Parametric Inversion of Experimental Electromagnetic Data Using Synthetically-Trained Neural Networks

by

Keeley Edwards

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada

Abstract

Electromagnetic imaging is a method that aims to reconstruct a spatial map of the material properties of a target through non-invasive measurements, allowing an observer to see inside objects, such as the human body, geological formations or large containers, without disturbing or entering them. This thesis evaluates the use of synthetically trained neural networks for performing phaseless parametric inversion of experimental electromagnetic imaging data. Fully-connected neural networks are used as a tool to obtain bulk parameters describing the geometry and complex-valued permittivity of materials within the region of interest (ROI) from experimental data. This data can subsequently be used for calibration and as prior information to facilitate full-phase 3D inversion of the data. These networks are applied to two different applications, one in agriculture and one in biomedical imaging, demonstrating that this method can be generalized to imaging systems at significantly different scales.

In the agricultural application, synthetically-trained neural networks are used to infer bulk parameters from raw, uncalibrated experimental data for two different grain storage bins. This presents a promising tool for calibrating data collected from uncooperative systems where conventional calibration methods are not practical. For commercial grain monitoring systems, where multiple measurements are taken daily, neural networks present a cost-effective long-term alternative to existing iterative parametric inversion methods.

For the biomedical application, a two-stage workflow is presented, demonstrating the suitability of the synthetically trained machine-learning-based parametric inversion to obtain the prior information needed for a second stage inversion using the Contrast Source Inversion (CSI) method. Noisy synthetic data representing a model of a human breast is used to simulate experimental data in the two-stage workflow, and preliminary bulk parametric inversion (neural network) results are given for experimental microwave

breast imaging data.

The results presented herein demonstrate that synthetically-trained neural networks can successfully infer bulk parameters from noisy synthetic and experimental data. Synthetic experiments for the microwave breast imaging case show that the inferred parameters are of sufficient quality to be used as prior information for full phase inversion and 3D image reconstruction.

Contributions

Journal Papers

K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, I. Jeffrey, "A Machine Learning Workflow for Tumor Detection in Breasts using 3D Microwave Imaging," *Electronics*, vol. 10, no. 6, 2021.
Personal contributions: Code development (including adding support for the breast imaging system), generation and analysis of results, manuscript preparation.

K. Edwards, N. Geddert, K. Krakalovich, R. Kruk, M. Asefi, J. LoVetri, C. Gilmore, and I. Jeffrey, "Stored Grain Inventory Management Using Neural-Network-Based Parametric Electromagnetic Inversion," IEEE Access, vol. 8, pp. 207 182–207 192, 2020.
Personal contributions: Code development (including adding support for multi-frequency data analysis, developing additional neural network architectures), generation and analysis of results, manuscript preparation.

Conference Papers

K. Edwards, J. LoVetri, C. Gilmore, I. Jeffrey, "A Machine Learning Method for Characterization of Complex Grain-Air Interfaces in Grain Storage Bins," Accepted to ANTEM 2021 as a 2-page conference paper (June, 2021).

K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, I. Jeffrey, "Recovery of Prior Information for Breast Microwave Imaging Using Neural Networks," Accepted to URSI GASS 2021 as a 2-page conference paper (March, 2021).

K. Edwards, K. Krakalovich, R. Kruk, V. Khoshdel, J. LoVetri, C. Gilmore, and I. Jeffrey, "The implementation of neural networks for phaseless parametric inversion," 2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science, Rome, Italy, 2020, pp. 1-3.

Provisional Patent (Application Filed)

K. Edwards, I. Jeffrey, C. Gilmore, J. LoVetri, M. Asefi, N. Geddert, R. Kruk, K. Krakalovich, Stored Grain Inventory Management Neural Network, Provisional filed August 28, 2020, U.S. Patent.

Code Base

The machine learning workflow discussed in this work was implemented in Python and TensorFlow Keras. The resulting code base (built off of preliminary work by K. Krakalovich and R. Kruk) can be configured through a single input file to train a neural network and evaluate test examples. A user manual has been written, providing instructions for configuring the code through the input file.

Acknowledgements

Firstly, thank you to my advisors, Dr. Ian Jeffrey and Dr. Colin Gilmore, for your continued support and encouragement throughout my M.Sc. Your knowledge, guidance, and friendship are sincerely appreciated.

I would also like to thank my colleagues in the Electromagnetic Imaging Lab and at AGCO Winnipeg, specifically Vahab, Ryan, Kennedy, Nick, and Mohammad for their support and contributions to this research.

Thank you to my M.Sc. reviewing committee Dr. Christopher Henry and Dr. Stephen Pistorius for reviewing my thesis, evaluating my work and providing valuable feedback.

I wish to express my gratitude for the computational resources made available to me that were funded through the Canadian Foundation for Innovation, and for the funding received from AGCO and Mitacs for my research on stored grain monitoring. In addition to their financial support, I would like to thank AGCO Winnipeg and its employees for their support, expertise and access to experimental data.

Contents

List of Tables **ix**

List of Figures **x**

1 Introduction **1**

1.1 Bridging the Gap Between Synthetic and Experimental Data 2

1.2 Thesis Scope 2

1.2.1 Stored Grain Monitoring and Inventory Management 2

1.2.2 Tumor Detection and Monitoring in a Model of a Human Breast . 3

1.2.3 Code Base 3

1.3 Thesis Outline 4

2 Background **5**

2.1 Inverse Scattering Problems 5

2.2 Electromagnetic Imaging 8

2.2.1 Prior Information 10

2.2.2 The Forward Model 10

2.2.3 The 24-Transceiver Imaging System 11

2.3 Parametric Inversion 12

2.4 Neural Networks 16

2.4.1 Supervised Machine Learning 17

2.4.2 Fully-Connected Neural Networks 17

2.4.3 ReLU Activation 17

2.4.4 Some Additional Machine Learning Terminology 18

3 The Bulk-Parametric Inversion Neural Network **20**
3.1 Labelled Synthetic Dataset . . . 21
3.2 Data Pre-Processing . . . 21
3.3 Network Architecture . . . 23
3.4 Loss Function . . . 23
3.5 Performance Metrics . . . 24

4 Stored Grain Inventory Management and Monitoring **26**
4.1 Introduction . . . 26
4.2 Experimental Grain Bins . . . 27
4.3 Experimental Data Collection . . . 28
4.4 Modeling the Grain Bin . . . 28
4.4.1 Grain-Air Interface . . . 29
4.4.2 Labelled Data . . . 30
4.5 Neural Network Architectures . . . 31
4.6 Results . . . 33
4.6.1 Experimental Test Data and Ground Truth . . . 34
4.6.2 Labelled Synthetic Dataset . . . 35
4.6.3 Single- vs. Multi-frequency Networks . . . 36
4.6.4 Neural Network Training . . . 38
4.6.5 Height, Cone Angle, and Volume Prediction Accuracy . . . 39
4.6.6 Permittivity Prediction Accuracy . . . 45
4.6.7 Performance on Synthetic Test Sets . . . 46
4.6.8 Complex Grain-Air Interfaces . . . 46
4.7 Chapter Summary . . . 48

5 Microwave Imaging of a Human Breast Model **50**
5.1 Introduction . . . 50
5.2 Faceted Air-Filled Chamber and its Model for Microwave Breast Imaging 51
5.3 A Two-Stage Workflow for Prior Information Extraction and Data Inversion 53
5.3.1 Stage 1: Bulk Parameter Inference Using Neural Networks . . . 54
5.3.2 Stage 2: 3D Image Reconstruction and Tumour Detection . . . 57
5.3.3 Calibration . . . 58

5.4 Results 58
5.4.1 Data Generation 59
5.4.2 Tumor Detection Test Samples 60
5.4.3 Parametric Inference of Fibroglandular Region Parameters 60
5.4.4 CSI-Based Tumor Detection From Predicted Prior Information . . 63
5.4.5 Monitoring Response to Tumor Treatment 65
5.4.6 Preliminary Bulk Parameter Inference Results on Experimental Data 66
5.4.7 Multi-Frequency Neural Network Performance 67
5.5 Chapter Summary 69

6 Conclusions and Future Work **70**

Bibliography **73**

List of Tables

3.1 Approximate time required for each step of the machine learning workflow. 25

4.1 Description of experimental grain bins. 27
4.2 Summary of labelled experimental stored grain data sets. 34
4.3 Summary of synthetic stored grain data sets. 35
4.4 First evaluated neural network architectures. 36
4.5 Initial results for average absolute error in height and cone angle predictions. 38
4.6 $\underline{\mathbf{p}}_4$ neural network architectures and training characteristics. 39
4.7 Average absolute error for height and cone angle predictions for Bin 1 and Bin 2. 40
4.8 Average absolute error for $\underline{\mathbf{p}}_6$ grain bin predictions. 48

5.1 Summary of fibroglandular region variations in the labelled dataset. . . . 59
5.2 Error in fibroglandular region bulk parameter predictions for the noise-free, tumor-free synthetic test set. 60
5.3 Fibroglandular region bulk parameter predictions for synthetic examples with -80dB noise. 61
5.4 Error in Error in fibroglandular region bulk parameter predictions for the noisy, tumor-containing test cases. 62
5.5 Fibroglandular region bulk parameter predictions for calibrated experimental examples. 67
5.6 Fibroglandular region bulk parameter predictions from the single-frequency neural network. 68
5.7 Fibroglandular parameter predictions from the multi-frequency neural network. 69

List of Figures

2.1 Depiction of experimental and model-based imaging systems. 8
2.2 Renders of regions of high relative permittivity for stored grain and microwave breast imaging chambers. 9
2.3 Illustration of the electromagnetic imaging system. 12
2.4 Visual representation of the target characterized by bulk parameters. . . 13
2.5 Comparison of cooperative and uncooperative systems. 15
2.6 Fully-connected neural network. 18
2.7 Plot of ReLU and Leaky ReLU activation functions. 19

3.1 Transmitter-by-transmitter normalization scheme. 22
3.2 Feature normalization / scaling illustration. 22
3.3 General bulk parametric inversion neural network architecture. 23

4.1 Cross-section render of grain bins. 28
4.2 Render of the grain bin imaging system installation. 29
4.3 Simple stored grain-air interfaces. 30
4.4 Description of the 4- and 6-parameter labels for stored grain. 31
4.5 Synthetic labelled data generation workflow. 32
4.6 Fully-connected, 5-hidden layer neural network for a single frequency and four output parameters. The number of nodes in each layer are given below each layer. 32
4.7 Fully-connected, 6-hidden layer neural network for $n_{\text{freq}} = 4$ frequencies and 4 output parameters. 33
4.8 Fully-connected, 7-hidden layer neural network for an arbitrary number frequencies, n_{freq}, and 6 output parameters. 33

4.9 Bulk parametric inversion results for the simplex method and the first neural networks evaluated. 37
4.10 Experimental bulk parametric inversion results for the Hopper-style grain bin. 41
4.11 Experimental bulk parametric inversion results for the flat-bottomed grain bin. 42
4.12 Tolerance plots for experimental stored grain test data. 44
4.13 Tolerance plots for experimental test data and synthetic test data. 47
4.14 Prediction profile of the grain-air interface from $\underline{\mathbf{p}}_6$ networks. 49

5.1 Photograph of the experimental microwave breast imaging setup with breast phantoms (blue). 52
5.2 Two-stage, machine-learning-enabled microwave breast imaging workflow. 53
5.3 Renders of the breast tissue regions within the faceted air-filled chamber. 55
5.4 Stage 1 of the two-stage microwave breast imaging workflow. 56
5.5 Stage 2 of the two-stage microwave breast imaging workflow. 57
5.6 True and predicted values for noisy-synthetic data from three different breast fibroglandular region geometries. 62
5.7 CSI reconstructions using neural network predicted prior information for different tumor locations and breast geometries. 64
5.8 CSI reconstructions based on neural network predicted prior information for two different tumor sizes and no tumor. 65
5.9 Alternative model of the microwave breast imaging system with a domed ABC chamber top. 66

Glossary

ABC Absorbing Boundary Condition

APIs Application Programming Interfaces

CNNs Convolutional Neural Networks

CSI Contrast Source Inversion

DGM Discontinuous Galerkin Method

EIL Electromagnetic Imaging Laboratory

FEM Finite Element Method

GANs Generative Adversarial Networks

ISP Inverse Scattering Problem

MAE Mean Absolute Error

MC Moisture Content

MWI Microwave Imaging

MSE Mean Square Error

NN Neural Network

PEC Perfect Electric Conductor

ReLU Rectified Linear Unit

ROI Region of Interest

SI Simplex Inversion

VNA Vector Network Analyzer

To my family and to Chai.

Chapter 1

Introduction

Electromagnetic imaging is a technique that aims to characterize the electromagnetic properties of material targets and reconstruct a map of these properties. It is a type of inverse problem in which the general goal is to gain an understanding of a target from measurement data. More specifically, electromagnetic imaging is an of Inverse Scattering Problem (ISP), a family of ill-posed problems (problems in which a small change in observation / measurement can yield a large change in the calculated solution [1]) which are often non-linear and computationally expensive to solve [2]. Reconstructing images of physical systems typically requires accurate prior information about the target to provide an initial guess (starting point) for the inversion algorithm [3–6]. Inversion algorithms struggle when there is high contrast between material parameters, and the inclusion of prior information can help to reduce the contrast [7]. This prior information can be provided to the algorithm in the form of bulk parameters about the ROI, where the parameters may include information on the average complex-valued permittivity and / or geometry of the material [8]. This thesis is focused on applying machine learning, specifically supervised machine learning using fully-connected neural networks, to accurately obtain this prior information by considering only the phaseless component of the data and facilitate the inversion of full-phase electromagnetic measurement data.

1.1 Bridging the Gap Between Synthetic and Experimental Data

Supervised machine learning relies on large datasets for model training that are representative of, and extensively span, the range of anticipated test cases [9, 10]. For many applications of electromagnetic imaging obtaining large experimental labelled datasets is difficult, if not impossible. It is, however, possible to synthetically model the measurement system using forward solvers and build large synthetic labelled datasets. These datasets can be specifically designed to span the expected range of experimental test cases.

This work presents a method for extracting bulk information from experimental data using a neural network that was trained using only synthetic data. It represents a tool that can use synthetic data obtained from a computational model of the system to: 1) reduce the need for costly and challenging calibration and measurement steps that are required to process experimental measurement data, and 2) provide the prior information necessary to support full inversion (image reconstruction) of the data [8].

1.2 Thesis Scope

An end-to-end workflow for training a machine learning model on synthetic electromagnetic data is presented and used to determine important prior information about experimentally measured electromagnetic imaging targets. Synthetic data based on a computational model of the system is used to train the neural network, which can subsequently be used to evaluate raw experimental data. This workflow can be adopted for, and applied to numerous different applications, and this work presents results for two applications: stored grain monitoring and inventory management, and tumor detection and monitoring in a model of a human breast.

1.2.1 Stored Grain Monitoring and Inventory Management

Electromagnetic imaging of grain stored in a large metallic grain bin can provide insight into the geometry of the grain-air interface along with the moisture content of the stored grain, making it a valuable tool for inventory management and monitoring for signs

of grain spoilage [11]. This work presents and evaluates several fully-connected neural network architectures for processing single- and multi-frequency measurement data for two different grain bins with distinct geometries, and types of stored grain (wheat and corn). Large synthetic labelled datasets representing the various possibilities for grain fill patterns with simple grain-air interfaces were generated in both cases. Neural networks were trained on the synthetic labelled data and used to predict the grain fill properties of uncalibrated experimental measurements taken from each of the two bins. A brief investigation of the ability of an expanded neural network architecture to predict more complex grain-air interfaces is also presented.

1.2.2 Tumor Detection and Monitoring in a Model of a Human Breast

Microwave imaging of the human breast is being studied as a non-invasive tool to monitor for the presence of tumors using non-ionizing radiation [12]. To demonstrate the utility of the machine learning workflow presented herein for this type of problem, a large synthetic dataset was generated for a model of a human breast. The synthetic model represents an experimental breast phantom and imaging chamber developed by the Electromagnetic Imaging Laboratory (EIL) [13, 14]. A neural network was designed to accept data from this chamber as inputs where the outputs of the network are bulk parameters representing the geometry and complex-valued permittivity of the fibroglandular region of the breast. For this application, noisy synthetic test data containing tumors of various sizes and locations was evaluated by using the network to infer bulk geometry and permittivity of the fibroglandular region of the breast. These inferred parameters were then used as prior information for inversion of the data to demonstrate that the neural network can provide prior information of sufficient quality to enable full inversion of the data for tumor detection and localization as well as tumor monitoring over time. Preliminary results for the inference of bulk parameters from calibrated experimental data are presented.

1.2.3 Code Base

This workflow has been refined and expanded following on the work of previous members of the EIL (K. Krakalovich and R. Kruk). It handles the necessary data packaging, and

pre- and post-processing, in addition to neural network training and inference (testing). The entire workflow can be configured through a single input file for convenient use. At this time, the code base includes a number of different network architectures, and additional architectures can be easily added in the future.

1.3 Thesis Outline

Chapter 2 provides additional background on inverse problems, electromagnetic imaging, parametric inversion, and neural networks.

Chapter 3 gives an application-agnostic description of the machine learning workflow. It describes the structure of the data and processing steps that enable the workflow, and provides a general overview of the neural network. A brief analysis of the computational cost compared to iterative, optimization-based parametric inversion is provided.

Chapter 4 focuses on the agricultural application of the workflow: stored grain monitoring and inventory management. Experimental results are presented for two different grain bin geometries and grain types. The model is extended to a more complex parameter set, and results are presented for noisy synthetic examples.

Chapter 5 explores the biomedical application of the workflow for tumor detection and monitoring in a model of a human breast. This chapter proposes an end-to-end data-to-image workflow that is enabled by the parametric inversion neural network. The performance of the network for noisy synthetic test cases containing tumors is evaluated, and preliminary results for experimental measurements are discussed.

Chapter 6 provides conclusions and focuses on possible areas for future work and extensions of this research.

Chapter 2

Background

2.1 Inverse Scattering Problems

Inverse problems are a family of problems where the goal is to determine the properties of a region of interest (ROI) by measuring the response to a known source. They can be used for imaging (tomography) in a wide range of fields including geology [15], remote sensing [16], agriculture [11], and medicine [12]. There are broadly two main types of inverse problems: inverse source problems, and inverse scattering problems. In the inverse source problem, the goal is to characterize a source embedded within the ROI, representing a problem that is inherently non-unique. In the inverse scattering problem, a source outside the ROI is applied and the response is measured at receiver points outside the ROI. The solution of the inverse scattering problem with complete information[1] is unique [2]. Practically, the solution is always non-unique as it is not possible to obtain complete information from a real-world imaging system.

Although related, the characteristics of these two types of problems are distinct. The two problem types are introduced for the reader's awareness, but as inverse scattering problems (ISPs) are the focus of this work, the remaining discussion focuses on inverse scattering problems.

ISPs provide a means of seeing inside an object without physically entering the object, providing a powerful imaging tool. In the case of electromagnetic imaging an "image" is a spatial map of the electromagnetic properties (conductivity, permittivity, permeability)

[1]Complete information, as defined in [2], is the response of the object for all possible positions of transmitters and receivers at all frequencies, which for practical purposes is impossible.

within the ROI. ISPs are difficult problems to solve, as they are non-linear, and ill-posed (unstable). Additionally, without complete information, which is impossible to obtain in practical situations, they are non-unique (i.e. have infinitely many solutions) [2]. It is these characteristics that make ISPs challenging and computationally expensive to solve.

The complement of the ISP is the forward problem (also called the direct problem), in which the goal is to determine the response of a system based on knowledge of the properties of that system. Where the inverse problem is non-unique and ill-posed, the forward problem has a unique solution and can be solved using computational models of the system in question. Obtaining a solution to the forward problem is limited by the ability to accurately model the system in question, where in many cases accuracy is traded for model simplicity and reduced computational cost.

