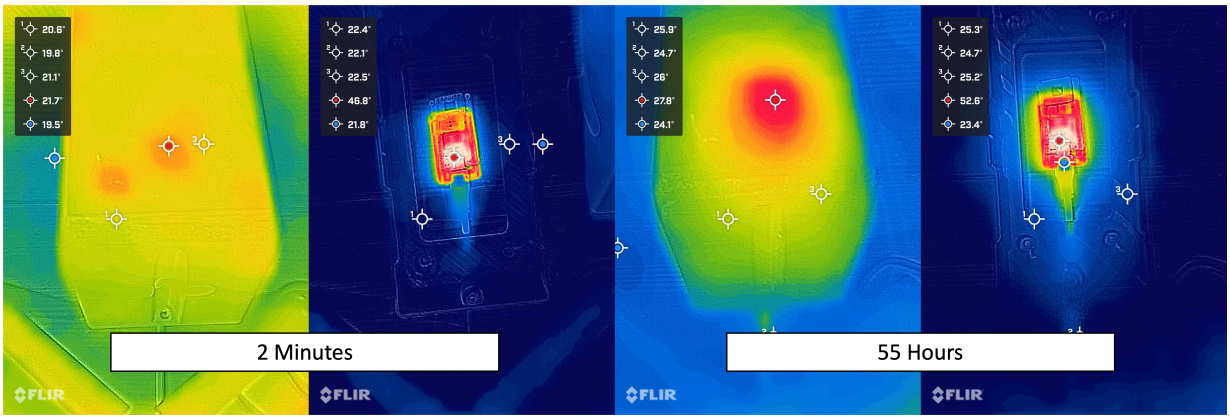


Figure F.3: Heat map of 60 GHz radar.

packaging. The temperature difference between the two at the end of the second day was about 1-2 degrees different. The radar in the packaging was slightly hotter than the one without, which is expected as the heat is trapped and contained inside the packaging. Examining just the radar in packaging from Figure F.3, the device heated up to a maximum temperature of 52.6 degrees Celsius from 46.8 degrees Celsius. At approximately an hour or less, the radar reached a steady-state and running it for a longer time did not change the reading. This is important to know how the the temperature of the radar changes over time as it is actively running, since its performance is temperature-dependent. For example, the gain of the PA in the radar would drop as the temperature increases. The performance difference between 2 minutes and 55 hours were not compared and will be left as a future study.

Appendix G

Gallium Nitride Power Amplifiers

G.1 Gallium Nitride High Electron Mobility Transistors

As previously mentioned, GaN HEMT technology is commonly used in high frequency PAs because of improved efficiency, output power, and power density. Unfortunately, self-heating becomes an issue due to their high-power density and high electron mobility. Combined with the heating of the radar, the signal degradation effects are non-negligible over long periods of time if there is no adequate heat dissipation. To study the self-heating effects of a GaN PA, a single finger can be first studied and extended to multifingers. All the simulations presented in this thesis for the electrothermal modelling of the GaN transistors are done in Sentaurus Technology Computer Aided Design (TCAD).

A structure of a 5 layer GaN HEMT is depicted in Figure G.1. The silicon nitride layer found in the source and drain access regions is used for as the passivation layer. The passivation layer shields the surface of the semiconductor from its surrounding environment. Below the nitride layer is the aluminum gallium nitride (AlGaN) layer doped with a specified donor concentration. Underneath the AlGaN layer is an unintentionally-doped (UID) AlGaN layer with a smaller specified donor concentration. The GaN region acts as a buffer and the silicon substrate acts as a thermal contact for the device.