Electromagnetic inverse problems are governed by Maxwell's curl equations [2]:

$$\begin{aligned} -\nabla \times \mathbf{H} + (j\omega\varepsilon' + \sigma)\mathbf{E} &= -\mathbf{J} \\ \nabla \times \mathbf{E} + j\omega\mu\mathbf{H} &= -\mathbf{K} \end{aligned} \tag{2.1}$$

assuming a linear, non-dispersive, isotropic medium; where $\mathbf{H}$ are the total magnetic fields, $\mathbf{E}$ are the total electric fields, $\mathbf{J}$ are the impressed electric currents, $\mathbf{K}$ are the impressed magnetic currents, ε' is the permittivity, σ is the conductivity, μ is the permeability, and ω is the angular frequency. Herein, a single variable, $\varepsilon = \varepsilon' + j\varepsilon''$, referred to as the complex-valued permittivity, is used to account for both the permittivity, ε', and the effective conductivity, where $\varepsilon'' = -\sigma/\omega$, such that:

$$j\omega\varepsilon' + \sigma = j\omega\varepsilon. \tag{2.2}$$

In order to reconstruct the fields ($\mathbf{H}$ and / or $\mathbf{E}$), knowledge of ε and μ is required. By assuming some known background model, Equations 2.1 can be formulated in terms of the background physics and incident fields $\mathbf{H}^{inc}$ and $\mathbf{E}^{inc}$ that are supported by the sources $\mathbf{J}$ and $\mathbf{K}$ in the background medium:

$$\begin{aligned} -\nabla \times \mathbf{H}^{inc} + j\omega\varepsilon_b\mathbf{E}^{inc} &= -\mathbf{J} \\ \nabla \times \mathbf{E}^{inc} + j\omega\mu_b\mathbf{H}^{inc} &= -\mathbf{K} \end{aligned} \tag{2.3}$$

where $\varepsilon_b = \varepsilon_b' + j\varepsilon_b''$ and μ_b, represent the background physics of the problem.

The scattered fields relative to a known background arising from some target within the ROI are then defined as the difference between the total fields and the incident fields:

$$\begin{aligned} \mathbf{H}^{sct} &= \mathbf{H} - \mathbf{H}^{inc} \\ \mathbf{E}^{sct} &= \mathbf{E} - \mathbf{E}^{inc}. \end{aligned} \tag{2.4}$$

It is straightforward to show from Equations 2.1 and 2.3 that the scattered fields satisfy:

$$\begin{aligned} -\nabla \times \mathbf{H}^{sct} + j\omega\varepsilon_b \mathbf{E}^{sct} &= -j\omega(\varepsilon - \varepsilon_b)\mathbf{E} \\ \nabla \times \mathbf{E}^{sct} + j\omega\mu_b \mathbf{H}^{sct} &= -j\omega(\mu - \mu_b)\mathbf{H} \end{aligned} \tag{2.5}$$

where $\mathbf{H}^{sct}$ and $\mathbf{E}^{sct}$ are the scattered magnetic and electric fields, respectively, and $\mathbf{H}$ and $\mathbf{E}$ are total fields. In terms of Equations 2.5, the inverse problem aims to determine ε and μ, given knowledge of the sources, scattered and/or total fields, and the background physics. To mitigate some of the challenges associated with the non-linearity of inverse problems, rather than solving for ε and μ directly, the problem is solved for the contrast, χ:

$$\begin{aligned} \chi_\varepsilon &= \frac{\varepsilon - \varepsilon_b}{\varepsilon_b} \\ \chi_\mu &= \frac{\mu - \mu_b}{\mu_b} \end{aligned} \tag{2.6}$$

where χ_ε and χ_μ are referred to as the electric contrast and the magnetic contrast, respectively. This work is concerned with the reconstruction of non-magnetic material properties ($\chi_\mu = 0$), and thus the term contrast refers to electric contrast herein. This research is primarily focused on obtaining an initial guess of the background physics. As reconstructions are generally more successful for smaller contrasts, using this initial guess as prior information for full inversion reduces the non-linearity of the ISP in an attempt to improve inversion results.

To solve either the forward problem or the inverse problem computationally, the ROI is subdivided into many smaller regions using a mesh, as shown in Figure 2.1. The equations above are then enforced for every element in the mesh, yielding a linear system of equations that can then be solved for the fields [17].

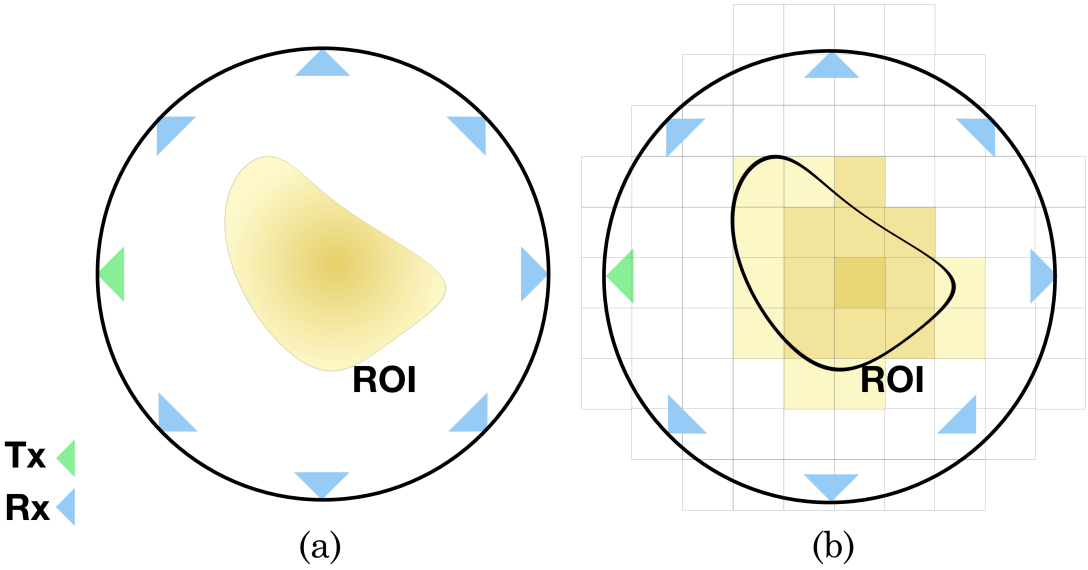


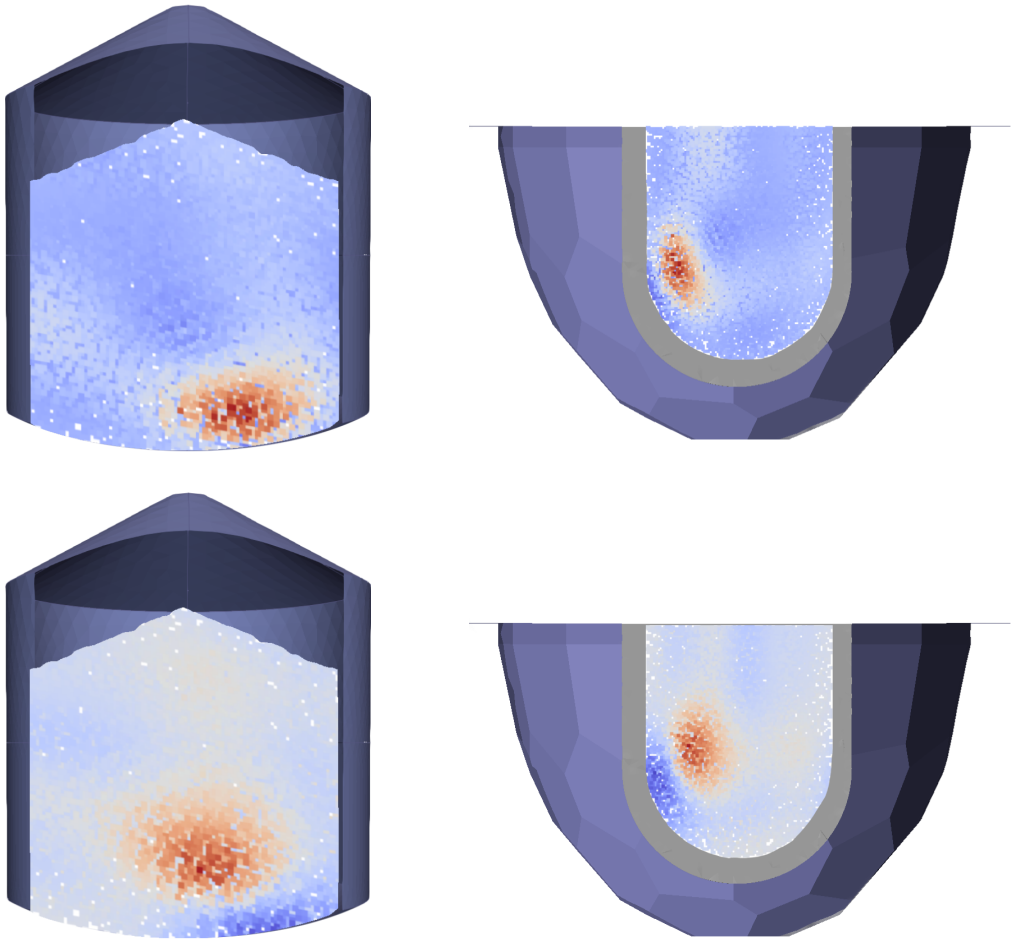
Figure 2.1: Example (a) experimental set-up for the inverse scattering problem and (b) model-based problem showing the meshed elements. White regions represent an air background and yellow shades represent varying permittivity of the target.

2.2 Electromagnetic Imaging

The 2D and 3D reconstruction of complex targets from electromagnetic data has been an active area of research for the inverse scattering community for several decades [9, 11, 12, 18]. These reconstructions rely on differences in the relative complex-valued permittivity ($\varepsilon = \varepsilon' + j\varepsilon''$) of the materials within the ROI to create images in the form of permittivity maps (Figure 2.2). The Contrast Source Inversion (CSI) method is one option for an iterative reconstruction algorithm that can be used to solve non-linear inverse scattering problems [19]. In the context of electromagnetic imaging, following data acquisition, calibration, and pre-processing, CSI can be used to generate permittivity maps. This process is used to validate the output of the neural networks developed in Chapter 5, and a more detailed discussion on CSI is provided in Section 5.3.2.1.

Two active areas of research are stored grain monitoring [20] and microwave imaging of the human breast [12]. The complex-valued permittivity of grain stored in large metallic grain bins is related to the moisture content of the grain, a characteristic which can affect the value of the commodity, and may be indicative of grain spoilage [21]. Being able to image the grain stored within in a bin can help to reduce waste and maximize profits from

the sale of the grain. In microwave breast imaging the various tissues within the breast, including adipose, fibroglandular, and tumor tissues, have significantly different relative permittivities such that the permittivity maps generated using electromagnetic imaging techniques can be used to detect and localize tumors [14, 22]. In both cases, one of the goals of electromagnetic imaging is to detect regions of high permittivity within the ROI. Examples of point-cloud permittivity maps resulting from image reconstruction using the CSI method are graphically depicted in Figure 2.2; where regions of high permittivity are shown in red. The air-filled imaging chamber is shown in purple to provide added context to the figure.



(a) (b)

Figure 2.2: Renders of regions of high relative permittivity in (a) the grain storage bin and (b) the breast imaging chamber. The real part of the reconstructed permittivity is shown in the top row, and the imaginary part is shown in the bottom row. Point-clouds represent the result of CSI where higher permittivities are shown in red and lower permittivities in blue. The relative size of the grain bin and breast imaging chamber are not to scale.

2.2.1 Prior Information

Independent of the application, accurate prior information about the target is imperative to the success of the reconstructed image [14, 23]. Information such as the approximate location, geometry, and permittivity of the target are all examples of prior information relevant to the ISPs discussed herein. Incorporating this prior information effectively fixes the estimate of the background physics of the materials, ε_b. Obtaining and improving prior information is a focus across the inverse scattering community [5–7, 24] and both 2D and 3D electromagnetic imaging has benefited from improved use of prior information [25, 26]. The source of the prior information can vary and can include spatial and/or non-spatial prior information from another imaging modality [3, 4, 27], another algorithm applied to the same data [28], or knowledge of the expected materials [29, 30].

2.2.2 The Forward Model

A computational model, referred to as the forward model, of the system is required to solve the electromagnetic ISP and reconstruct an image of the ROI [31–33]. Forward models solve forward (direct) problems: problems that seek to determine the fields when all aspects of the physical problem are assumed known. Throughout ISPs, forward models are used to simulate the fields for estimates of the target properties so that the simulated fields can be compared to the fields measured from the physical system.

When attempting to recover a high contrast target, such as a region of high moisture content in grain or a tumor in a human breast, the scattered fields arising from the target are of interest for inversion. As scattered fields cannot be measured directly, they are defined as the difference between the total fields and the incident fields relative to a known background (Equations 2.5); the background can thus be chosen depending on information available about the system. Prior information can be used to parameterize a ROI assuming no high contrast targets (for example, tumor-free breast tissue), describing it in terms of bulk geometry and / or bulk permittivity. The fields obtained from solving the forward problem for the parametric representation of the ROI serve as the incident fields (Equations 2.3) and can be used to compute the scattered fields arising from a high contrast target.

Any given imaging chamber can be modelled using CAD/meshing software, such as

`Gmsh` [34], an open-source software package that allows for the creation of a geometry and corresponding mesh files in two or three dimensions. The models contain information about the chamber geometry, boundary conditions at the surfaces (absorbing boundary condition (ABC) or perfect electric conductor (PEC)), and may optionally contain the location of the transceivers. A Finite Element Method (FEM) forward solver [17] or a more recently developed Discontinuous Galerkin Method (DGM) forward solver [33, 35] can be used to generate synthetic data based on the mesh file, along with information about the locations of the transceivers within the chamber.

2.2.3 The 24-Transceiver Imaging System

Although stored grain monitoring and breast imaging are clearly different applications, the physics of the two problems are fundamentally similar and the electromagnetic imaging systems developed by the Electromagnetic Imaging Laboratory (EIL) and used in this work are similarly designed [11, 13]. The most significant difference between the two applications is the size of imaging chambers (the microwave imaging chamber is approximately 20 cm in diameter at the opening whereas commercial grain bins can be as large as 40 m in diameter) and, consequently, the corresponding operating frequencies (tens of MHz for stored grain, and GHz for breast imaging).

The physical imaging chamber is modeled as having PEC walls, and having either an open or closed, PEC or non-PEC top. 24 transceivers (which may be antennas or field probes [11, 13]) are arranged along the walls of the chamber and are designed to measure the fields tangent to the chamber wall; this work considers the ϕ component of these fields, H_ϕ. The antennas are connected through cables to a switch and vector network analyzer (VNA) that measures the signals from the antennas as S-parameters [11, 36]. The result is a 24×24 matrix of S-parameters (referred to as the S-matrix, or simply S), where the $S_{i,j}$ entry corresponds to the ith transmitter and jth receiver. The matrix diagonal, $S_{i,i}$ where $i = 1, 2, \ldots 24$, represents the back-scattering measurement. Figure 2.3 shows the overall system setup.

The two forward solvers used with these systems (FEM and DGM) output field data (electric or magnetic) in x, y, z coordinates. To align with the experimental systems, a post-processing step based on the known polarizations of each transceiver converts these to the ϕ projection of the measurement tangent to the chamber surface. Where

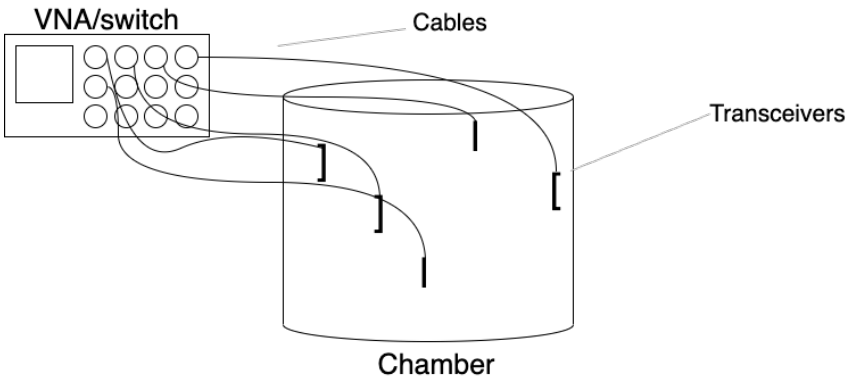


Figure 2.3: General depiction of the electromagnetic imaging system (not to scale or antenna / cable quantity), showing transceivers along the chamber walls, connected to a VNA/switch through coaxial cables.

in the experimental system the data are collected as S-parameters, the data from the forward solvers are H_ϕ field measurements. This work aims to demonstrate that fully-connected neural networks trained on phaseless synthetic (H_ϕ) data can accept phaseless experimental (S-parameter) data as test examples and accurately infer bulk parameters characterizing the contents of the ROI, (Figure 2.4).

2.3 Parametric Inversion

Electromagnetic imaging of complex, high-contrast targets is difficult without accurate prior information [7, 25]. Two examples of high-contrast targets include regions of high moisture content in stored grain, and tumors in the human breast. By initially seeking prior information as knowledge of the background, in the form of bulk parameters, the contrast of the target can be effectively reduced to simplify the non-linear inversion.

The background is taken to be everything within the imaging domain except the high-contrast target, and a set of bulk parameters can be used to describe this background, denoted as a vector, $\underline{\mathbf{p}}$. In the case of stored grain monitoring, the background is chosen to be the grain stored within the large metallic bin assuming uniform moisture content of the grain. The simplest form of this background described herein is characterized by grain height (h), cone angle of repose (θ) and complex-valued permittivity ($\varepsilon = \varepsilon' + j\varepsilon''$) such that $\underline{\mathbf{p}} = [h, \theta, \varepsilon', \varepsilon'']$. This choice allows for the detection of variations in the moisture content of the grain and will, in theory, help to identify regions of high moisture content during full inversion. For the breast imaging application, the background is chosen to

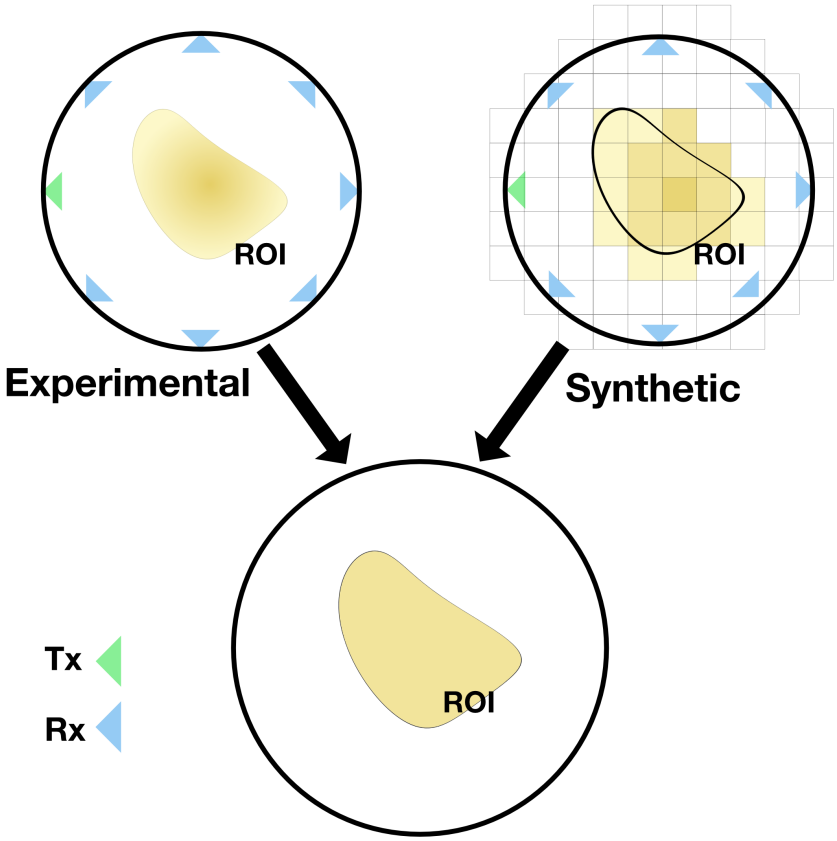


Figure 2.4: Parameters can be obtained to describe the bulk properties of the target from phaseless experimental (S-parameter) measurement data, or from phaseless synthetic (H_ϕ) data obtained from a computational model of the system.

be the tumor-free breast tissue, allowing for tumors (high-contrast targets) to be identified. The background then includes an adipose region with known characteristics (height, radius, complex-valued permittivity), and a fibroglandular region where the height (h), radius (r), and complex-valued permittivity ($\varepsilon = \varepsilon' + j\varepsilon''$) are to be determined by the network such that $\underline{\mathbf{p}} = [h, r, \varepsilon', \varepsilon'']$. Independent of application, these bulk parameters are a combination of geometry and complex-valued permittivity. For relatively simple bulk parameterization of the background (up to 6 parameters are evaluated in this work), appropriately designed optimization methods can be used to calculate background information. Other choices of background, such as non-uniform grain moisture content, are possible and may lead to improved results; however some choices may come at an increased problem complexity and computational cost.