A 2D electron gas (2DEG) layer exists between the GaN and AlGaN layers. This 2DEG layer is confined to a plane and limits the scattering of any electrons from the body or dopant to two dimensions. This makes the mobility of the electrons extremely high and gives these devices the name high electron mobility transistors. The 2DEG

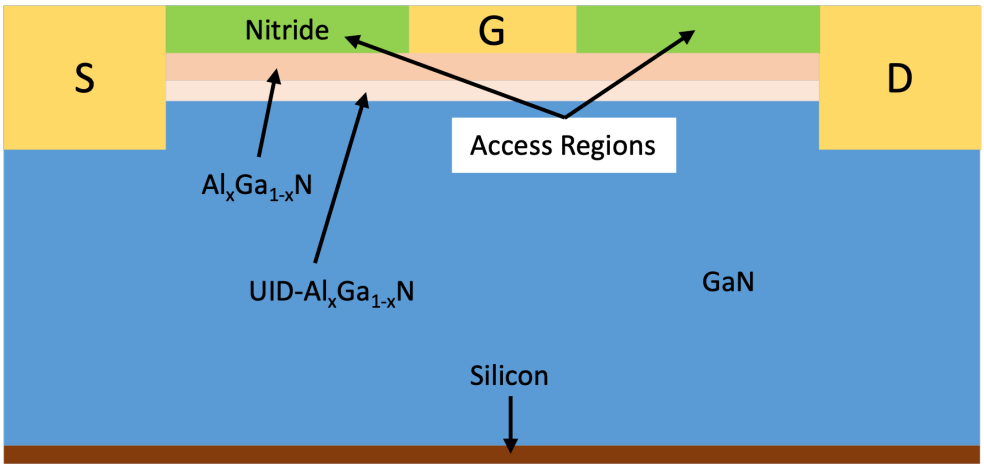


Figure G.1: Cross-section of a GaN HEMT. Adapted from [81].

emerges due to the different Fermi levels of the AlGaN and GaN layers and band bending between the two. The asymmetry of GaN creates a dipole and induces an electric field. Piezoelectric polarization may be observable if mechanical stress is applied along with the dipolar polarization already occuring. Piezoelectric polarization occurs when the lattice is stretched or strained, causing a molecule to carry a net postive charge and net negative charge on each side.

G.2 Sentaurus TCAD Electrothermal Simulation

Sentaurus is a software offered by Synopsys to simulate multiphysics effects in wide variety of semiconductor devices in 2D or 3D. It uses combines multiple tools into one seamless software suite where each tool has its own programming language, which adds to its steep learning curve. The foundation of the solver in Sentaurus is FEM, but it works differently from HFSS that has adaptive meshing. In Sentaurus, the meshing over each geometry or region must be specified as well as the physics. To properly understand the simulation, a strong background in semiconductor physics and mathematics is required. Some work has been done with Sentaurus in simulating GaN transistors at DC, but this largely remains a work in progress and can be a whole project on its own. Nevertheless, some concepts will be introduced to better understand the presented results. Scripts of the single finger device from Chen's work [81] has been used for the following single and multifinger simulations. For more details of the device structure and simulation setup, refer to [81].

The same structure shown in Figure G.1 is simulated in Sentaurus with a gate length of 0.1 μm. The device was symmetrical with both access region lengths 0.95 μm. The device was simulated with an ambient temperature of 300K. A gate-dependent piezoelectric polarization model is used. There are two models (strain and stress) that can be used to compute polarization effects in GaN devices. An anisotropic crystal lattice in the z-direction is assumed. A thermionic emission model was enabled for both electrons and holes. It is used to look at carrier transport at a heterojunction interface. Joule heating, recombination heating, Thomson heating, and Peltier heating effects are taken into consideration.

The simulated drain current vs drain voltage (IV) curve results of the cross section of the single finger GaN HEMT are shown in Figure G.2. The device was simulated at DC and gate current leakage was assumed to be negligible. Furthermore, the device 2D and current distribution assumed to be 1D to improve simulation time. In Figure G.2, the IV curves with gate voltage swept from -3 to 0 for 3 different transport/temperature models. The dashed lines represent the simulation using only the drift diffusion equation without any temperature coupling. These drift diffusion equations are defined as [115]

$$\vec{J_n} = \mu_n(n\nabla E_C - 1.5nkT\nabla \ln m_n) + D_n(\nabla n - n\nabla \ln \gamma_n) \tag{G.1}$$

$$\vec{J_p} = \mu_p(p\nabla E_V + 1.5pkT\nabla \ln m_p) - D_p(\nabla p - p\nabla \ln \gamma_p). \tag{G.2}$$

- $\vec{J_n}$ and $\vec{J_p}$ are the electron and hole current densities.
- n and p are the electron and hole densities.
- E_C is the conduction band energy.
- E_V is the valence band energy.
- m_n and m_p are the effective masses of the electrons and holes.
- μ_n and μ_p are the electron and hole mobilities.
- k is the Boltzmann constant.
- T is the lattice temperature.
- D_n and D_p are the electron and hole diffusivities.
- γ_n and γ_p are defined accordingly for Fermi statistics, more information in [115].