In 2019, Gilmore *et al.* presented a phaseless parametric inversion technique using the Nelder-Mead Simplex Method [37] for the stored grain application [8]. For the phaseless parametric inversion technique, each of four bulk parameters serves as an optimization

parameter; the objective function that was used in this previous work is:

$$F(\underline{\mathbf{p}}) = \sum_{i,j,i\neq j} \left|\left| \alpha_i(\underline{\mathbf{p}})|S_{i,j}| - |H_{\phi,i,j}(\underline{\mathbf{p}})| \right|\right|^2 \tag{2.7}$$

where uncalibrated total field measurements are given by $S_{i,j}$, and synthetic field values at the field probes are given by $H_{\phi,i,j}$. $\underline{\mathbf{p}} \in \mathbb{R}^4$ is the vector of four design variables characterizing grain but could, in theory, be any parameter set of prior information. In [8] parameters were grain height, the cone angle of repose, and bulk complex-valued permittivity of the grain ($\varepsilon = \varepsilon' + j\varepsilon''$, where the real and imaginary parts each constitute a single parameter).

Transceiver-dependent data scaling parameters, α_i, in Equation 2.7 serve to account for differences in the transceivers. These scaling parameters are calculated as the ratio of the average values of the measurements and simulated fields for each transmitter:

$$\alpha_i(\underline{\mathbf{p}}) = \frac{\sum_j |H_{\phi,i,j}(\underline{\mathbf{p}})|}{\sum_j |S_{i,j}|} \tag{2.8}$$

This method is able to successfully determine bulk parameters of stored grain, enabling subsequent full reconstruction of moisture content (in the form of complex-valued permittivity) within an industrial grain bin [8].

Two key aspects of this parametric inversion method that should be highlighted are the use of 1) phaseless and 2) uncalibrated total field measurement data. Both of which are motivated, in part, by the challenges associated with calibrating some experimental systems. In a cooperative system, a calibration target can be introduced in the experimental system (and modeled in the model system). The measurements taken with this target in the experimental system can then be compared to those from the model system to calibrate the experimental system to the model [38]. Conversely, in uncooperative systems it is not practical to introduce a calibration target, and thus conventional calibration methods cannot be used. Figure 2.5 shows the difference between model and experimental systems for cooperative (a) and uncooperative (b) systems, using the grain bin as an example of an uncooperative system.

A full system calibration would provide the information necessary to map the experimental system to the synthetic model of the system, for each measurement point (transmit / receive pair), in the form of a calibration matrix. This calibration matrix

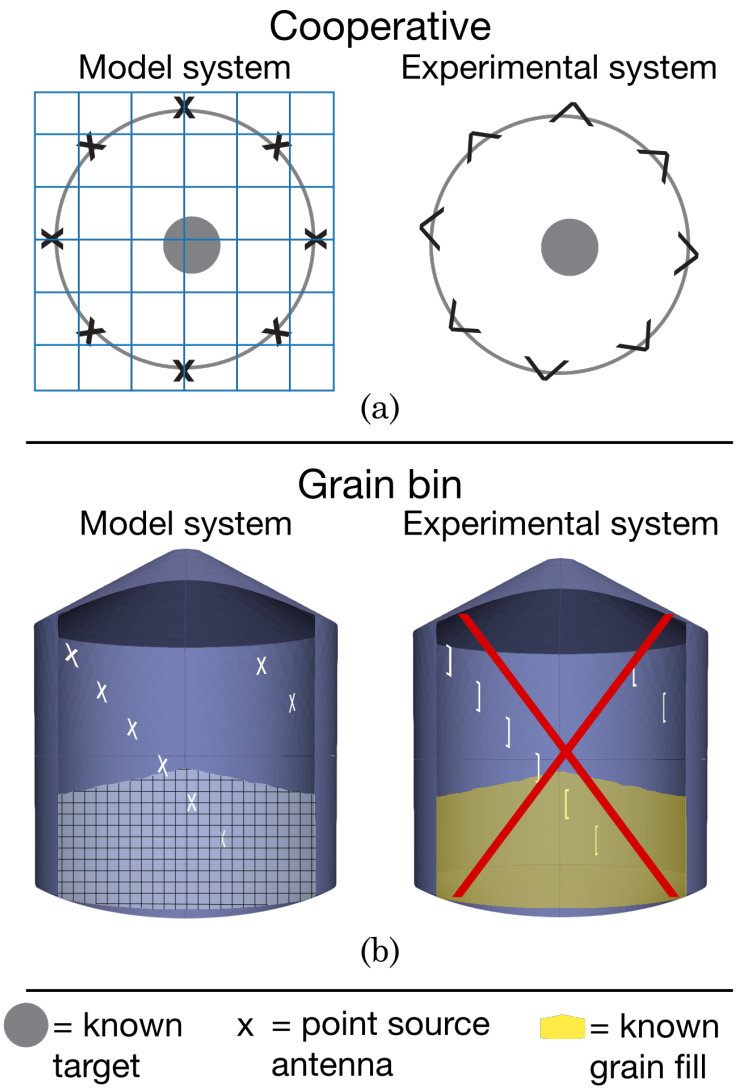


Figure 2.5: Comparison of model (left) and experimental (right) systems for (a) a cooperative system, where a known target can be introduced for calibration and (b) an uncooperative system, where a known target measurement cannot be taken (indicated by the red 'X').

would account for switch and cable losses within the system (which are often excluded from computational models for simplicity and to reduce cost), and would support full phase inversion [11]. In the case of the industrial grain bin it is not practical (and in some cases not possible) to introduce a known target into an empty bin to take a measurement and calibrate the imaging system. Furthermore, the calibration of the system is sensitive to external (environmental) factors and would require frequent re-calibration, which if done using conventional methods would require manipulating the grain within the bin to some known state.

The optimization-based method of phaseless parametric inversion, proven for the in-

dustrial grain bin application [8], provides a way to obtain the prior information needed for full inversion as well as a means of calibrating experimental measurements to the model system. One drawback of the method is that each experimental measurement requires many iterations of the simplex method, where each iteration requires a forward solver call, to determine the bulk parameters for the measurement.

This work explores the use of neural networks as an alternative method for inversion transferring this per-measurement computational cost to a one-time upfront cost for generating a large synthetic training set and training the neural network. By doing so, the machine-learning-based method provides a cost-effective long-term solution, where once the network is trained for a given system, bulk parameters from measurement examples can be determined in seconds.

2.4 Neural Networks

Over the last couple of decades, machine learning has been applied to the inverse scattering problem for both synthetic and experimental data [9]. Applications include parametric inversion of limb measurement data [39], and pixel-based inversion of various targets [6, 40, 41]. There are three major ways in which machine learning has been applied to ISPs: 1) image-to-image (pixel-based) inversion aims to improve the quality of the reconstructed image, where the original image may have been obtained using any number of methods [9], 2) data-to-image inversion where the network accepts measurement data and produces a reconstructed image [42, 43], and 3) parametric inversion, which is the focus of this thesis, is a data-to-parameter approach that accepts measurement data and outputs bulk parameter data that can be used to characterize a ROI and serve as useful prior information for a full inversion.

The novel contribution of this work is the use of synthetically trained neural networks for inference on experimental data, in some cases eliminating the need for other calibration. While this is accomplished using neural networks, the primary goal was to focus on applying machine learning in practical electromagnetic imaging systems. Therefore, this work does not focus on exploring a wide variety of neural network architectures or developing new ones, but rather focuses on demonstrating that machine learning can be used as a tool for experimental electromagnetic imaging. As this work is not intended to

be a machine learning thesis, the following sections provide a brief overview of relevant concepts and terminology used in this work. For a more informative and in-depth discussion on neural networks and their application to inverse problems, the reader is referred to [9, 44, 45].

2.4.1 Supervised Machine Learning

Machine learning methods can be classified by their supervision model: supervised, semi-supervised, and unsupervised learning are three classes of machine learning methods. Supervised learning, the focus of this work, relies on labelled datasets that can be used for network training. These datasets can be very large, and for many specialized applications it can be difficult to obtain a labelled dataset that sufficiently covers the problem space.

Semi-supervised and unsupervised learning methods offer an alternative to supervised learning. In recent years, Generative Adversarial Networks (GANs) have demonstrated the ability to create new training examples based on a limited training set, where the resulting training set could be used for supervised learning problems [46, 47].

2.4.2 Fully-Connected Neural Networks

Fully-connected, feed-forward neural networks are the focus of this work, and for the sake of brevity, other types of network architectures are not discussed. Fully-connected neural networks are networks where every layer is a fully-connected layer (Figure 2.6). A fully-connected layer is one in which every node in the nth layer is connected to every node in the $(n+1)$th layer.

Fully-connected networks benefit from the fact that they do not assume any structure / relationship between each of the network inputs. The use of fully-connected networks in this work is motivated by the fact that there is no immediately obvious spatial mapping from the measurement data to the 3D position within the imaging chamber for the systems used herein.

2.4.3 ReLU Activation

The activation function of a node determines the output of the node in response to the input at that node, and is another choice that must be made when designing a neural

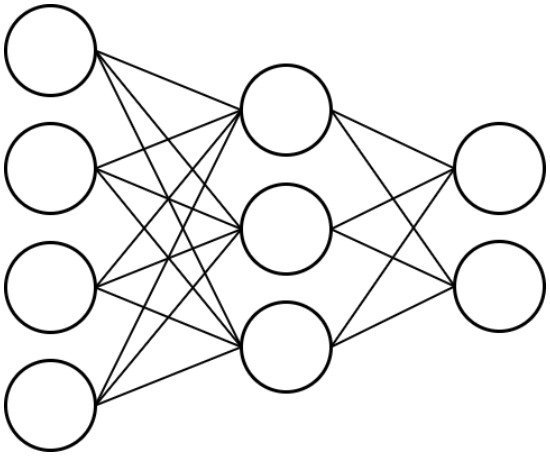


Figure 2.6: Simple depiction of a fully-connected neural network where every node in each layer is connected to every node in the following layer.

network. One of the most common, and often appropriate, activation functions is the rectified linear unit (ReLU) [48]. This activation function is linear for outputs greater than zero, and zero otherwise (Figure 2.7). ReLU is useful in preventing saturation, a problem seen when using other activation functions such as tanh and sigmoid [49]. ReLU only "activates" when the value at the node is greater than zero, this can lead to sparse models, which may be beneficial in preventing overfitting, but the effect can also be detrimental as a node may never activate, leading to something referred to as the dead or dying neuron problem [50]. A slight modification on the ReLU activation function is the leaky ReLU, where the slope of the activation function in the negative region is very small, and controlled by a parameter, α [50]. Figure 2.7 compares the ReLU and leaky ReLU activation functions. While leaky ReLU was briefly evaluated, ReLU was generally found to provide better results and was used for all results presented in this work.

2.4.4 Some Additional Machine Learning Terminology

The loss function is used to calculate the model error, and aims to minimize the error between the training labels and the output of the neural network; the loss function is user-defined when designing a neural network. During training, the weights in the network (value of each connection between nodes) are updated to minimize this loss. The loss function can be any user-defined function, but common choices are mean square error (MSE), and mean absolute error (MAE), both of which are provided as loss functions in machine learning application programming interfaces (APIs) [51]. As these may not be

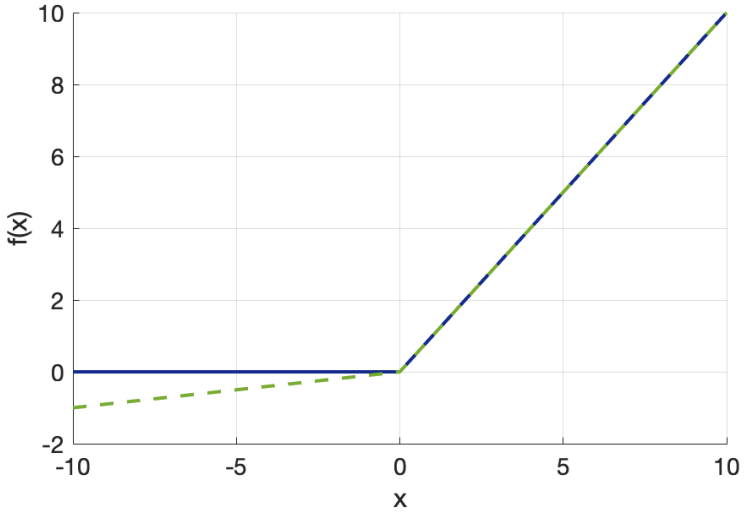


Figure 2.7: ReLU (solid blue) and leaky ReLU with $\alpha = 0.1$ (dashed green) activation functions.

suitable in all cases, machine learning frameworks also support the creation of custom loss functions.

The batch size and the number of epochs are two additional choices made when training a neural network. The training set typically consists of many (hundreds or thousands) training samples, where a sample is a single example in the training set. The user can set the batch size, which defines how many samples the network sees during training before updating the weights in the network. For example, if the training set consists of 10,000 samples, and the batch size is set to 1,000, the network will be updated 10 times during one pass through the entire dataset. In addition to the batch size, the user can set the number of epochs that the network is trained for. One epoch is defined as the number of times the network sees the entire dataset. Continuing the example above, if the number of epochs for training was 100, then the network would see a total of 1,000,000 samples during training.

Chapter 3

The Bulk-Parametric Inversion Neural Network

As discussed in Section 2.4, this work focuses on using supervised learning to train fully-connected neural networks for performing bulk parametric inversion to obtain prior information that can be used for electromagnetic imaging. For the networks to be useful practically, it is imperative that they can be trained on *synthetic* data but used to evaluate *experimental* data due to the challenges and limitations inherent in obtaining large experimental datasets; that is to say, it is impractical to obtain enough labeled experimental data to train a neural network using supervised machine learning. This presents an interesting problem, as there are unknown differences between model-based synthetic data generated using a forward solver and experimental data obtained from a system operating in a lab or in the field. The workflow discussed herein was originally developed for stored grain inventory management, and was later applied to microwave breast imaging. In the case of stored grain monitoring, system calibration can be difficult if not impossible, so one goal was to develop a neural network that can accurately recover bulk information about the chamber contents from experimental data without the need for calibration. In order to facilitate this, the parametric inversion is performed on phaseless (magnitude-only) data, a method that was previously shown to be successful using iterative optimization algorithms [8].

3.1 Labelled Synthetic Dataset

This section generally discusses the labelled synthetic dataset to provide context for the data that will serve as inputs to the neural network. Details of specific labelled synthetic datasets are provided in Chapters 4 and 5.

A large synthetic dataset with thousands of training examples covering the various configurations of the parameters of interest is generated for a given problem using a forward solver (FEM or DGM). The output of the data generation step is a large labelled dataset, where each example at a single frequency consists of 552 complex-valued $H_{\phi(i,j)}$ measurements, where the back-scattering measurement has been discarded[1]. Each example (552 measurements) is paired with a set of labels describing the parameterized model.

3.2 Data Pre-Processing

The data is subject to a pre-processing stage prior to being passed to the neural network for training. To mitigate the effect of unknown switch and cable losses, only phaseless data is considered and a per-transmitter normalization is applied to each example (Figure 3.1). The data from each transmitter is normalized to zero mean and unit variance for that transmitter where an input example, H_ϕ, is partitioned by transmitter into a vector, $\underline{\mathbf{x}}_{Tx_i}$, the mean, μ_{Tx_i}, and standard deviation, σ_{Tx_i}, for that transmitter are calculated, and the partitioned data is then normalized using z-scores:

$$\underline{\mathbf{z}}_{Tx_i} = \frac{\underline{\mathbf{x}}_{Tx_i} - \mu_{Tx_i}}{\sigma_{Tx_i}} \tag{3.1}$$

where the resulting vector, $\underline{\mathbf{z}}_{Tx_i}$ represents a given transmitter's measurements scaled to zero-mean with unit variance.

Following the transmitter-by-transmitter normalization, the full labelled synthetic dataset is partitioned 85%, 10%, 5% into training, validation, and testing sets, respectively. The training set mean, μ, and standard deviation, σ, of each feature are determined (Figure 3.2) and then the training, validation, and test examples are scaled using z-scores

[1]For the 24 transceiver system used herein, each of 24 transmitters has one measurement from each of 23 receivers (as the self (back-scattering) measurement is ignored). This results in $24 * 23 = 552$ data points per measurement.

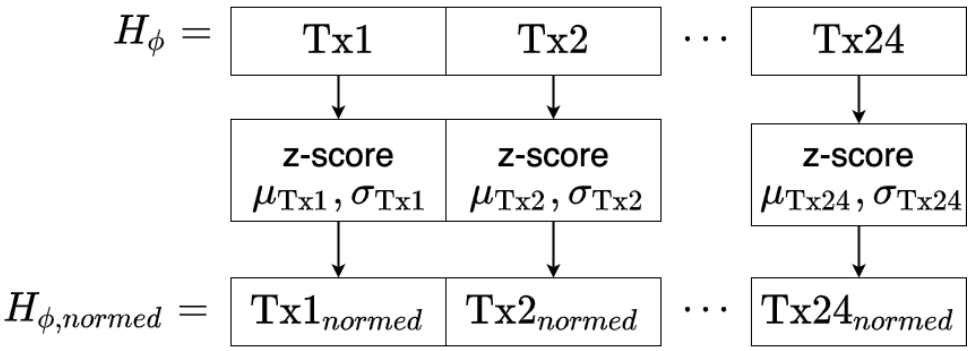


Figure 3.1: Transmitter-by-transmitter normalization scheme for data. The raw, phaseless example, H_ϕ, is partitioned by transmitter and normalized using z-score to give a normalized input example, $H_{\phi,normed}$.

(a variation of Equation 3.1):

$$z_{ij} = \frac{a_{ij} - \mu_j}{\sigma_j} \tag{3.2}$$

Similar to Equation 3.1, z_{ij} is the representation of a_{ij} scaled to zero mean with unit variance. In the case of Equation 3.2, the data to be scaled, a_{ij}, is a single data point (1 of 552 in the sample), which will be passed to one of the neural network input nodes.

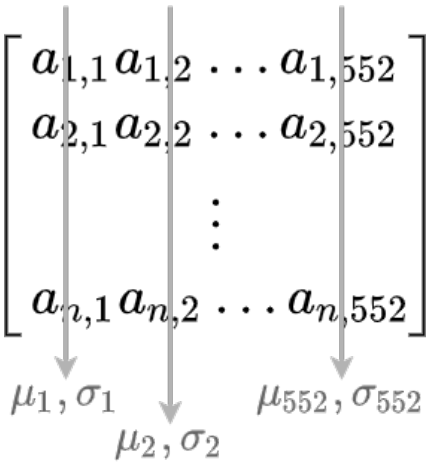


Figure 3.2: Mean and standard deviation calculation for feature scaling scheme, where n is the number of examples in the training set. The mean and standard deviation are computed for each feature (column).

The vectors $\underline{\mu}$ and $\underline{\sigma}$ obtained according to Figure 3.2 are stored with the model and used to scale each test example in a pre-processing step prior to passing it to the neural network for prediction.

3.3 Network Architecture

Several network architectures have been evaluated, all of which are fully-connected networks (Figure 2.6). Different variations perform better depending on the number of frequencies and number of parameters.

The workflow is configurable to accept single- or multi-frequency data, where in the multi-frequency case the input examples are a vertical concatenation of frequencies, resulting in a single column vector of length $552 * n_{\text{freq}}$.

The general network architecture is summarized in Figure 3.3, where n_{freq}, the number of hidden layers, and the size of the output layer vary depending on application.

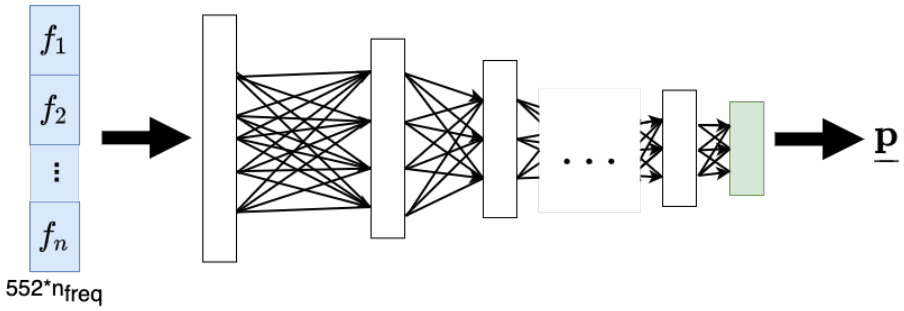


Figure 3.3: General bulk parametric inversion network architecture. The input layer to the fully-connected network is shown in blue and the output layer in green. The number of hidden layers (white) and nodes in each hidden layer varies depending on application.