The drift diffusion current equations describe the summation of the drift and diffusion current processes occurring. The names are as they suggest, where the drift current arises due to a shifting of charge carriers while the diffusion current occurs as charge carriers distribute themselves over an area. Understanding the physics and equations presented in this section is beyond the scope of this project, and is a considerable portion of a much larger project. As the drift diffusion equation is coupled with the default lattice temperature equation in the Default Temperature model, the drain current shows degradation due to self-heating. The default lattice temperature model only simulates an equilibrium temperature and has the following form [115]

$$\frac{\partial W_L}{\partial t} + \nabla \cdot \vec{S_L} = \left.\frac{\partial W_L}{\partial t}\right|_{collision} + \frac{1}{q}\vec{J_n} \cdot \nabla E_C + \left.\frac{dW_n}{dt}\right|_{collision} + \frac{1}{q}\vec{J_p} \cdot \nabla E_V + \left.\frac{dW_p}{dt}\right|_{collision}. \tag{G.3}$$

- W_n, W_p, W_L are the energy densities of the electrons, holes, and lattice respectively.
- $\vec{S_L}$ energy flux of the lattice.

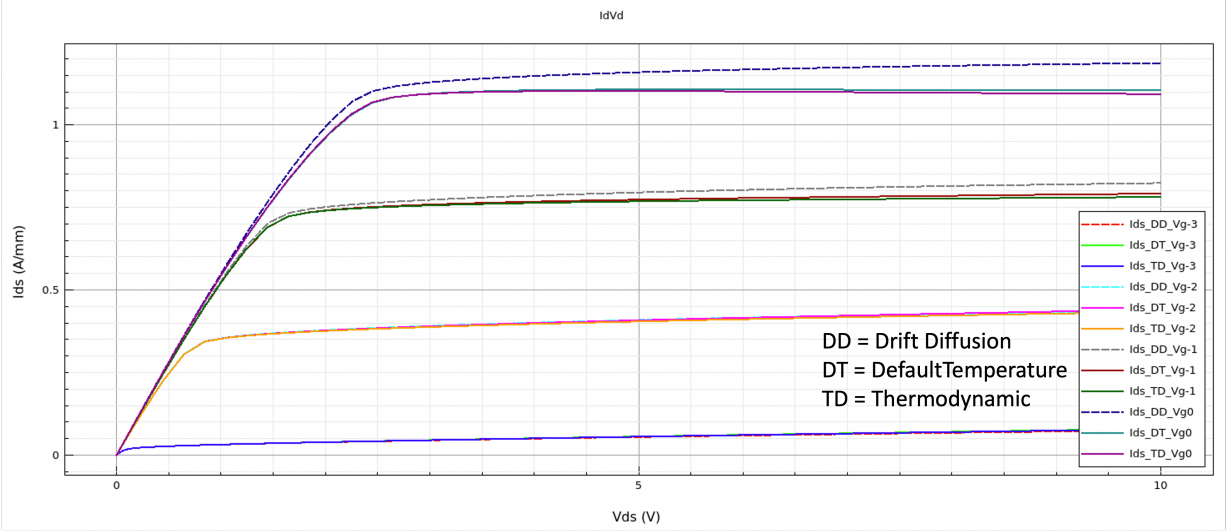


Figure G.2: IV curves computed using different temperature models.