3.4 Loss Function

The loss function chosen for a neural network varies depending on application, but in general is chosen to minimize the error between the network output and the training data labels. The loss function used in this work is the residual sum of squares loss function:

$$L(\underline{\mathbf{p}}) = \frac{(\underline{\mathbf{p}} - \underline{\mathbf{p}}_{true})^T(\underline{\mathbf{p}} - \underline{\mathbf{p}}_{true})}{\underline{\mathbf{p}}_{true}{}^T\underline{\mathbf{p}}_{true}} \tag{3.3}$$

where $\underline{\mathbf{p}}_{true}$ represents the true labels and T denotes the transpose of a vector.

As the bulk parameters to be recovered will vary in scale for any given application, each parameter is weighted equally in the loss function by normalizing the labels prior to training using z-scores (Equation 3.2, where the index j represents a parameter, rather

than a feature, when normalizing labels).

3.5 Performance Metrics

Training a neural network on synthetic data to evaluate experimental data requires that the physical system be somewhat stable over time or for the network to accurately infer bulk parameter information despite changes in the physical system. This is particularly important for uncooperative systems (stored grain monitoring) where calibration is not practical. Although data from cooperative systems (microwave breast imaging) can be calibrated regularly to account for system drift, it is still important that the synthetic data accurately to represent the experimental system.

The performance of the neural networks discussed herein will be primarily measured by their ability to accurately infer bulk parameters for each application. While the accuracy of inferred predictions is important, the computational cost associated with creating large synthetic datasets and training a neural network for each unique imaging system must also be considered. Accepting that dataset creation is expensive, a brief comparison is made to the existing iterative optimization-based parametric inversion method from [8]. In practice, both these computationally expensive methods use a shared-resource computer running CentOS7 equipped with 1.5 TB of RAM and an NVIDIA Tesla V100 GPU. Timing for the optimization-based parametric inversion is given for a typical use case of 8 cores on the above computer. The timing for data generation is measured using (up to) 32 cores, with the network training being GPU-accelerated.

Current methods for parametric inversion are iterative, and scale linearly with the number of frequencies involved, so that the total computation time, t_{comp}, is:

$$t_{\text{comp}} = n_{\text{freq}}(t_{\text{inv}} * k) \tag{3.4}$$

where k is the number of iterations, t_{inv} is the time to perform inversion for a single iteration, and n_{freq} is the number of frequencies, as noted previously.

Parametric inversion must be applied to every measurement example to be evaluated; in the stored grain case, this may amount to tens of examples per day. A single frequency parametric inversion using the simplex method in [8] takes approximately 3 hours ($\pm$ 1 hour).

In contrast, the machine learning workflow relies on a large upfront cost for generating the labelled synthetic dataset. For any given chamber, a dataset need only be generated once at the appropriate frequencies, and can then be used repeatedly as long as the antenna positions within the chamber remain unchanged.

A single frequency dataset consisting of 10^4 examples can be generated in approximately 8 hours using the DGM forward solver. Packaging of an 8 frequency dataset can be completed in under 30 minutes and GPU-accelerated training of a six-hidden-layer network takes around 3 minutes (around 10 minutes without the use of GPU). Once the network is trained, the model can be easily moved to another computer for use. As a baseline, a 2013 MacBook Pro (2.4 GHz Intel Core i5 processor, 8GB RAM) can package and evaluate 14 experimental datasets in under 1 minute.

Table 3.1: Approximate time required for each step of the machine learning workflow.

Step	**Training time**
Data generation (per frequency)	8 hours
Data packaging	<30 minutes
Network training	3 minutes
Data testing	<1 minute

One drawback of the neural network compared to the iterative optimization is reduced flexibility. Although the network can be trained for any system, each system requires the dataset generation stage, making one-off measurements impractical using this approach. As the iterative optimization requires a forward solve at every iteration, there is no large upfront cost, which may make it more suitable for parametric inversion when the anticipated number of results to evaluate is small.

The iterative optimization-based inversion makes use of additional system information, either in the form of the results from a height scan of 15 height steps or the backscattering measurement, to provide an initial guess for the iterative inversion in an attempt to reduce the chances of the algorithm getting stuck in a local minimum. Although this is not a significant additional cost, it does require extra forward solves. In contrast, the neural network does not require an initial guess as it is not susceptible to such local minima in the same way, and can accurately predict the parameters directly from the uncalibrated phaseless data.

Chapter 4

Stored Grain Inventory Management and Monitoring

Material in this chapter is based off of and includes excerpts (Sections 4.3, 4.6.4, 4.6.5, 4.6.6, 4.6.7) from work previously published in IEEE Access [52].[1]

4.1 Introduction

Following harvest, grain is stored in large metallic bins until it is sold. Between harvest and sale, it is common for grain to be incrementally added to and removed from a bin. As grain is sold by weight and its value is sensitive to moisture content, accurate inventory management and monitoring for moisture content is desirable. Conventional methods for monitoring are dangerous and impractical, as they can involve physically entering the bin to take manual measurements using a tape measure to determine the height and estimate the cone angle of the grain. Alternative systems for inventory management of grain stored in large metallic grain bins can help reduce food waste, and maximize profit from the sale of the commodity. The EIL has previously reported an industrial electromagnetic grain bin monitoring system that can be used for inventory management and is capable of detecting regions of high moisture content [11]. Such regions may be indicative of grain spoilage, and their early detection is valuable to the farmer. Full 3D inversion and

[1] K. Edwards, N. Geddert, K. Krakalovich, R. Kruk, M. Asefi, J. LoVetri, C. Gilmore, and I. Jeffrey. Stored grain inventory management using neural-network-based parametric electromagnetic inversion. *IEEE Access*, 8:207182–207192, 2020.

image reconstruction of the measurement data from this system can be accomplished using CSI [11] or deep-learning [43], however both of these methods rely on knowledge of the bulk parameters of the grain within the bin. An iterative optimization method was originally developed to extract these bulk parameters from phaseless data [8]. This work presents a machine-learning-based approach to infer the bulk parameters of stored grain, and results are given for two different experimental systems.

The chapter is organized as follows: Sections 4.2 and 4.3 summarize the relevant features of the experimental grain bins and data collection system, Section 4.4 describes how the bins and grain-air interface are modelled, and Section 4.5 provides details of the neural network model, dataset creation, training, and testing. The model is assessed for experimental accuracy in Section 4.6.

4.2 Experimental Grain Bins

This research tested the performance of the neural networks on two different grain bins, whose properties are given in Table 4.1. Bin 1 is used for storing wheat and has a Hopper-style bottom, where the main (cylindrical portion) is 3.98 m in height with a diameter of 4.71 m. The effective bin volume is $\approx 90m^3$ ($\approx$ 2500 bushels). The bin has 24 transceivers arranged in staircases radially along the bin walls. Bin 2 has a flat bottom and a radius of 7.28 m. There is a sweep auger present on the floor of the bin (which is not accounted for in the computational model), the bin has an eave height of 5.24 m and an effective bin volume of $\approx 260m^3$ ($\approx$ 7400 bushels). In Bin 2, the 24 transceivers are located radially around the bin wall in pairs.

The two bin geometries are shown (not to scale) in Figure 4.1.

Table 4.1: Description of experimental grain bins evaluated. Frequency step sizes are given in parentheses.

Bin ID	Grain Type	Eave Height [m]	Bin Diameter [m]	Base Style	Frequencies [MHz]
1	Wheat	3.96	4.71	Hopper	[80-95] (5)
2	Corn	5.24	7.28	Flat	[65-86] (3)

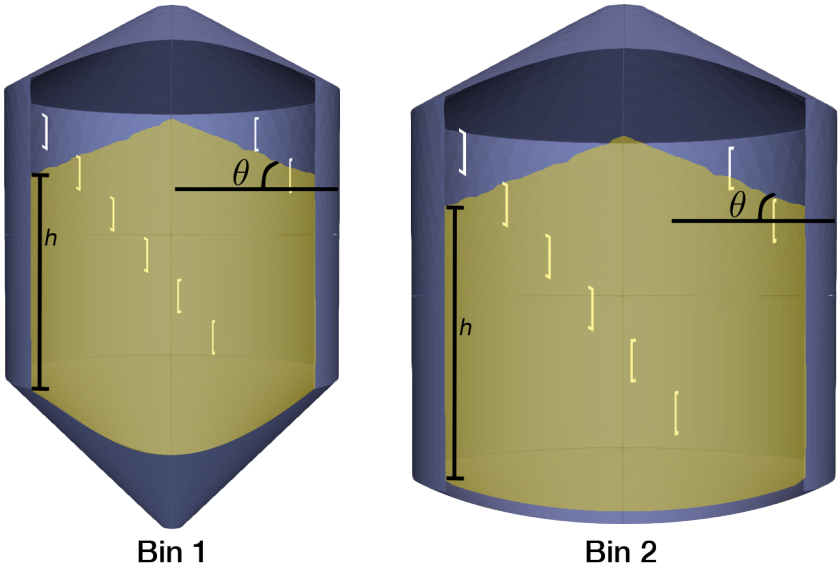


Figure 4.1: Cross-section render of a Hopper-style (left) and flat bottom (right) grain bin. Antennas are shown in white. The height, h, and grain angle of repose, θ, are marked.

4.3 Experimental Data Collection

The grain bin electromagnetic data acquisition system consists of 24 transceivers (a typical number) located at various positions (heights and angles) around the bin wall [11]. Each transceiver is connected, by coaxial cable that runs along the bin wall and through a switch, to a VNA as shown in Figure 4.2. The transceivers are designed to measure the surface current corresponding to the ϕ projection of the magnetic field on the bin wall, H_ϕ [36, 53]. The measurements used consist of VNA S-parameter data $S_{i,j} \in \mathbb{C}$ for all transmitter/receiver pairs (i, j), $i \neq j$. This amounts to an S-parameter matrix $S \in \mathbb{C}^{24\times 23}$ at a fixed frequency. We refer to this matrix simply as the S-matrix. Data is collected over a wide-range of frequencies (e.g. for the bins in this work, data is collected for frequencies from 10 MHz to 100 MHz). These frequencies are varied with bin size, with larger bins using lower frequencies. S-matrix data constitutes the raw electromagnetic measurement data used for bulk parameter estimation. Additional details of the data collection system can be found in [11, 20, 54, 55].

4.4 Modeling the Grain Bin

Each grain bin has a unique geometry and antenna installation. The bins are modelled using `gmsh` along with a collection of files that specify the antenna locations and polar-

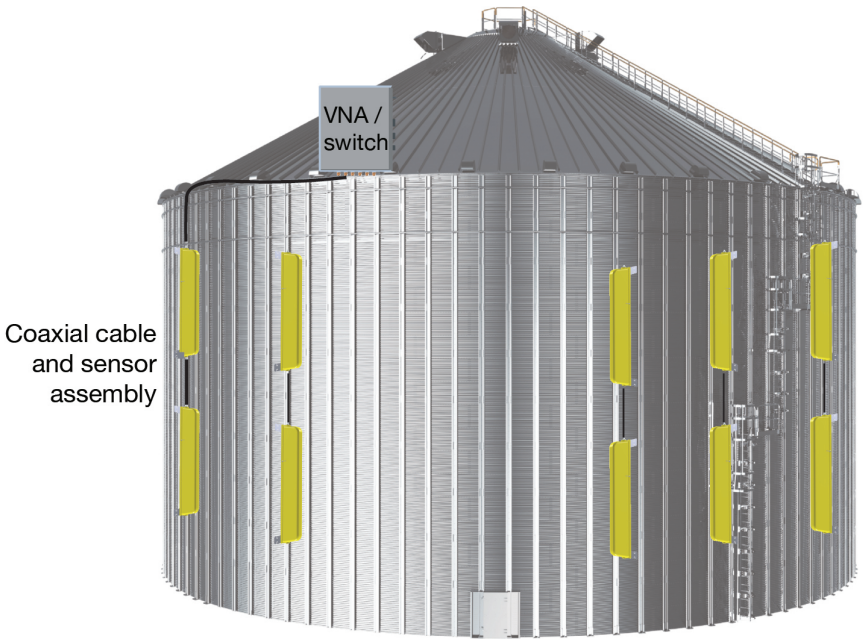


Figure 4.2: Data acquisition system installation. A transceiver and switching system (top of bin) is connected to the antennas inside the bin by coaxial cables. Not all antennas are shown.

izations, for use with the forward solver when generating synthetic data. The synthetic models of the bins do not model the antenna cables or other temperature cables in the system, or any augers that may be present at the base of bin. Synthetically, the antennas are modelled as point sources, though in practice they are shielded half-loop field probes [11, 53].

4.4.1 Grain-Air Interface

The bulk parametric inversion in stored grain aims to characterize the grain fill within the bin and to provide an estimate of the average moisture content of the grain. The bulk complex-valued permittivity parameters will provide information on the moisture content, while the remaining parameters will provide information about the geometry of the grain fill. These parameters can be thought of as the descriptors of the grain-air interface.

The grain-air interface can have several symmetric or asymmetric profiles, depending on where the grain is added to the bin and where it is removed from the bin. This thesis explores the characterization of three symmetric grain-air interfaces. These interfaces represent a top-center load point and a bottom-center draw, and are shown in Figure 4.3.

The fully peaked and fully inverted cone shapes were studied first, as they can be

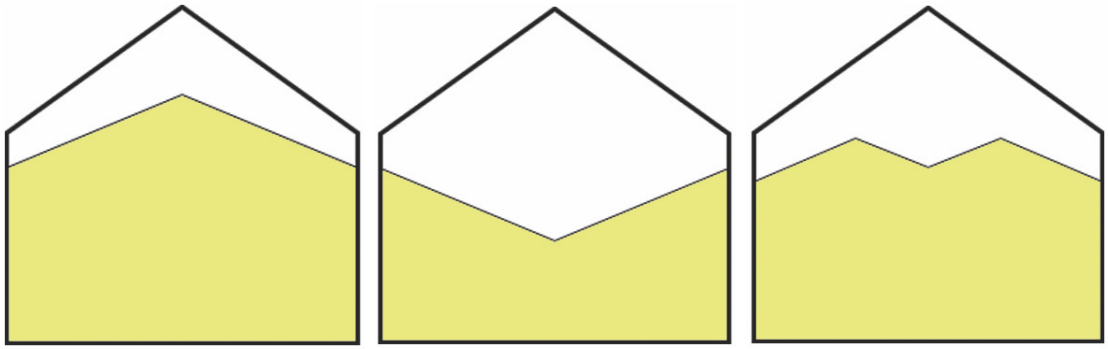


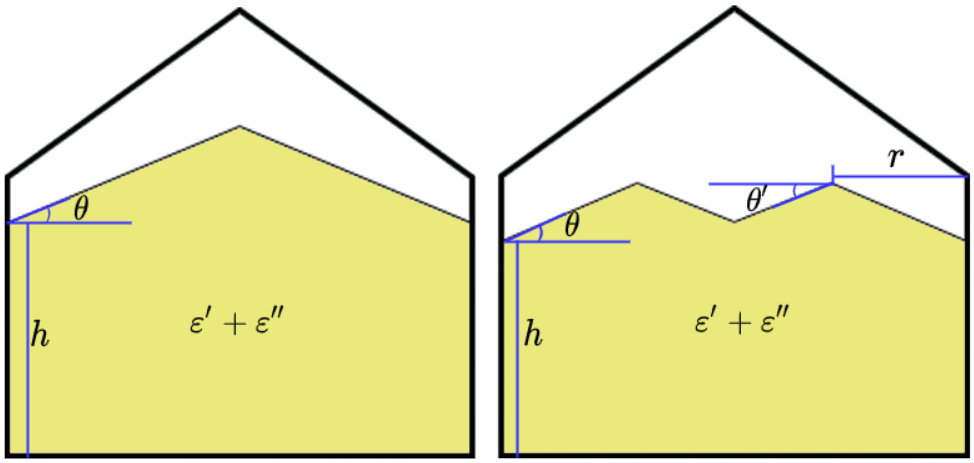
Figure 4.3: Fully peaked (a), fully inverted (b), and partially inverted (M-shape) (c) grain-air interfaces. ©2021 IEEE.

characterized by two parameters: grain height, and one of cone angle of repose or grain volume. The partially-inverted interface was the subject of subsequent experiments, where the parameter set was expanded from 4 parameters to 6 parameters. Neural networks restricted to fully peaked and fully inverted cones are referred to as $\underline{\mathbf{p}}_4$ networks, whereas those that are capable of characterizing partially inverted interfaces are referred to as $\underline{\mathbf{p}}_6$ networks.

4.4.2 Labelled Data

The simple grain-air interface of stored grain within a large metal grain bin can be parameterized by grain height (h), cone angle (θ), and bulk complex-permittivity ($\varepsilon = \varepsilon' + j\varepsilon''$, which accounts for two parameters, ε' and ε''). The grain volume can be inferred from the above parameters, or can optionally be used in place of or in addition to cone angle of repose. This work considers two 4-parameter sets for simple grain-air interfaces: $\underline{\mathbf{p}}_{4A}$, comprised of grain height, cone angle of repose, and bulk complex-valued permittivity, and $\underline{\mathbf{p}}_{4V}$, comprised of grain height, grain volume, and bulk complex-valued permittivity. Grain volume is of particular interest to grain handlers, and using volume in place of cone angle of repose may lead to more accurate results where the grain-air interface does not have a simple cone shape when using a simple (4-parameter) model. A single 6-parameter set expanding characterization capability to partially inverted cones is also evaluated: $\underline{\mathbf{p}}_6$ is comprised of grain height, cone angle of repose, inner angle of repose (θ'), peak radius (r), and bulk complex-valued permittivity. Figure 4.4 shows how stored grain is characterized by each of these parameters. It should be noted that when $\theta = -\theta'$

the 6-parameter interface simplifies to the equivalent 4-parameter full cone geometry.



(a) (b)

Figure 4.4: Visualization of the 4-parameter (a) and 6-parameter (b) labels relative to a grain bin.

When generating a synthetic data example, the bin geometry and antenna files are constant, but the grain fill within the bin is varied, resulting in the generation of a unique mesh file for each grain level and profile. Each mesh file is partitioned into two volumes: grain and air, where the grain volume is constrained by the bin walls and the grain-air interface as characterized by the parameter set, and the air volume fills the rest of the bin. A given mesh file can be used for data generation at multiple permittivities, where the bulk permittivity of the grain volume in the mesh file is passed as an input parameter to the forward solver. Figure 4.5 summarizes the steps for labelled data generation, highlighting the stage at which each parameter is incorporated.

As noted in Section 4.2, each bin configuration always consists of 24 transceivers at known locations along the bin wall, and while the bin dimensions and exact location of the transceivers within the bin are unique to each installation, the overall structure of the data remains consistent across bins.

4.5 Neural Network Architectures

The EIL's current optimization based approaches to parametric inversion process multi-frequency data one frequency at a time. The frequency whose predicted parameters are ultimately used for inversion is the frequency whose parameters give the highest correlation coefficient between the experimental data and the synthetic example generated for

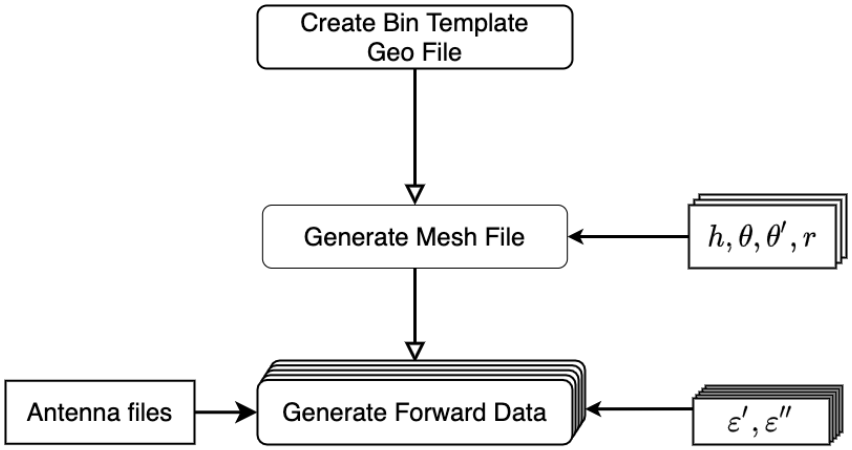


Figure 4.5: Synthetic labelled data generation workflow. The bin is modeled and a mesh file representing the bulk geometry parameters is generated using CAD software. This mesh file, along with information on the location of the antennas, and the specific bulk permittivity is passed to the forward solver to generate a synthetic data example for the specific parameter set, $\underline{\mathbf{p}}$. Note that θ' and r are only given and used for the $\underline{\mathbf{p}}_6$ network.

those parameters. However, experience has shown that incorporating data from multiple frequencies improves results. To provide a one-to-one comparison between the neural network and the iterative optimization-based inversion, the first neural networks developed were designed for single frequency input data.