The GaN HEMT has a parasitic source resistance R_S comprised of a contact resistance and channel resistance. Increasing the temperature of the device lowers the channel's electron mobility which in turn increases R_S [112]. Thus, with self-heating effects the

saturation current would be expected to decrease by a factor of $I_{sat} g_m R_S$ where I_{sat} is the saturation current and g_m is the transconductivity of the device [112]. In comparison to the drift diffusion model and the default temperature models, the thermodynamic model offers the most accurate heating. The model uses a different temperature equation and can simulate a lattice at equilibrium or non-equilibrium temperature, unlike the default temperature model. As one would expect, the equation is much more complicated

$$\frac{\partial}{\partial t}(c_L T) - \nabla \cdot (\kappa \nabla T) = -\nabla \cdot [(P_n T + \Phi_n)\vec{J_n} + (P_p T + \Phi_p)\vec{J_p}] \tag{G.4}$$

$$-\frac{1}{q}\Big(E_C + \frac{3}{2}kT\Big)(\nabla \cdot \vec{J_n} - qR_{net,n}) \tag{G.5}$$

$$-\frac{1}{q}\Big(-E_V + \frac{3}{2}kT\Big)(-\nabla \cdot \vec{J_p} - qR_{net,p}) + \hbar\omega G^{opt}. \tag{G.6}$$

- κ is the thermal conductivity.
- c_L is the lattice heat capacity.
- G^{opt} is the optical generation rate of photons.
- $R_{net,n}$ and $R_{net,p}$ are the recombination rates of the electrons and holes.
- P_n and P_p are the absolute thermoelectric powers.

From Figure G.2, it can be seen that there is greater drain current degradation with the thermodynamic model than the default temperature one. Sentaurus also offers the hydrodynamic model where the lattice temperature can be different from the carrier temperatures. This model offers the most realism and accuracy for a simulation. Unfortunately, it is extremely difficult to get to converge for GaN. A comparison of the cross-sectional heat map of the GaN HEMT using the default temperature model and the thermodynamic model is shown in Figure G.3. The heat map was generated for a drain voltage of 10V and a gate voltage of 0V. Noticeably, the gate edge on the drain side is hotter than the gate edge on the source side. This is due to the DC biasing on the drain contact and a voltage drop occurring across the gate contact. The source is assumed to be grounded. The heat dissipates into the surroundings, with the coolest regions being at the bottom of the device near the silicon substrate. Because of this large voltage drop, there will be a large lateral electric field along that edge contributing to a greater heat density and generating more Joule heat. Therefore, the edge has the hottest lattice temperature [81].

In the simulations presented, the device was assumed to be symmetrical such that the length between the source and gate is the same as the length between the gate and drain,

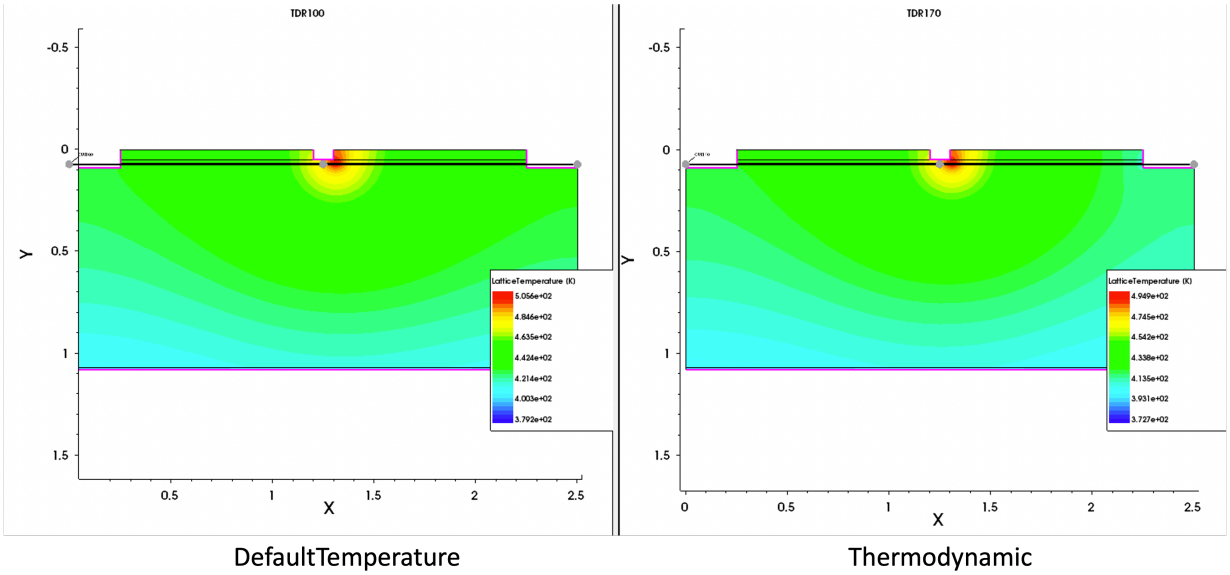


Figure G.3: Cross-Sectional heat map of GaN HEMT for $V_D = 10V$ and $V_G = 0V$.

$L_{sg} = L_{gd}$. Note these are the access regions of the device. Chen comments that increasing L_{sg} lowers the drain current and lowers the device temperature [81]. This is a possibility to reduce self-heating, but it comes with the trade-off that there will be decreased power output and slower device speed. One may ask why does the source access region affect the drain current more than the drain side. The saturation current equation is defined with the source side in mind:

$$I_{D,sat} = \frac{1}{2}\mu_p C_{ox}\frac{W}{L}(V_{GS} - V_T)^2 \tag{G.7}$$

where $V_{GS} = V_G - V_{si}$, V_{si} is the internal voltage drop across the source access region. This depends on the resistance of the region, which further depends on the length, $R = \frac{\rho L_{sar}}{A_{sar}}$. The source access resistance can be defined as

$$R_{sar} = \frac{R_{s0}}{W_{sar}}. \tag{G.8}$$

The current flow over the source access resistance is

$$I(si, s) = \frac{V(si) - V(s)}{R_{sar}}. \tag{G.9}$$

If the source is grounded, then $V(s) = 0$ and

$$V_{si} = I_{si} R_{sar}. \tag{G.10}$$

Equation G.10 can be substituted into the expression for V_{GS} and into Equation G.7 to yield

$$I_{D,sat} = \frac{1}{2} \mu_p C_{ox} \frac{W}{L} (V_G - I_{si} R_{sar} - V_T)^2. \tag{G.11}$$

From Equation G.11, it becomes clear that if the source access resistance increases then the drain current decreases. This can be kept in mind when analyzing the results of the device if the access regions were not symmetric.

G.3 Multifinger GaN HEMT Electrothermal Simulation

Most transistors exist as multifingered ones rather than single fingered. The single finger simulation offers an approximation of the heating that occurs and much shorter computation time, but running a multifinger simulation for higher accuracy is desirable. The layout of the multifinger transistor is important, and whether it is cloned from the drain side or the source side makes a difference in the results. A 4-finger transistor was set up in Sentaurus by cloning the single finger transistor on the drain contact. The IV curves using the three different transport/temperature models are plotted in Figure G.4. Notice that the total current is 4 times the current for a single finger because they are in parallel.

The cross-sectional heat map of the 4-finger transistor is shown in Figure G.5 with a cutline plotted at y=0.0725 for the lattice temperature distribution in 1D. The centred peaks are smaller than the outer peaks because the source is in the centre. The heat dissipates at the middle source because there is 0 V and is cooler. If the drain was in the centre instead, the heat would be expected to accumulate and the peaks at the centre would be higher. However, when an alternate layout cloning the single finger transistor from the source, the peak temperature of the alternate layout is lower than the peak temperature of the original layout, even though there are more drain contacts present. The access regions are taken to be both 0.95 μm in length each and the gate contact 0.1 μm. Work was done with varying access region lengths asymmetrically, but is not presented here as Equation G.11 explains the differences mathematically.