The architecture of a single-frequency neural network is shown in Figure 4.6.

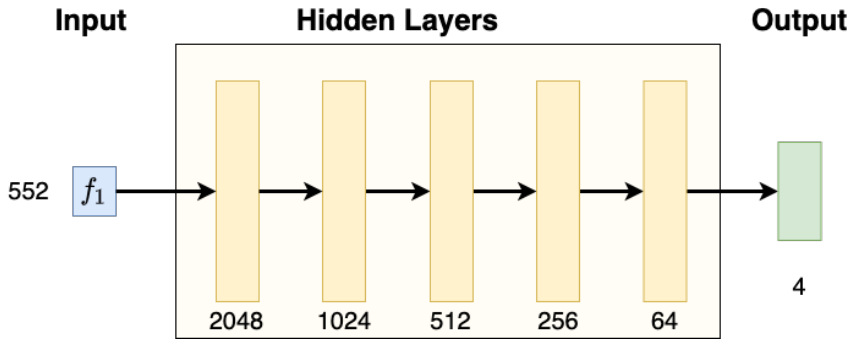


Figure 4.6: Fully-connected, 5-hidden layer neural network for a single frequency and four output parameters. The number of nodes in each layer are given below each layer.

A multi-frequency neural network was then developed to test whether incorporating data from multiple frequencies would improve the accuracy of the predicted parameters. To maintain a fully-connected network with a single channel input, data from multiple frequencies is concatenated to create a $552 * n_{\text{freq}}$ element column vector. To account for the increased length of the input vector, an additional hidden layer was added to the multi-frequency network architecture as shown in Figure 4.7. This architecture can, in

theory, be used for data from any number of frequencies (include single-frequency data); this work evaluates the network for up to 8 frequencies.

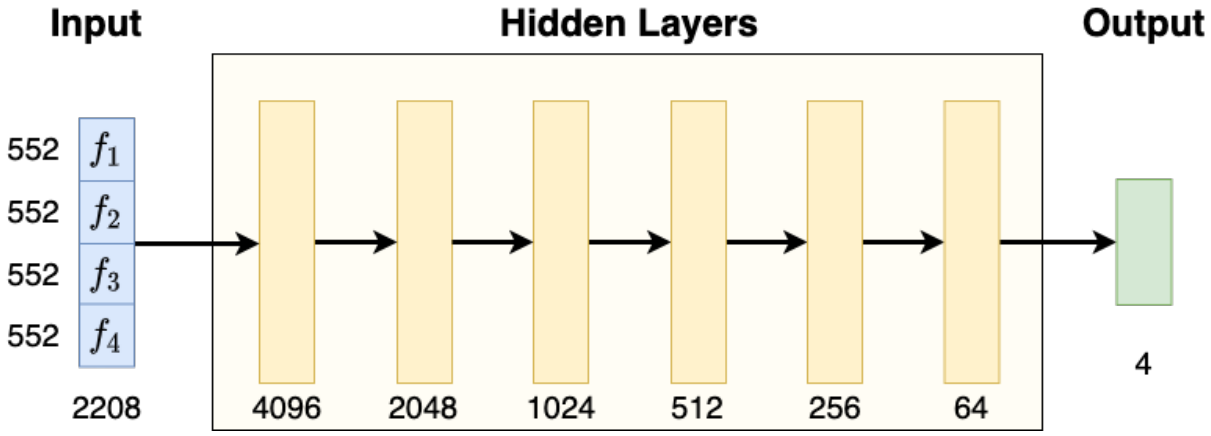


Figure 4.7: Fully-connected, 6-hidden layer neural network for $n_{\text{freq}} = 4$ frequencies and 4 output parameters.

This multi-frequency network architecture was further modified to evaluate the ability to characterize a 6-parameter model. Several network architectures, with varying numbers of input layers were evaluated. Of the architectures tested, a seven-hidden layer architecture (Figure 4.8) produced the best results.

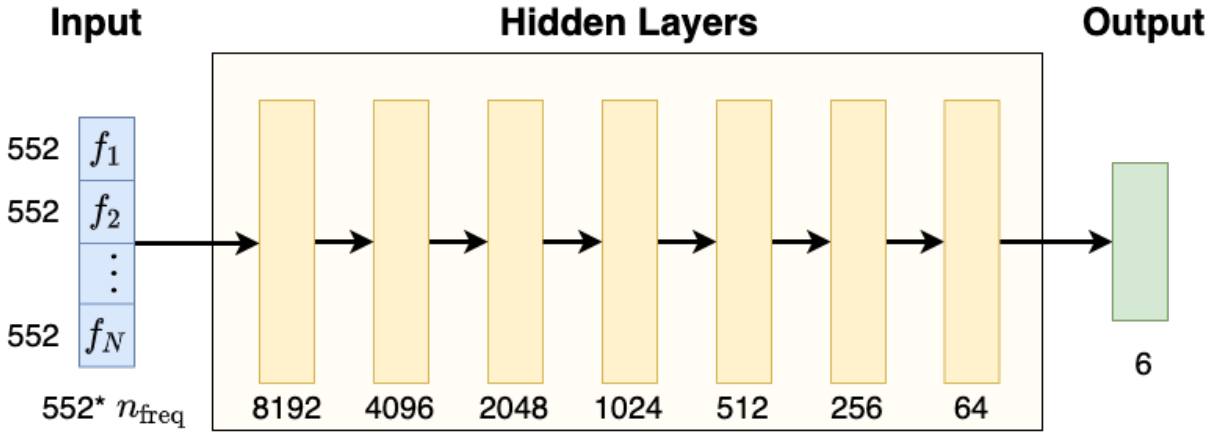


Figure 4.8: Fully-connected, 7-hidden layer neural network for an arbitrary number frequencies, n_{freq}, and 6 output parameters.

4.6 Results

The fully-connected neural network architectures described in Section 4.5 were synthetically trained for 4-parameter and 6-parameter predictions. The 4-parameter networks

were used to predict parameters ($\underline{\mathbf{p}}_{4A}$ and $\underline{\mathbf{p}}_{4V}$) for experimental data from the two bins described in 4.2. The 6-parameter network was used to predict parameters ($\underline{\mathbf{p}}_{6}$) for noisy synthetic data generated for Bin 2.

4.6.1 Experimental Test Data and Ground Truth

Results are presented for the two different grain bins with different stored commodities as outlined in Section 4.2.

The selected frequencies used to train and evaluate the neural network are based on the EIL's experience with parametric inversions. For Bin 1, inversions are most successful in the 90–95 MHz range; this work evaluates data from 70–105 MHz for the neural network. This frequency range corresponds to approximately 1.1–1.65 wavelengths across the bin diameter in free space, or 2.2–3.3 wavelengths in grain (exact number of wavelengths varies with commodity). Chosen wavelengths in microwave imaging are typically around 2-3 wavelengths to achieve a reasonable resolution while considering the non-linearity of the problem and the computational cost of 3D electromagnetic simulations in the bin. For Bin 2, frequencies in the 65–86 MHz range were taken in 3-MHz steps, giving approximately 1.6–2 wavelengths across the bin diameter in free space, or 3.2–4 wavelengths in grain.

Experimental labelled data from Bin 1 consists of 14 samples taken from hard red winter wheat, each at various heights and cone angles. The average moisture content (MC) was measured to be 13.6% using a Dickey-John GAC2100. The bulk density was measured to be 0.80 and 0.85 g/cm^3. Grain height measurements were taken manually and ranged from nearly empty (0.130 m) to completely full (3.52 m), and the measured cone angles ranged from $-23.7°$ to $23.6°$. The Bin 2 labelled dataset contains 46 experimental measurements for corn. The average moisture content of the corn was between 13.0% and 13.6% and the density of the corn was measured to be between 0.75 and 0.79 g/cm^3. Table 4.2 provides a description of the labelled experimental datasets for Bins 1 and 2.

Table 4.2: Summary of labelled experimental stored grain data sets.

Bin	# samples	Height range [m]	Angle range [°]	MC avg [%]
1	14	[0.130, 3.52]	[-23.7, 23.6]	13.6
2	46	[0.000, 4.46]	[-38.4, 29.0]	[13.0, 13.6]

4.6.2 Labelled Synthetic Dataset

4.6.2.1 $\underline{\mathbf{p}}_4$ Data

For Bin 1 and Bin 2, meshes were generated for the corresponding bin geometry for various heights, from 0 m to the eave height of the bin, and cone angles, from $-30°$ to $30°$, to cover the range of feasible fully inverted and fully peaked grain-air interfaces. For $\underline{\mathbf{p}}_4$ datasets, permittivity combinations from 3.00 to 5.00 (real) and -0.600 and -0.200 (imaginary) in step sizes of 0.167 and 0.0333, respectively, were allowed.

4.6.2.2 $\underline{\mathbf{p}}_6$ Data

A $\underline{\mathbf{p}}_6$ dataset was generated for Bin 2, where mesh files were created for various geometries modelling the more complex partially inverted grain-air interface. The values of θ and θ' were permitted to be any combination of positive and negative angles selected from the set, where $\theta \neq \theta'$ gives a partially inverted cone and for $\theta = -\theta'$: $\theta > 0$ corresponds to a fully peaked cone and $\theta < 0$ corresponds to a fully inverted cone. To simplify the search space for the proof-of-concept $\underline{\mathbf{p}}_6$ experiments, the permittivity combinations were reduced. The real part of the complex-valued permittivity was either 4.2 or 4.6, and the imaginary part of the permittivity was the same as for the $\underline{\mathbf{p}}_4$ case.

In all cases, the DGM forward solver was called on each mesh multiple times, once for each complex-valued permittivity. Table 4.3 provides an overview of the discrete parameter-space sampling for each of the three training sets discussed in this work. Note that for Bin 2, samples were only generated for negative cone angles where the grain height is such that the inverted cone does not extend below the bin floor.

Table 4.3: Summary of synthetic stored grain data sets; ranges in square brackets, step sizes in parentheses. Extended from Table 3 in [52].

Type	Bin	Height range [m]	Angle range [°]	ε' range	ε'' range	M-Angle range [°]	M-radius range [m]
$\underline{\mathbf{p}}_4$	1	[0, 3.96] (0.330)	[-30.0, 30.0] (5)	[3.00, 5.00] (0.167)	[-0.600, -0.200] (-0.0333)	-	-
$\underline{\mathbf{p}}_4$	2	[0, 5.25] (0.350)					
$\underline{\mathbf{p}}_6$	2			[4.2, 4.6] (0.4)		[-30.0, 30.0] (5)	[0, 3.64] (1.21)

4.6.3 Single- vs. Multi-frequency Networks

This section is based on work presented at URSI GASS 2020 [56]. [2]

The first neural network model (Figure 4.6) evaluated is a 4-hidden layer, single-frequency network trained for Bin 1. In an attempt to improve the results from this network, several architectures were evaluated for multi-frequency measurement data. As previously described, the method for incorporating multi-frequency data is concatenation of frequencies into a single input vector. This concatenation results in a larger input layer, as data from additional frequencies is included. To account for the increased size of the input layer, different multi-frequency networks have additional hidden layers. Table 4.4 describes the architecture of three neural networks tested on the 14 experimental examples from Bin 1.

Table 4.4: First evaluated neural network architectures. ©2020 IEEE.

Architecture	Number of frequencies	Hidden layers	Output size
NN-1	1	4	4
NN-2	3	5	4
NN-3	6	6	4

From these three network architectures, 5 different networks were trained. The simplex method inversion of predicting parameters has shown good results for 70–90 MHz data, so frequencies were chosen from near this range for evaluation. The average error in height and cone angle predictions for each of the 5 neural networks tested, as well as for two single-frequency simplex method tests are given in Table 4.5. Figure 4.9 plots the predictions for each of the neural networks and the two single-frequency simplex method results. The x-axis for the height and cone angle graphs are independently sorted by increasing true value of the parameter (y-axis).

These results show that the 6-frequency neural network has the best performance among the neural networks. This is likely due to a combination of incorporating data from multiple frequencies and adding depth to the neural network. The neural network achieves more accurate cone angle predictions than the simplex method, but the simplex method performs slightly better on height predictions. It is worth noting that the simplex method

[2] K. Edwards, K. Krakalovich, R. Kruk, V. Khoshdel, J. LoVetri, C. Gilmore, andI. Jeffrey, "The implementation of neural networks for phaseless parametric inversion," in *2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science.* ©IEEE, 2020.

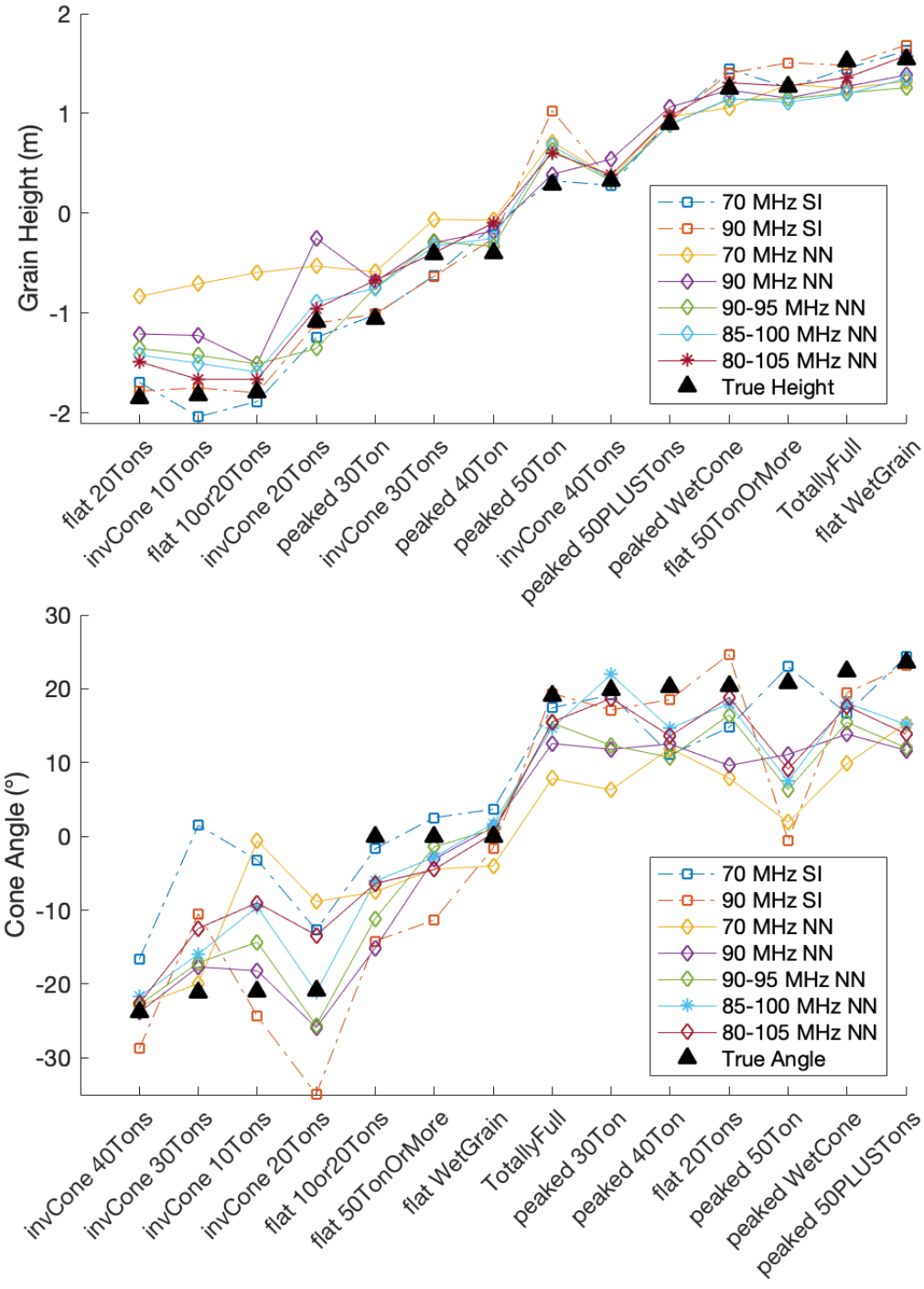


Figure 4.9: Plots of inversion results (neural network and simplex) for the predicted height (top) and predicted cone angle (bottom). In each plot the neural network with lowest error is marked with an asterisk for emphasis. ©2020 IEEE.

Table 4.5: Average absolute error for height and cone angle with each parametric inversion obtained from the first neural network architectures evaluated. ©2020 IEEE.

Frequency (MHz)	Method	Height (m) Error	Cone angle (°) Error
Neural Networks			
70	NN-1	0.4467	9.695
90	NN-1	0.2915	6.716
90-95	NN-2	0.2244	6.258
85-100	NN-3	0.2042	5.010
80-105	NN-3	0.1560	5.690
Simplex method			
70	SI	0.1175	6.388
90	SI	0.1384	6.698

uses additional prior information (in the form of back-scattering, $S_{i,i}$, measurements) to determine the approximate height of the grain prior to starting the inversion.

As the exact permittivity of the grain in Bin 1 at the time the measurements were taken is not known, the permittivity predictions are compared to the simplex method predictions to determine whether the predicted values are in a reasonable range. Permittivity predictions from 70 MHz were used for comparison as on average the simplex method performed better at 70 MHz than at 90 MHz (Table 4.5). Across the 14 experimental data measurements the simplex method at 70 MHz predicted an average real permittivity of 4.112 ($\sigma = 0.223$) and average imaginary permittivity of -0.572 ($\sigma = 0.0984$). The average real permittivity prediction for the neural networks ranged from 4.169 to 4.251 ($\sigma_{max} = 0.219$) and the average imaginary permittivity prediction ranged from -0.502 to -0.435 ($\sigma_{max} = 0.0690$), showing reasonable agreement between the two methods. Further discussion on permittivity prediction accuracy is reserved for later sections.

4.6.4 Neural Network Training

Following the initial investigation of different network architectures, a multi-frequency network architecture was trained for Bin 1 using 4 frequencies, and Bin 2 using 8 frequencies (selected from the optimal range for each bin based on prior experience). Each network was trained with a batch size of 1000 for up to 100 epochs. An early stopping condition was added such that if there is no improvement in the loss for 15 epochs the training stops and uses the weights from the epoch with the lowest loss. Table 4.6 com-

pares the network parameters and training conditions for the $\underline{\mathbf{p}}_4$ networks. Note that although the number of inputs doubles for Bin 2 (compared to Bin 1), the number and size of the hidden layers remains unchanged.

Table 4.6: $\underline{\mathbf{p}}_4$ neural network architectures and training characteristics.

Bin	Number of frequencies	Number of inputs	Number of layers	Number of trainable parameters	Batch size
1	4	2280	6	20,209,732	1000
2	8	4416	6	29,253,700	1000

4.6.5 Height, Cone Angle, and Volume Prediction Accuracy

Once trained on synthetic data, the neural networks were tested using experimental data. Table 4.7 shows the average absolute error of the height and cone angle predictions relative to the measured (experimental) values for the DGM neural network (NN) for Bin 1 compared to two single-frequency simplex inversion (SI) at 70 MHz and 90 MHz. The 4-parameter inversion was run once for $\underline{\mathbf{p}}_{4A}$ (angle) and once for $\underline{\mathbf{p}}_{4V}$ (volume), demonstrating the configurability of the neural network. As volume is a function of grain height, cone angle, and bin geometry, any one of the three parameters of interest (height, cone angle, volume) can be calculated from the other two, thus we can predict all 5 parameters without changing the 4-parameter architecture. Italicized values in Table 4.7 are derived from the other two predictions for that NN.