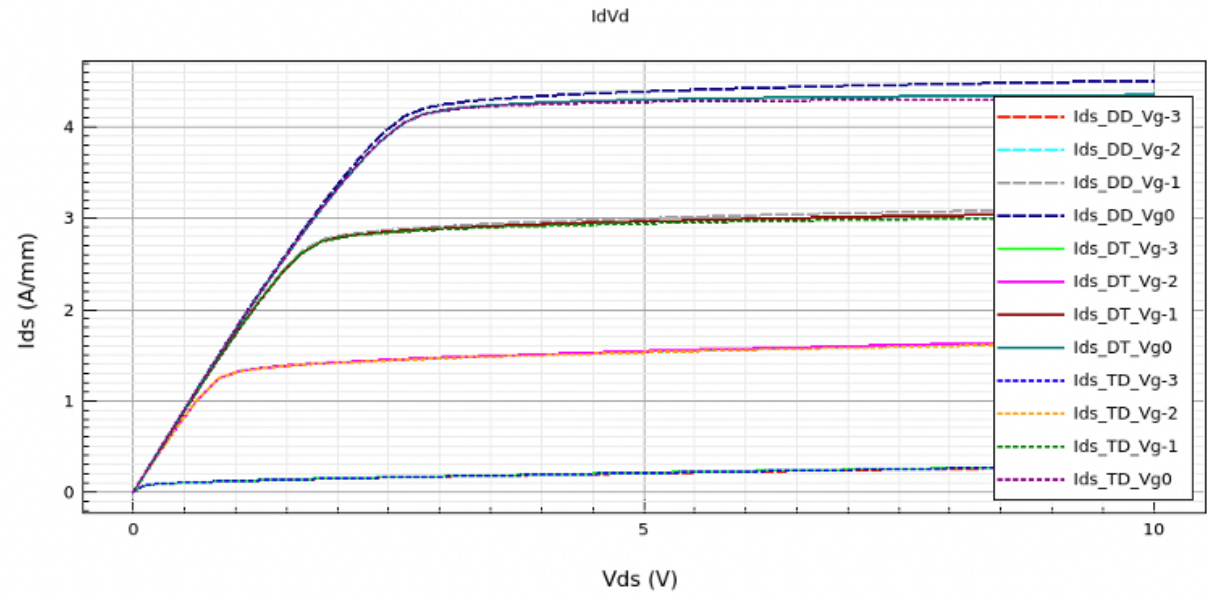


Figure G.4: IV curves with different temperature models for 4-finger transistor.

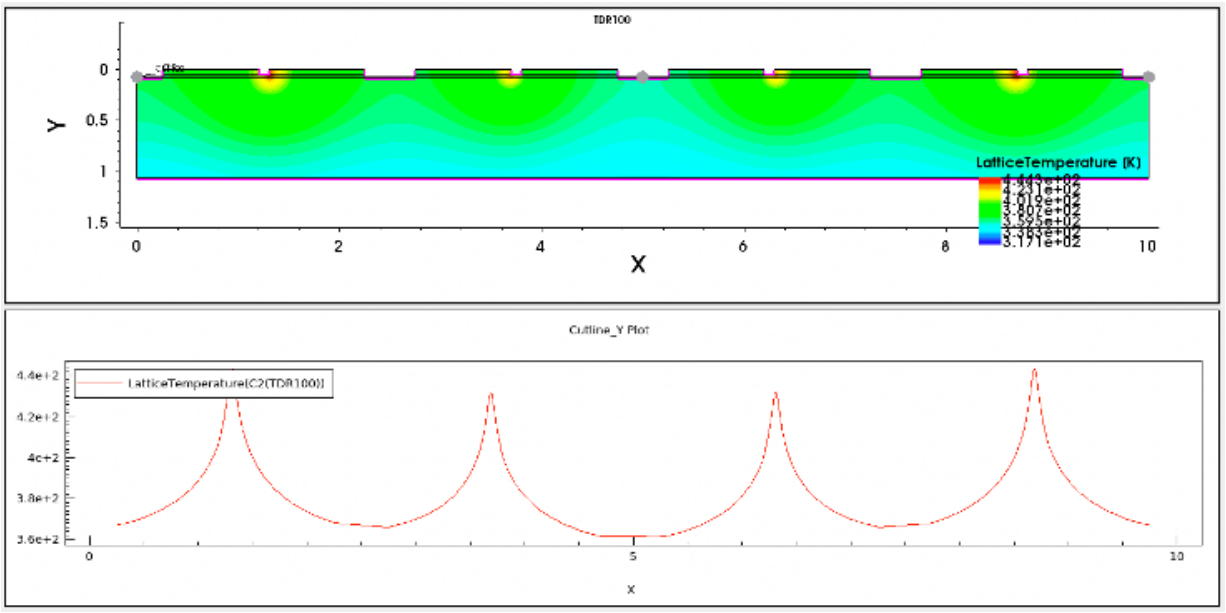


Figure G.5: Cross-sectional heat map and lattice temperature of 4-finger GaN HEMT.

Figure G.6 displays the electrostatic potential and lateral electric field across the four fingers taken at the cutline y=0.0725. The different transport/temperature models are compared as well. The gate voltage was swept from -2 to 2 while the drain voltage was swept from 0 to V_{dmax}, which was set to 10 V. From the lateral electric field plot, it is noticeable that a larger gate voltage corresponds to a smaller amplitude. Here, the default temperature model and thermodynamic model align nicely. When different drain voltage values were taken, a larger drain voltage corresponded to a larger lateral electric field amplitude.

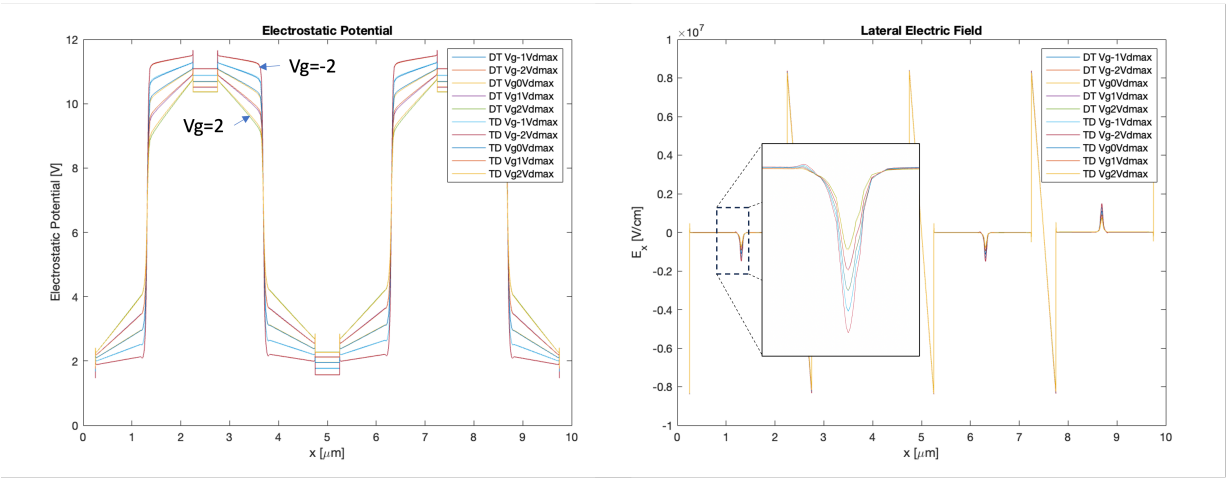


Figure G.6: Electrostatic potential and lateral electric field.

Figure G.7 displays the lattice temperature and total heat of the device along the same cutline. The differences between the default temperature model and the thermodynamic model are more noticeable in the lattice temperature plots, especially near the metal contacts. In both of these plots, a larger gate voltage leads to more heating in the device. Likewise, when different drain voltages were taken, the larger drain voltage would lead to a higher lattice temperature and a greater total heat generated. Discontinuities are present in the total heat plot and occur at each of the edges of the metal contacts.

Figure G.8 displays the electron mobility and electron velocity at the same cutline. Increasing either the gate voltage or drain voltage increased the temperature, leading to more scattering and a lower electron mobility. Likewise, increasing the voltages also led to a higher thermal/kinetic energy which would mean a higher electron velocity. Each of the plots presented look similar to the single finger plots, but are repeated for each of the four fingers. For information on the single finger plots, the reader may refer to Chen's work in [81].