For both bins, the height prediction accuracy of the NN is close to, or better than, the height discretization accuracy used to generate the synthetic training samples. The error in cone angle predictions is slightly lower using the proposed NN compared to SI. We note that as the bin is filled, the cone contributes proportionally less to the overall volume of grain. Consequently, for significantly filled bins, inaccuracies in height will arguably have a larger effect on volume prediction than (slightly) erroneous angle predictions. It is not surprising, therefore, that the simplex method, which produces better height predictions, also produces more accurate volume predictions, an effect that, as expected, is more pronounced for fuller bins.

The results presented from the neural network are for the same experimental data with no pre- or post-processing steps to refine the results (other than scaling for the

neural network as mentioned above). It should be noted that some preprocessing and prior information is used to avoid local minima in the SI process [8], whereas none is used for the neural networks.

Table 4.7: Average absolute error for height and cone angle with each parametric inversion for Bin 1 and Bin 2. Results are given for simplex inversion, as well as for $\underline{\mathbf{p}}_{4A}$ and $\underline{\mathbf{p}}_{4V}$ neural network predictions. Derived quantities are italicized.

Frequency [MHz]	Method	Height [m] Error	Cone angle [°] Error	Volume[1] [%] Error
Bin 1				
80-95	NN $\underline{\mathbf{p}}_{4A}$	0.285	5.45	*6.06*
80-95	NN $\underline{\mathbf{p}}_{4V}$	0.208	*5.74*	4.17
70	SI	0.117	6.39	*1.34*
90	SI	0.138	6.70	*2.25*
Bin 2				
65-86	NN $\underline{\mathbf{p}}_{4A}$	0.406	7.21	*7.33*
65-86	NN $\underline{\mathbf{p}}_{4V}$	0.352	*7.53*	5.78
Best	SI	0.234	8.50	4.52
Bin 2: True Height >0.5 m				
65-86	NN $\underline{\mathbf{p}}_{4A}$	0.193	6.37	*4.48*
65-86	NN $\underline{\mathbf{p}}_{4V}$	0.175	*6.90*	3.33
Best	SI	0.0990	7.18	3.03

[1]Volume average absolute error reported as a percentage of the full bin.

Fig. 4.10 and Fig. 4.11 show that the error in cone angle predictions is greater and less consistent than the error in height and volume predictions, motivating the continued use of the $\underline{\mathbf{p}}_{4V}$ networks. The error in predictions for the empty or nearly empty bin is significantly larger than for samples where the true height is at least 0.5 m full. When the bin is nearly empty, the impacts of using uncalibrated data are most pronounced. In a near-empty bin, without loss through the bulk grain volume, the effects of multiple reflections and modelling error are much more significant. Specifically, the presence of unmodelled augers and ladders contribute larger effects in empty bins. For this reason, it is expected that the parametric inversion for uncalibrated data will not perform as well for empty or nearly empty bins given our current models, but may be improved by refined modelling in the future. The bottom section of Table 4.7 shows the prediction accuracy for Bin 2 limited to experimental samples where the true height is greater than 0.5 m. These results demonstrate that the neural network performs significantly better than the height discretization step size of the synthetic data when the bin is at

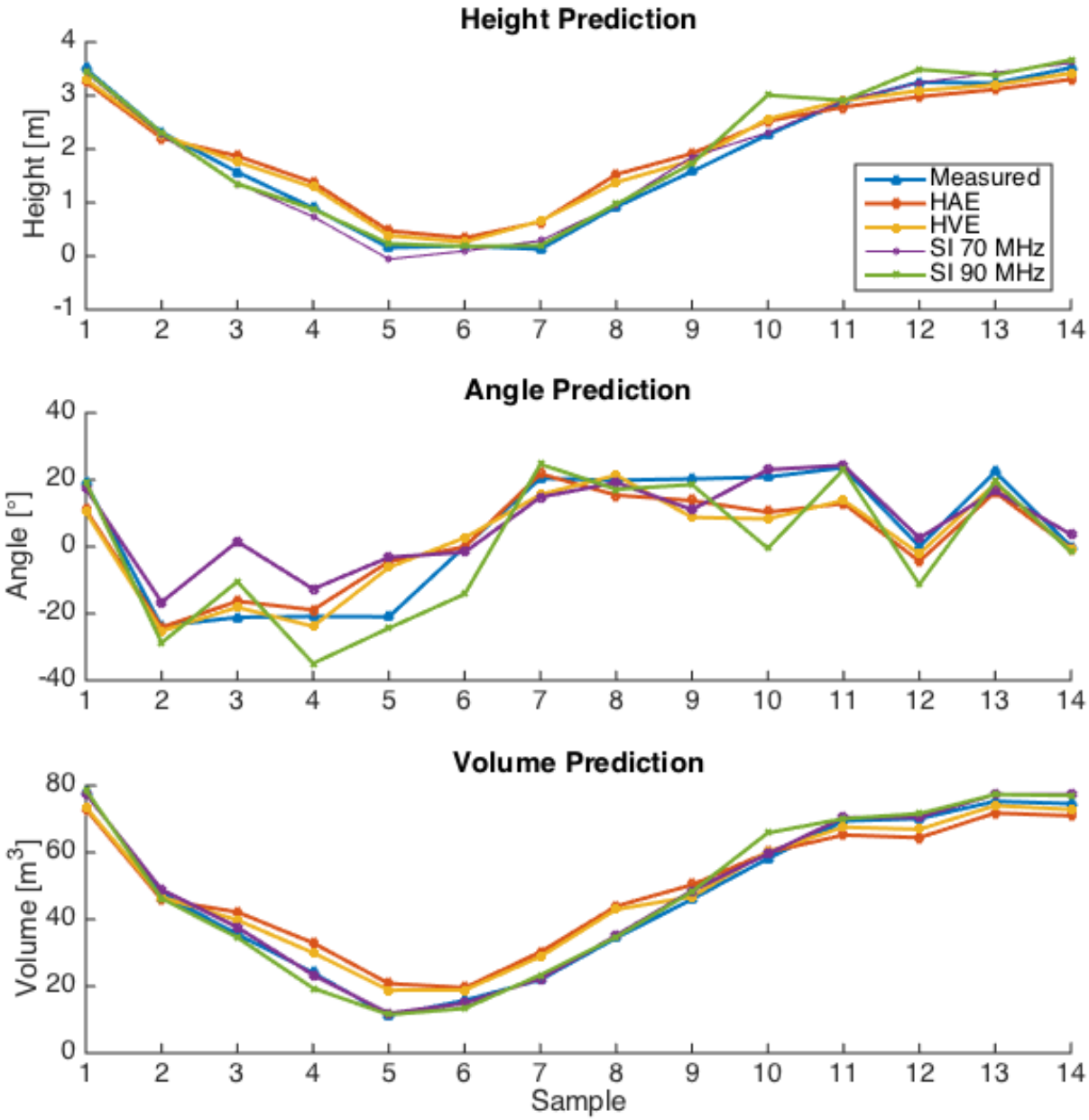


Figure 4.10: Experimental results (neural network and simplex) for the predicted height (top), predicted cone angle (middle), and predicted volume (bottom) for Bin 1 (Hopper-style) as a function of sample number.

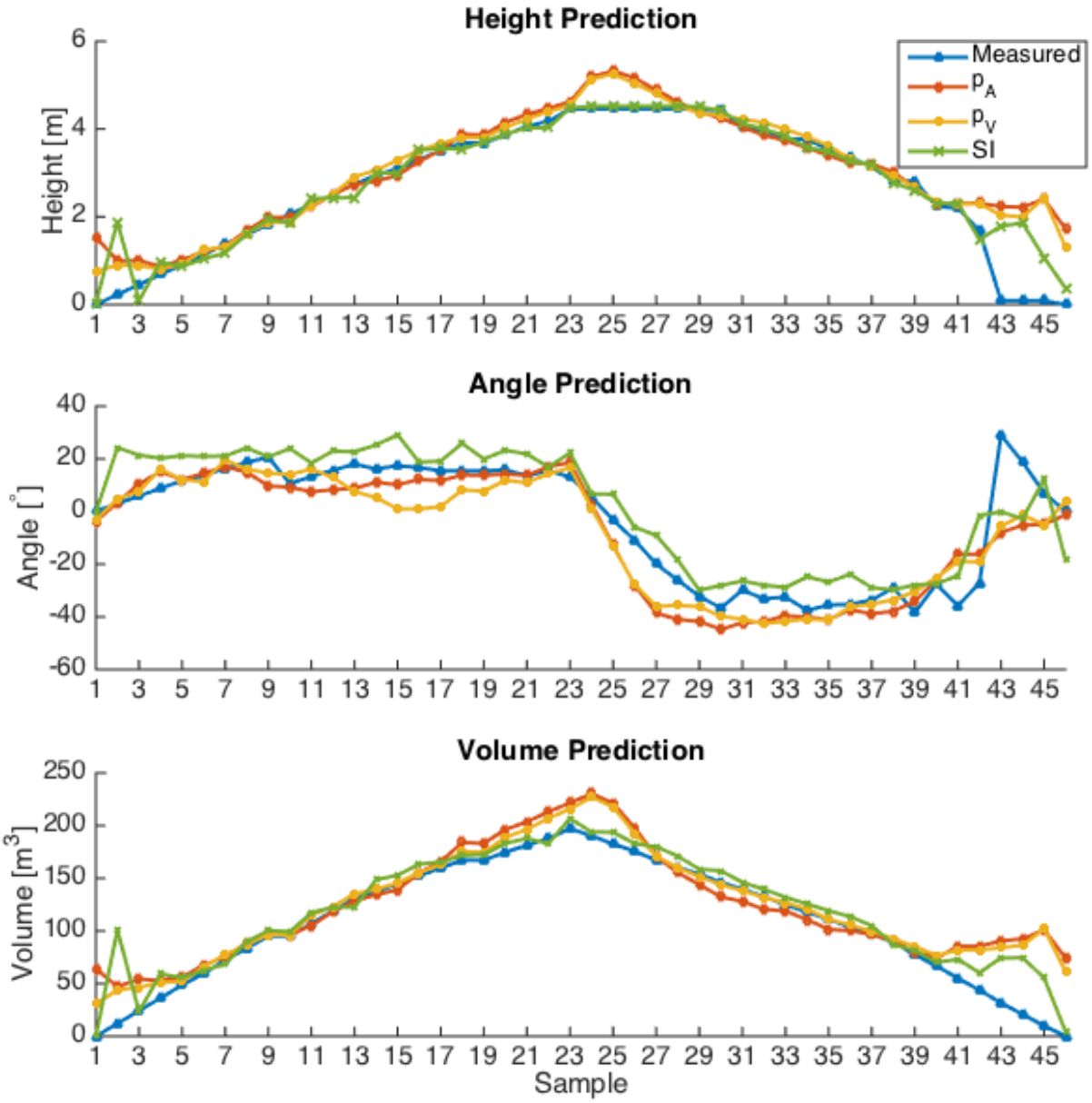


Figure 4.11: Experimental results (neural network and simplex) for the predicted height (top), predicted cone angle (middle), and predicted volume (bottom) for Bin 2 (flat-bottom) as a function of sample number.

least 0.5 m full. Additionally, for Bin 2 the neural network overestimates the height and volume parameters when the bin is transitioning from being filled to being emptied. The experimental samples for Bin 2 proceed from an empty bin to a full bin with an upward cone then back to an empty bin. This includes the transition from an upward cone, through a partially inverted cone [57], into a downwards cone. This transition highlights a limitation of our simple four parameter model that does not account for partially inverted cones and suggests the need to include more diverse surface models in the neural network training set. Note that the data samples from Bin 1 did not include any partially inverted cone configurations.

The tolerance plots in Fig. 4.12 provide a visual representation of the network performance for the $\underline{\mathbf{p}}_{4V}$ networks, consequently the angle tolerance is based on the derived angle prediction. For a fixed tolerance on the horizontal axis, the vertical axis counts the number of test sample predictions (as a percentage of the total test samples) that fall within the tolerance of the measured value. The tolerance is defined as the percent error relative to the measured value. Note that errors for measured values near zero are calculated relative to the associated discretization step sizes summarized in Table 4.3.

From the figure it is clear that height, volume, and cone angle each contribute to the overall prediction error, and also that inference struggles most with determining cone angle. This result is not surprising as features of the measured data, including transceiver loading, may be most prominent when transceivers change from being in grain to out of grain; a change in angle does relatively little to affect the raw data.

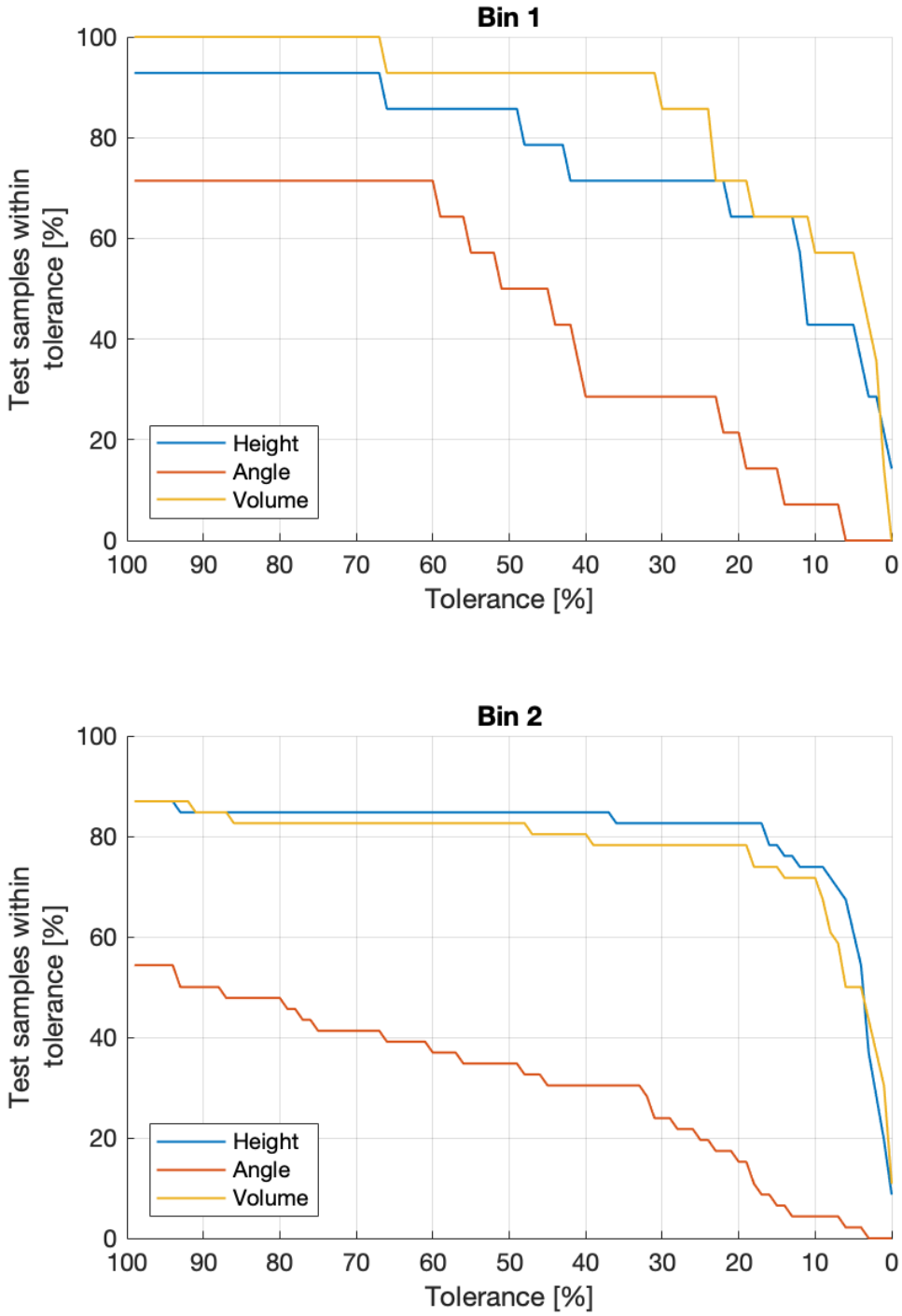


Figure 4.12: Tolerance plots for experimental test data ($\underline{\mathbf{p}}_{4V}$ network) for Bin 1 (top) and Bin 2 (bottom). Height, cone angle, and volume tolerances are shown in blue, red, and yellow, respectively. The percentage of test sample predictions (y-axis) within the percent tolerance of the measured data (x-axis).

4.6.6 Permittivity Prediction Accuracy

True values for the bulk permittivities ($\varepsilon = \varepsilon' + j\varepsilon''$) of the grain are not available for the Bin 1 experimental data set, posing a challenge for quantitative validation of the permittivity results. Instead, we compare the neural network predictions to those obtained through simplex inversion to provide preliminary validation. The simplex method (at 70 MHz) predicted an average real permittivity of 4.11, with a standard deviation of 0.223, and an average imaginary permittivity of -0.572 with a standard deviation of 0.0984 over the 14 experimental samples at various heights and cone angles. The neural networks tested predicted the average bulk real component of the grain permittivity between 4.17 and 4.25, with a maximum standard deviation of 0.219, and the average bulk imaginary component between -0.502 and -0.435, with a maximum standard deviation of 0.0690. Although not a conclusive demonstration of absolute accuracy, the neural network predicted values are similar to those obtained by the simplex method; they could be used for bulk inventory quantification, or for calibration and as prior information for subsequent imaging refinements. Furthermore, we compared the predicted permittivities with Nelson and Kraszewski's composite model of the complex permittivity of cereal grain [21] at a moisture content of 13.6% with bulk densities ranging from 0.80 to 0.85 g/cm^3 at 90 MHz. The model predicts a range of bulk real permittivities between 3.90 and 4.13 and bulk imaginary permittivities between -0.38 and -0.42, which further supports that the neural network provides reasonable predictions of both real and imaginary bulk permittivities of the grain.

For Bin 2, when excluding the predictions where the true height is less than 0.5 m full, the bulk real permittivity predictions range from 4.29 to 4.77 and imaginary permittivity predictions range from -0.235 to -0.436. Based on Nelson and Kraszewski's composite model, we expect bulk real permittivities to range from 3.56 to 4, and bulk imaginary permittivities to range from -0.32 to -0.39. We do note that the composite model is at room temperature and the outdoor bin was at a higher temperature. This temperature difference contributes to the differences seen. However, it also suggests that there is a need for improvement in predicting bulk permittivity of the grain for some bins and/or commodities. We are currently investigating ways to improve bulk permittivity predictions, as they are an important parameter for determining moisture content.

4.6.7 Performance on Synthetic Test Sets

While the previous results validate the primary objective of this work, namely to apply a synthetically trained network to uncalibrated experimental data for stored grain monitoring, there is value in assessing the performance of the synthetically trained network on synthetic data. To assess the suitability of our neural network architecture for the multi-frequency parametric inversion problem, we visualize tolerance plots for a synthetic test set for each bin. When comparing the tolerance plot for the synthetic test sets to the tolerance plot for the experimental test sets (Fig. 4.13), we observe that the network is capable of fitting nearly all of the synthetic test examples within 10% error of the true values. This suggests that much of the error in predicting parameters from experimental data can be attributed to how well the synthetic dataset represents the experimental dataset, and not necessarily to how well suited the neural network is to the multi-frequency parametric inversion problem. This result not only validates the suitability of this network architecture for the parametric inversion task, but also highlights the limitations of running an experimental test set on a synthetically trained neural network.

4.6.8 Complex Grain-Air Interfaces

This section is based on work that has been accepted to ANTEM 2021. [3]

The $\underline{\mathbf{p}}_4$ neural network results from the previous section are encouraging, and suggest that this neural network based approach may be useful in commercial stored-grain monitoring, however there is still room for improvement. It is well known that the $\underline{\mathbf{p}}_4$ parameter set is an oversimplification of the many grain-air interfaces that may exist in a bin throughout the loading and unloading cycle [57]. The $\underline{\mathbf{p}}_6$ parameter set that supports characterization of partially inverted cones is one option for extension of the $\underline{\mathbf{p}}_4$ set. When evaluating the results from the $\underline{\mathbf{p}}_4$ network, Figure 4.11 shows that when the bin is transitioning from being filled to being emptied the height and volume predictions are overestimated and the error in the angle predictions increases. For a center draw bin such as Bin 2, a partially inverted cone would be expected as the bin is being emptied.

The $\underline{\mathbf{p}}_6$ neural network described in Section 4.5 was trained for Bin 2. The M-shape parameters defined are not available as labels on the Bin 2 experimental dataset, and

[3] K. Edwards, J. LoVetri, C. Gilmore, and I. Jeffrey, "A machine learning method for characterization of complex grain air interfaces in grain storage bins," *Accepted to ANTEM 2021*, 2021

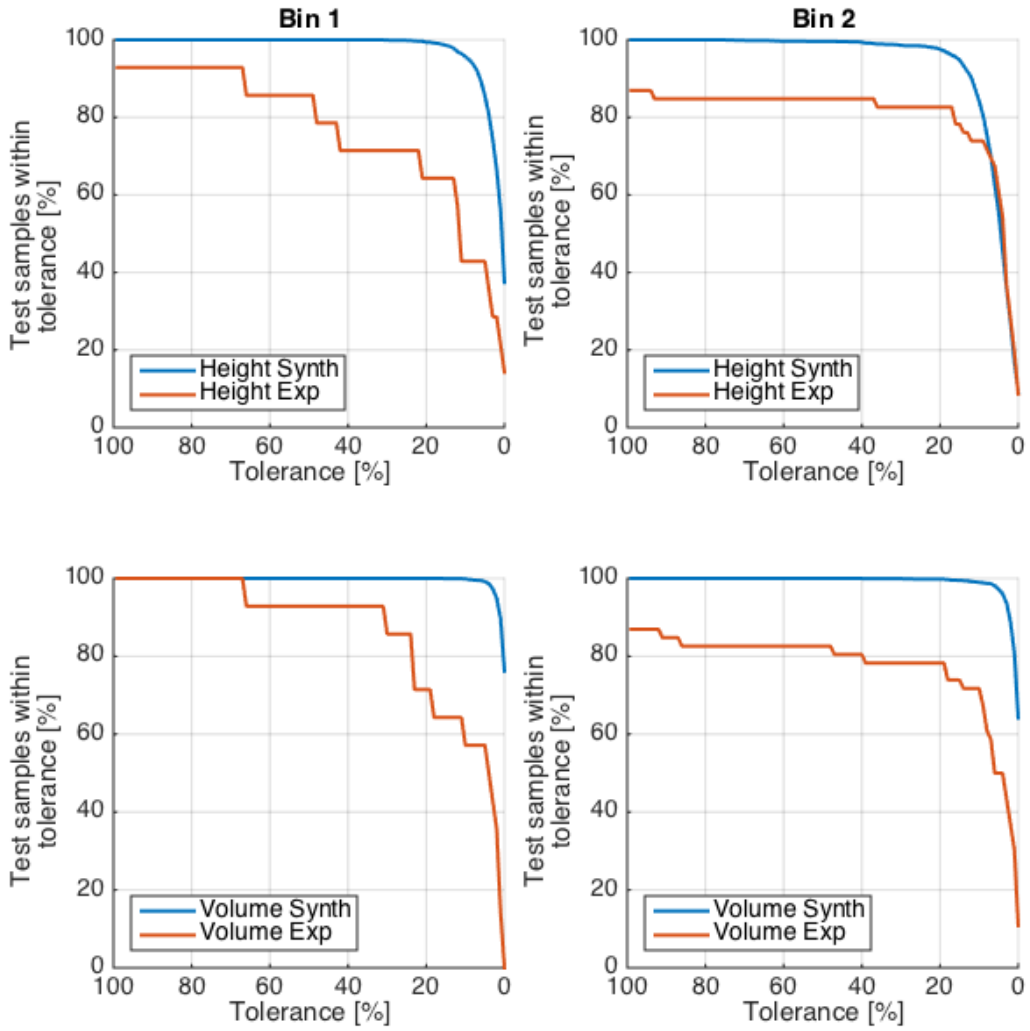


Figure 4.13: Tolerance plots for both synthetic and experimental test sets ($\underline{\mathbf{p}}_{4V}$ networks), height plots are shown on top and volume plots on bottom. Left: Bin 1 (14 samples) (red) and synthetic test set (1429 samples) (blue). Right: Bin 2 (46 samples) (red) and synthetic test set (1573 samples) (blue).

so a noisy synthetic test set for two different levels of random white Gaussian noise was created to investigate the feasibility of a $\underline{\mathbf{p}}_6$ neural network. Table 4.8 gives the error in each of the 6 parameters for both levels of noise.

Table 4.8: Average absolute error for $\underline{\mathbf{p}}_6$ predictions for noisy-synthetic test data.

Parameter	**10^{-2} Noise**	**10^{-4} Noise**
Height error [cm]	2.50	2.41
Angle error [°]	0.542	0.506
M-Angle error [°]	1.31	1.11
Peak Radius error [cm]	3.92	3.10
ε' error	6.40×10^{-3}	5.77×10^{-3}
ε'' error	3.30×10^{-3}	3.32×10^{-3}

To visualize how the $\underline{\mathbf{p}}_6$ parameter set can characterize the grain-air interface and show the error in predictions, the results from 5 different test examples are visualized in Figure 4.14. Table 4.8 and Figure 4.14 show that the $\underline{\mathbf{p}}_6$ network is able to successfully reconstruct complex grain-air interfaces from noisy synthetic data representing different fill levels and grain profiles within the bin.

4.7 Chapter Summary

It has been shown that a bulk parametric inversion neural network trained solely on synthetic data can be used to evaluate uncalibrated, phaseless experimental data obtained from an electromagnetic stored grain imaging system. In this work, two different grain bins were evaluated, having different geometries, and storing different commodities. The ability of the network to infer bulk parameters from both of these systems demonstrates that the workflow presented can be generalized as long as a computational model of the system can be created.

The extension of the parameter set from $\underline{\mathbf{p}}_4$ to include additional geometry parameters ($\underline{\mathbf{p}}_6$) gives encouraging results when evaluated on noisy synthetic data. The fact that the workflow can predict more complex parameter sets, suggests an opportunity for future investigation into the choice and complexity of the bulk parameter sets, providing additional capabilities for obtaining prior information for inversion.

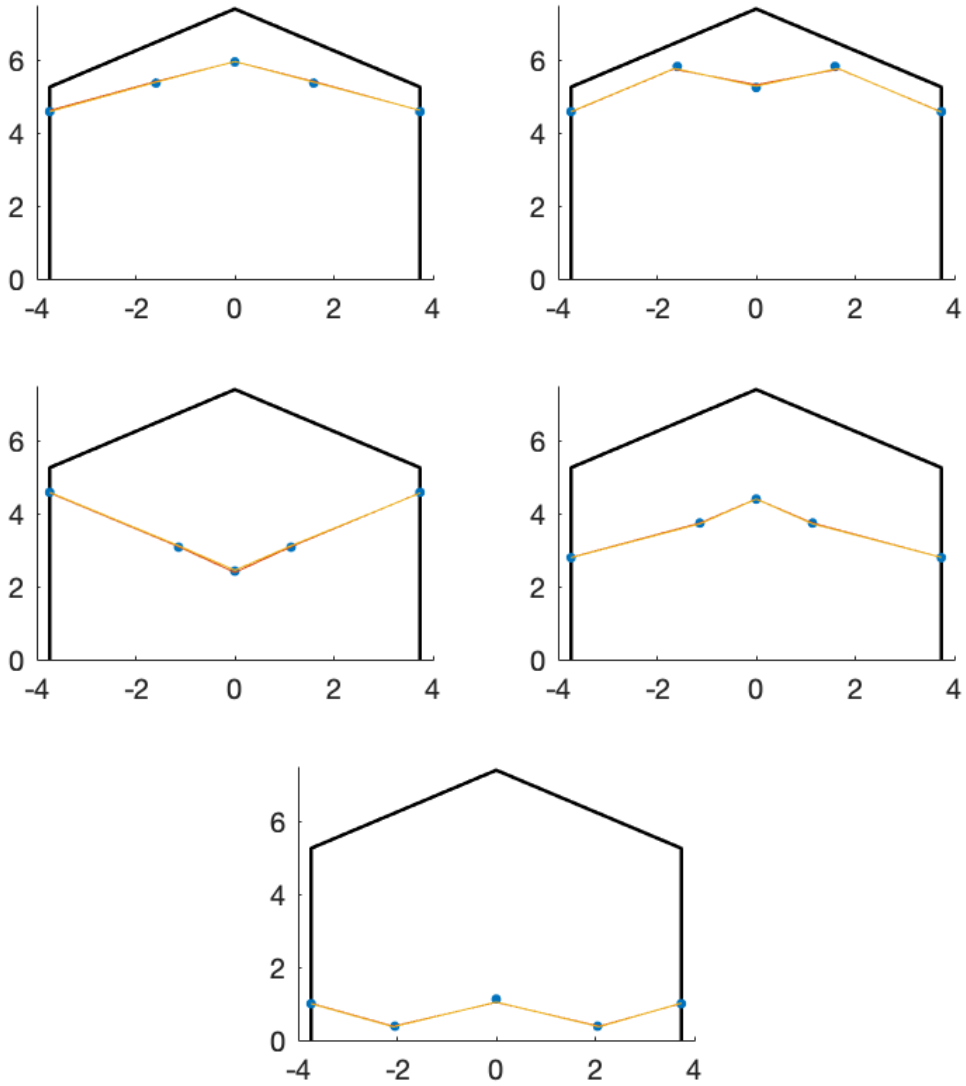


Figure 4.14: Predicted profile of the grain-air interface from $\underline{\mathbf{p}}_6$ networks for 5 different noisy-synthetic test examples. True values are marked as blue points. All axes are in meters. ©2021 IEEE

Chapter 5

Microwave Imaging of a Human Breast Model

Material in this chapter has been adapted from [58][1], a paper published in MDPI's Electronics Journal and included in the Special Issue on Microwave and Wireless Communications: Photonic and Microwave Sensing Developments and Applications.

5.1 Introduction

This chapter demonstrates the applicability of the machine learning framework to the microwave breast imaging problem. Microwave imaging using inverse scattering-based techniques is actively being studied as a technique for breast cancer detction [7, 59–61]. At microwave frequencies, different tissue types within the breast can generally be distinguished based on their complex-valued permittivity. This work uses a model of the human breast made up of adipose and fibroglandular tissues; small tumors can be introduced into the model, having a complex-valued permittivity that is higher than either of the other two tissues [22]. In the stored grain application there is a single bulk permittivity region, and a region of high moisture content is the analog of a tumor, having a higher permittivity than the bulk of the grain. In contrast, the breast imaging application consists of two separate tissue types, representing adipose and fibroglandular

[1] All articles published by MDPI are made available worldwide under an open access license, meaning no special permission is required to reuse all or parts of articles published by MDPI. Reprinted with permission from K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, and I. Jeffrey, "A Machine Learning Workflow for Tumour Detection in Breasts using 3D Microwave Imaging" *Electronics* 2021, 10, 674.

tissues, in addition to the high permittivity tumor tissue. The parameter set for the breast imaging work describes the geometry and permittivity of the fibroglandular region of the breast, where it is assumed that the geometry of the adipose region can be fixed in a support and the bulk permittivity is known. The goal is then to use the recovered parameters as prior information about the fibroglandular region, along with the known information about the adipose region to detect and monitor tumors. Therefore, it is desirable to extract information about the tumor-free breast in a way that is robust to the presence of tumors. This chapter presents a two-stage workflow in which a machine-learning-based first stage performs bulk parameter extraction of prior information for image reconstruction at the second stage.

The results are focused on noisy synthetic data, and demonstrate that a network trained on tumor-free data can successfully extract information about the fibroglandular region with and without a tumor present. These parameters are then used as prior information for CSI to demonstrate that they are of sufficient accuracy to detect and localize tumors within the fibroglandular region. Some preliminary results are presented for experimental measurement data. Unlike in the stored grain monitoring case, measurements from the breast imaging chamber can be calibrated. This chapter describes a method of calibrating the measurement data; experimental results are presented for calibrated data.

The chapter is organized as follows: in Section 5.2 we summarize the imaging system. In Section 5.3 we describe the stages of the framework wherein Section 5.3.1 provides details of the neural network model, dataset creation, training, and testing, and Section 5.3.2 discusses the application of the Stage 1 output to image reconstruction using CSI. The neural network is assessed for prediction accuracy, and the suitability of its output for inversion is shown in Section 5.4.

5.2 Faceted Air-Filled Chamber and its Model for Microwave Breast Imaging

We have designed our bulk-parameter extraction network around a previously developed breast imaging system that consists of an air-filled quasi-resonant chamber [13]. This chamber is easy to model, available for data collection, and has been studied experimentally [14, 36].

The imaging chamber consists of 24 magnetic field probes arranged on its 42 facets. Each field probe in the measurement system acts as a transceiver, resulting in a 24×24 matrix of $S_{i,j}$ measurements, where i is the transmitter index and j is the receiver index. In this system, these S-parameters represent the ϕ projections of the magnetic fields measured at the chamber walls.

We assume that the breast can be placed inside the chamber within a support that would fix the ROI boundary within the imaging chamber (Figure 5.1).

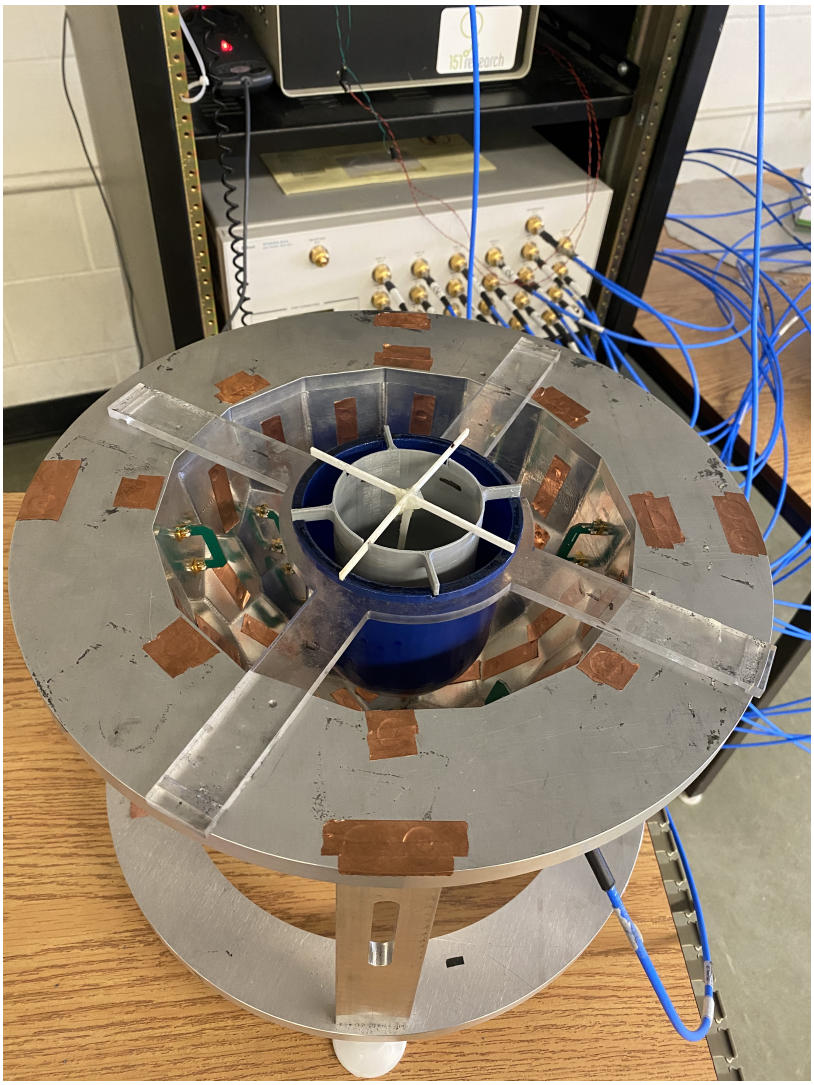


Figure 5.1: Photograph of the experimental microwave breast imaging setup with breast phantoms (blue).

A full-wave 3-D model of this system is constructed using either of our in-house FEM [32] or DGM [33] solvers, the latter of which supports high-order geometry simplifying modeling of curved regions [35]. These models output H-fields at specified receiver points for each transmitter, where we can convert between S-parameters and H-fields through

calibration (details of the transmitter and receiver probe numerical models can be found in [35].

The field data obtained from the system with a breast target is denoted H^{tot}_{breast} (this notation is shared for both S and H). Assuming some prior information of the target, we denote the total-field data for the prior information as H^{tot}_{prior}. Scattered fields of the breast relative to the prior information are:

$$H^{sct}_{target} = H^{tot}_{breast} - H^{tot}_{prior} \tag{5.1}$$

5.3 A Two-Stage Workflow for Prior Information Extraction and Data Inversion

Imaging multi-wavelength high-contrast targets, such as the human breast in air, is extremely difficult without the use of accurate background information to reduce the contrast in the non-linear imaging problem [7, 25]. However, seeking bulk parameters using standard optimization techniques is computationally expensive [8]. As a result of these two facts, we propose a workflow for microwave breast imaging that consists of two distinct stages: Stage 1, bulk parameter (prior information) inference, and Stage 2, image reconstruction. The point is to allow the bulk parameters reconstructed in Stage 1 to serve as prior information for full image reconstruction in Stage 2 as shown in Figure 5.2. Regardless of the exact nature of the bulk parameters, Stage 1 outputs them as a vector $\underline{\mathbf{p}}$. Both stages of the workflow accept a form of the measurement data from the imaging system, and are described in the following subsections.

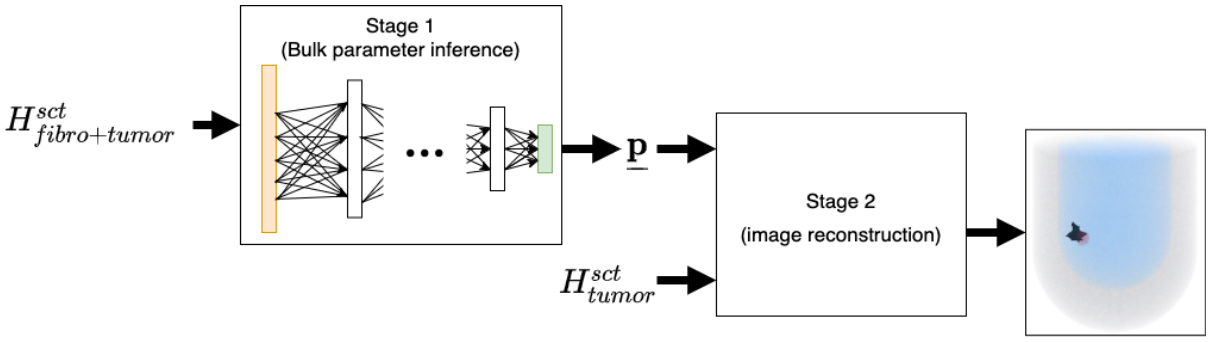


Figure 5.2: Bulk parametric inversion of scattered-field data to recover prior information (our vector of parameters $\underline{\mathbf{p}}$), followed by data-to-image reconstruction using neural network recovered prior information.

5.3.1 Stage 1: Bulk Parameter Inference Using Neural Networks

The goal of Stage 1 of our two-stage workflow is to estimate bulk parameters of the breast for use as prior information in Stage 2. In this work we choose our bulk parameters to define only the fibroglandular region of the breast, namely its tissue radius, height, and complex-valued permittivity. Thus, our proof-of-concept machine learning model assumes an adipose region of fixed extent, within which exists a centered fibroglandular region of varying size (height and radius). We note that this assumption for the adipose region is not as limiting as it may at first appear as the imaging system can be setup to fix the breast surface and thus the ROI boundary. While the ability to fix the location of the breast surface is a reasonable assumption, the choice of a single adipose region geometry with varying fibroglandular regions is limiting, as it does not account for a wide variety of breast sizes and densities. This limitation of our model is likely to reduce the generalizability of the network to accurately characterize a large variety of breasts, without further expanding the synthetic training set. Future work will consider making this model more flexible. We have chosen to ignore the equivalent of skin in our model, as previous work has shown that inclusion of the skin layer is not critical when attempting to accurately reconstruct the adipose region [62]. However, as shown in [25, 26], an approximate skin-layer can be introduced as required to improve the results.

Figure 5.3 shows renderings of the interior of the imaging chamber for two representative breast models.