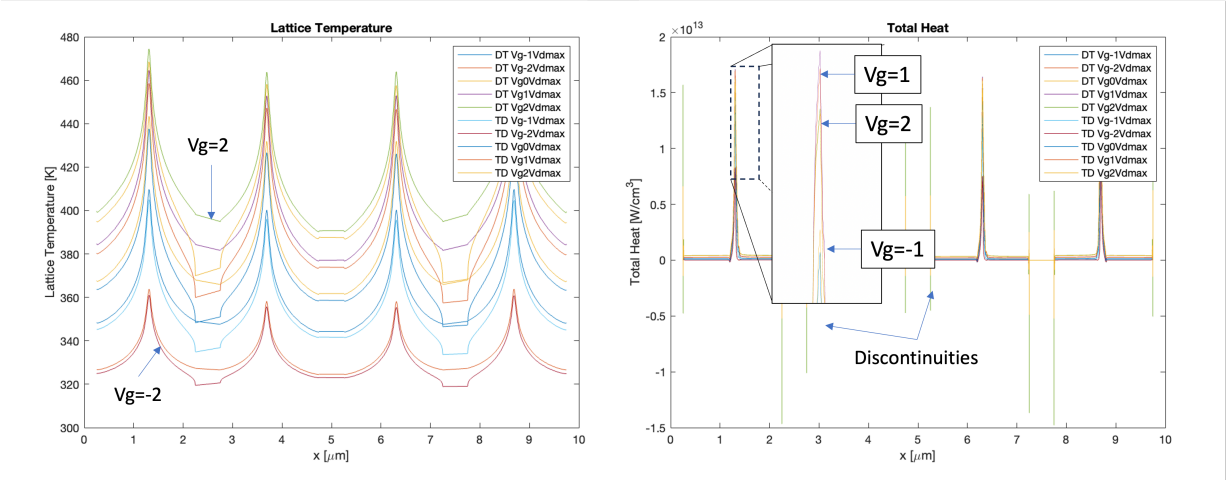


Figure G.7: Lattice temperature and total heat.

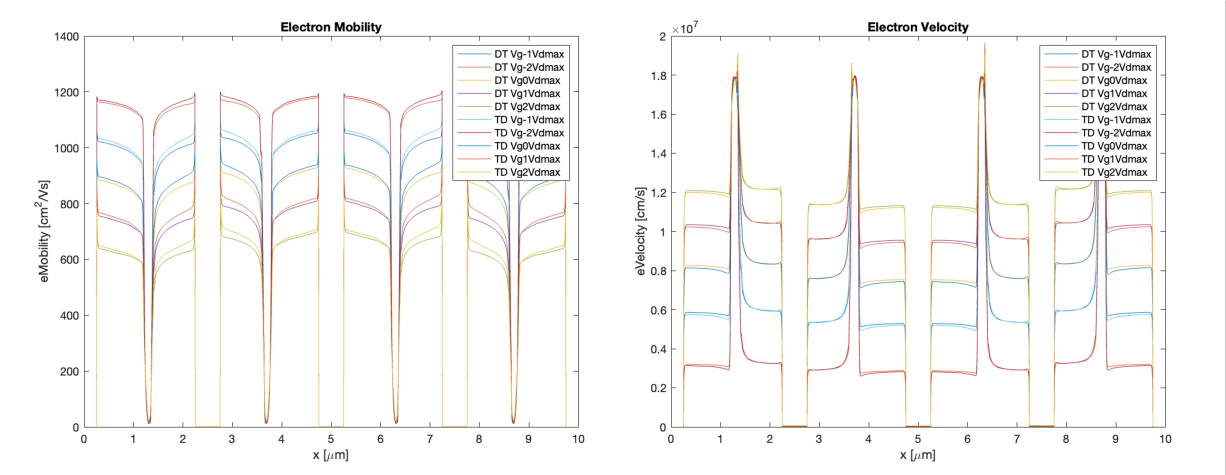


Figure G.8: Electron mobility and electron velocity.

Due to limited resources and time constraints, the electrothermal characterization of GaN HEMTs remain largely a work in progress. There is a possibility that the heat maps and results of the GaN HEMTs presented here are not realistic, as there is no experimental electrothermal data that could be generated to prove the simulated results.