For this stage of the workflow, the neural network accepts flattened matrices of data $H^{sct}_{fibro+tumor}$ as input. These fields, as defined in Equation (5.1), are obtained by taking the difference of total fields H^{tot}_{breast} (which may or may not contain a tumor) and $H^{tot}_{prior} = H^{tot}_{adipose}$, i.e., total fields for the fixed adipose region without the fibroglandular region and without any tumor. The goal is to pass $H^{sct}_{fibro+tumor}$ as an input sample to a trained neural network and have it return the parameters $\underline{\mathbf{p}}$ as shown in Figure 5.4. By design this parameter extraction does not attempt to detect or localize any tumors, but seeks an approximation to the fibroglandular region that is sufficiently accurate to allow tumor detection in Stage 2 as described in Section 5.3.2.

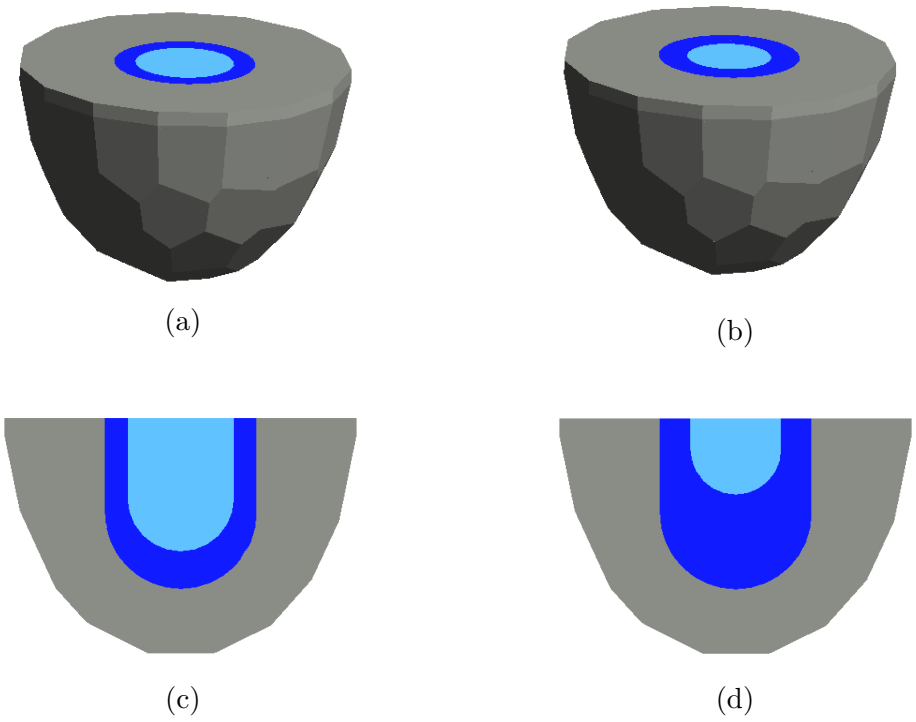


Figure 5.3: Renders of interior of faceted air-filled chamber (a),(b) and cross-section (c),(d) showing the three regions: air (grey), fat (dark blue), fibroglandular (light blue), for a medium (left) and small (right) fibroglandular region.

5.3.1.1 Labelled Data

In order to facilitate supervised learning, a large dataset that is representative of data obtained using the experimental setup is required. It is not practical to obtain a large number of labelled experimental measurements from this system, and synthetically generated data must be used. We have previously shown that bulk-parameter networks can be trained on synthetic data and successfully applied to uncalibrated experimental data [56]. Synthetic data should be chosen to cover a range of expected fibroglandular geometries and permittivities within the fixed adipose region.

When building the training data we can choose whether tumors should be included or excluded from the dataset. The approach taken herein is to exclude tumors from the training set altogether. This has the benefit of not requiring an extensive dataset representative of the many possible variations in tumor properties. Of course this approach necessitates demonstrating that the network is robust to the presence of tumors when predicting bulk parameters. That is, the network should accurately predict the $\underline{\mathbf{p}}$ parameters even from data obtained from phantoms containing tumors.

We have chosen to use single-frequency, phaseless data as Stage 1 input and to ignore

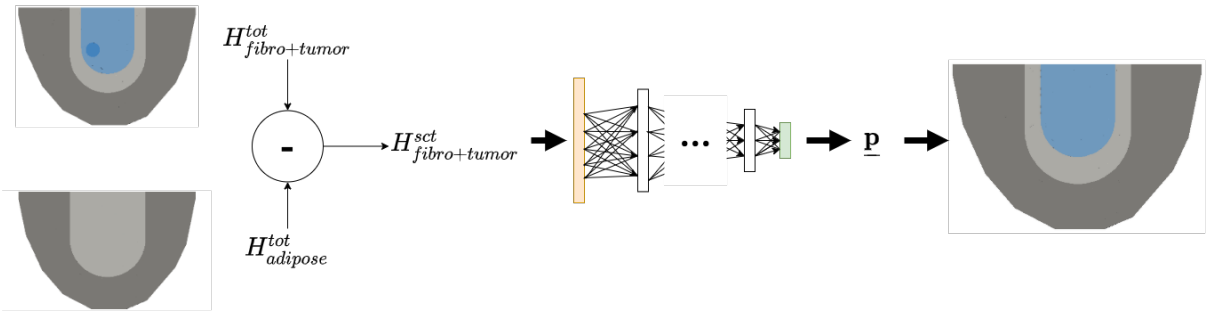


Figure 5.4: The Stage 1 bulk parameter inference network accepts breast-target data which may include a tumor, and outputs the bulk fibroglandular parameters.

back-scatter measurements, i.e., transmitter/receiver pair (i, i). Thus, each labelled sample consists of $(|H^{sct}_{fibro(\underline{\mathbf{p}})}|, \underline{\mathbf{p}})$, where $|H|$ represents the element-wise magnitude of the matrix H, and $H^{sct}_{fibro(\underline{\mathbf{p}})}$ is the now 24×23 matrix of scattered fields for the fibroglandular region parameterized by $\underline{\mathbf{p}}$, relative to the assumed known adipose background where the element-wise subtraction in Equation (5.1) is carried out in the complex domain. In this particular model, the parameter vector $\underline{\mathbf{p}}$ consists of four labels: the radius, height, as well as the real and imaginary parts of the complex-valued permittivity (respectively ε' and ε'') for the fibroglandular region.

5.3.1.2 Stage 1 Network Architecture

The adopted network architecture is a fully-connected neural network. Each input sample, $|H^{sct}_{fibro+tumor}|$ consisting of 24×23 real values, is flattened into a 552×1 vector to be compatible with the network's 552-node input layer. The rest of the network consists of 5 hidden layers, with a 4-parameter output vector. All layers in the network have ReLU activation. The full network architecture is **552**→2048→1024→512→256→64→**4** (where the input and output layer sizes are bolded for clarity). We use a residual sum of squares loss function (Equation 3.3).

5.3.1.3 Sample Pre-Processing

Before being passed to the neural network data are passed through a pre-processing step, where the complex measurement data are converted to magnitudes, and normalized by the maximum value on a per-transmitter basis. All test examples are subjected to the same pre-processing steps as the training data (conversion to magnitude-only, and per-transmitter normalization).

5.3.2 Stage 2: 3D Image Reconstruction and Tumour Detection

Once Stage 1 of the workflow has provided an estimate of $\underline{\mathbf{p}}$, Stage 2 is applied to detect and/or monitor tumors. In this inversion stage of the workflow, the inputs are the calibrated $H^{sct}_{tumor} = H^{tot}_{breast} - H^{tot}_{fibro+adipose}$. To calculate H^{sct}_{tumor} we have effectively introduced the need for a registered baseline measurement $H^{tot}_{baseline}$, which we take to be $H^{tot}_{baseline} = H^{tot}_{fibro+adipose}$. An appropriate registered baseline measurement implies Stage 1 parameters $\underline{\mathbf{p}}$ that are the same for both H^{tot}_{breast} and $H^{tot}_{baseline}$.

The overall two-stage workflow is summarized in Figure 5.2, and Figure 5.5 outlines the second stage of this workflow as presented in this work. The approach used here adopts the CSI algorithm in Stage 2. Previous work has demonstrated the ability of CSI to reconstruct the permittivity profile of the breast model from experimental data assuming accurate prior information is known [25]. Demonstrating that Stage 2 works with imperfect prior obtained from Stage 1 is a significant step towards practicality. Of course many other Stage 2 approaches are also possible i.e., any other inversion algorithm or a machine learning approach.

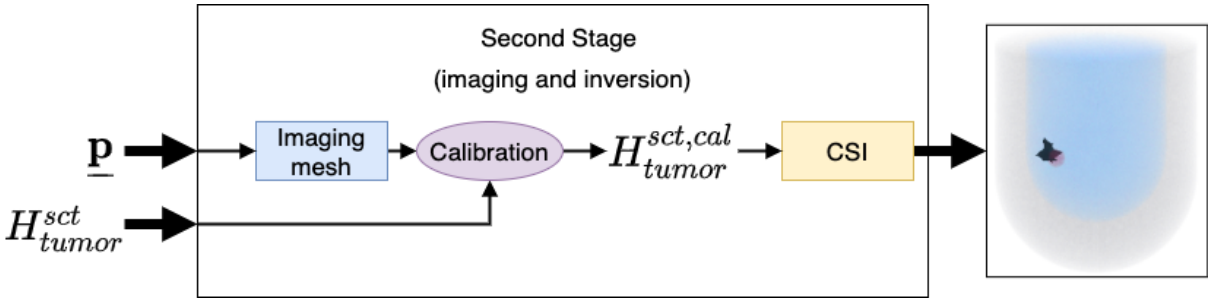


Figure 5.5: Workflow stage 2 employing CSI, where $\underline{\mathbf{p}}$ is obtained from the bulk parameter inference network (first stage), and specifies the fibroglandular region in the imaging mesh.

5.3.2.1 Contrast Source Inversion

CSI is a well-known iterative approach to solving the non-linear inverse scattering problem [19]. Our implementation of CSI used for this work is based on a partial differential equation formulations of Maxwell's equations namely the finite element method (FEM-CSI). This algorithm easily supports prior information in the form of inhomogeneous backgrounds, details can be found in [17, 32]. The parameters obtained from the neural network, $\underline{\mathbf{p}}$, are used as prior information for the inversion where the fibroglandular region height and radius predictions are used to generate a new model for which a forward solver

is called to generate $H^{tot}_{fibro(\underline{\mathbf{p}})+adipose}$ data for this tissue geometry, where the permitivitty of the fibroglandular region used for the forward solver is the complex-valued permittivity prediction from $\underline{\mathbf{p}}$. The background physics for the CSI inversion are then taken to be the assumed known complex-valued permittivity for adipose tissue, and the $\underline{\mathbf{p}}$ predicted complex-valued permittivity for fibroglandular tissue. An imaging model is generated according to these parameters, whose discretization is intentionally different from that used to generate the forward data for the test examples.

CSI seeks to reconstruct regions of contrast with respect to the provided prior background. As discussed in Section 5.3.2 we currently assume that the background is characterized by the baseline measurement of the breast in which the field measurements of the adipose and fibroglandular regions of the breast are known. In this configuration, CSI should reconstruct the tumor.

5.3.3 Calibration

As forward model discretization can result in significant simulated field errors relative to the scattering signature of the target, it is necessary to calibrate the test data to the imaging model (CSI prior information) to eliminate errors arising from the use of different models. The calibrated scattered fields are computed as

$$H^{cal}_{target} = \frac{H^{tot}_{prior}}{H^{tot}_{baseline}}(H^{tot}_{breast} - H^{tot}_{baseline}) \tag{5.2}$$

where $H^{tot}_{prior} = H^{tot}_{fibro(\underline{\mathbf{p}})+adipose}$ is the total field forward data for the prior information (imaging) model, and H^{tot}_{breast} and $H^{tot}_{baseline} = H^{tot}_{fibro+adipose}$.

5.4 Results

In what follows we demonstrate the proposed two-stage workflow on synthetic data. Initial testing on experimental data shows that Stage 1 is capable of recovering the geometry of the fibroglandular region.

5.4.1 Data Generation

A synthetic model of the air-filled chamber described in Section 5.2 was created. To simplify the model and speed up synthetic data generation, we elected to set the top plane of the system to a PEC. We note that in practice an impedance boundary condition may be more appropriate for modelling the patient chest wall across the top surface of the chamber. In the model, the 24 magnetic field probes, when transmitting, are approximated as 24 magnetic dipole sources placed tangential to the appropriate chamber facet that correspond to their location in the experimental system.

The synthetic data used for training and evaluating the network was generated using a high-order DGM forward solver [33, 35]. In all cases, second-order mesh geometry and third-order field solutions were used.

Data was generated at 1.1 GHz, corresponding to a freespace wavelength of around 27cm or roughly 1.2 times the radius at the top of the chamber. Each of the tissue types considered in this model were characterized based on an approximation of their complex-valued permittivity at 1.1 GHz: in all cases adipose was assigned a permittivity of $3-j0.6$, and tumors were assigned a permittivity of $56.3-j30$ [23]. As we have previously assumed that the experimental imaging system can be designed with a support that would fix the overall size and shape of the target to known geometry, for all synthetic training examples the size of the adipose region was fixed with a height of 10.9 cm and a radius of 4.8 cm.

To create a labelled dataset, synthetic breast models were generated for varying fibroglandular region heights and radii ranging from 5.30 to 9.80 cm and 2.85 to 4.10 cm, and for a wide range of complex-valued bulk permittivities for the fibroglandular tissue: the real part of the permittivity ranged from 15 to 25 and the imaginary part of the permittivity ranged from -25 to -15. The range of values used to generate the labelled dataset are summarized in Table 5.1.

Table 5.1: Summary of fibroglandular region variations in the labelled dataset.

Radius range [cm]	Height range [cm]	ε' range	ε'' range
[2.85, 4.10]	[5.30, 9.80]	[15, 25]	[-25, -15]

The labelled dataset consists of a total of 7,116 tumor-free examples, which are partitioned into training, validation and reserved test sets of 85%, 10%, and 5% of the set, respectively. Tumors were intentionally omitted from all training examples so that we

could later evaluate the robustness of the Stage 1 network to the presence of tumors.

Table 5.2 shows the average absolute error and standard deviation in each parameter for the testing set comprised of 5% of the noise-free synthetic examples described above.

Table 5.2: Error in fibroglandular region bulk parameter predictions for the noise-free, tumor-free synthetic test set.

Metric	Error in Radius [mm]	Error in Height [mm]	Error in ε'	Error in ε''
Average absolute error	0.0280	0.0248	0.0866	0.0963
Average absolute error (%)	0.97%	0.83%	0.58%	0.39%
Standard deviation	0.0204	0.0194	0.0670	0.0765

5.4.2 Tumor Detection Test Samples

In addition to the tumor-free labelled dataset described above, synthetic data was generated for three unique geometries representing small, medium, and large fibroglandular regions. For these unique geometries, one baseline (tumor-free), and two tumor-containing examples were generated, where the tumor is a spherical target with a radius of 9mm. In each tumor-containing example, the target was placed at one of two different locations within the fibroglandular region. For the medium fibroglandular case, data was also generated for a 4.5mm radius tumor. In all cases the complex-valued permittivity of the fibroglandular region was set to be $20 - j21.6$ [23]. A single fibroglandular permittivity was chosen to reduce the problem size for this proof of concept workflow, and the particular value was selected to correspond to an existing experimental phantom. This data corresponds to different examples of H^{tot}_{breast}. To simulate measurement noise, white Gaussian noise between $\pm 10^{-4} * \mathbf{max}(real(|H^{tot}_{breast}|))$ and $\pm 10^{-4} * \mathbf{max}(imag(|H^{tot}_{breast}|))$ was applied to the real and imaginary parts of the data, respectively, where 10^{-4} represents -80dB noise for the total field data.

5.4.3 Parametric Inference of Fibroglandular Region Parameters

Using the training set, a single frequency, magnitude-only (phaseless) neural network, described in Section 5.3.1.2 was trained to perform bulk parameter inference. The model

was generated using Python and TensorFlow Keras. This neural network was trained with a batch size of 100, for a maximum of 150 epochs with an early stopping patience of 30 iterations.

The ability of the neural network to predict the geometry and complex permittivity of the fibroglandular region was evaluated for synthetic cases with and without a tumour. In all cases the output of the network, $\underline{\mathbf{p}}$, describes only the fibroglandular region, and does not directly provide any information on the presence, size, or location of the tumor.

Table 5.3 presents the output of the neural network for three different fibroglandular region sizes, each with two different tumor positions. The true value of the parameters is provided for reference. For the medium fibroglandular case, two different sized tumors were tested at position one. All tumors are spherical, with radius as specified in column two of the table.

Table 5.3: Fibroglandular region bulk parameter predictions for synthetic examples with -80dB noise.

Position	Tumor radius [mm]	Radius [cm]	Height [cm]	ε'	ε''
Small fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		2.90	5.35	20.0	-21.6
No tumor	-	3.06	6.04	16.02	-21.85
T1	9	3.05	6.09	16.66	-22.08
T2	9	3.06	6.13	16.03	-22.07
Medium fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		3.40	8.50	20.0	-21.6
No tumor	-	3.39	8.46	19.90	-21.60
T1	9	3.39	8.47	19.71	-21.72
T1	4.5	3.40	8.44	20.03	-21.57
T2	9	3.39	8.45	20.48	-21.58
Large fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		4.05	9.75	20.0	-21.6
No tumor	-	4.07	9.71	19.81	-21.55
T1	9	4.07	9.73	19.74	-21.36
T2	9	4.06	9.72	19.79	-21.58

The results in Table 5.3 suggest that the neural network may be capable of accurately predicting the prior information needed for CSI. It also shows that the network is robust to the presence of tumors of different sizes and locations for the medium and large fibroglandular regions. The neural network is not as successful in reconstructing the bulk

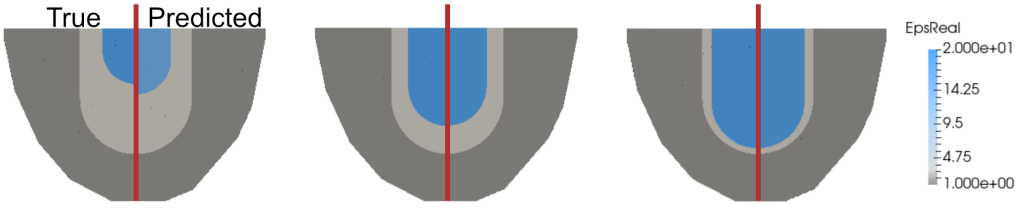
parameters when the fibroglandular region is small relative to the size of the fat region.

To further evaluate the impacts of noise on the parameter predictions from Stage 1, each of the test cases in Table 5.3 was evaluated for -40dB and -60dB noise (in addition to the -80dB noise presented above). Table 5.4 shows the average and maximum absolute error for each parameter across the 10 tumor-containing, noisy test cases. These results demonstrate that the network remains robust to noise when the noise applied is increased from -80dB to -60dB. For -40dB noise the network still gives reasonable results, but the prediction error begins to increase and in the case of the complex-valued permittivity predictions, the maximum error is nearly half of the true value.

Table 5.4: Error in Error in fibroglandular region bulk parameter predictions for the noisy, tumor-containing test cases.

Noise	Metric	Error in Radius [mm]	Error in Height [mm]	Error in ε'	Error in ε''
10^{-4} (-80dB)	Average	0.0545	0.0251	1.28	0.167
	Max	0.161	0.775	3.98	0.478
10^{-3} (-60dB)	Average	0.0549	0.0250	1.41	0.283
	Max	0.180	0.779	4.56	0.679
10^{-2} (-40dB)	Average	0.115	0.444	4.06	2.92
	Max	0.233	1.31	7.47	10.1

Figure 5.6 shows the neural network prediction versus true prior information for the 9mm tumor in position T1 for each fibroglandular size.



(a) (b) (c)

Figure 5.6: True values (left) versus neural network predictions (right) for the fibroglandular region for the T1 test examples for the (a) small fibroglandular case, (b) medium fibroglandular case, and (c) large fibroglandular case. The associated differences in the predicted complex-valued permittivity are provided in Table 5.3.

5.4.4 CSI-Based Tumor Detection From Predicted Prior Information

To demonstrate that the output of the Stage 1 network is accurate enough to be used as prior information for CSI, the parameters, $\underline{\mathbf{p}}$, extracted by the network for the tumor detection test samples were used as prior information for CSI reconstructions of the fibroglandular region, where the goal is to detect and localize the tumor. For each test example, CSI was run to recover the tumor (high contrast region). The test sample data was calibrated according to the procedure outlined in Section 5.3.3. The synthetic test data is generated using a completely independent forward solver that was used to create the training set data. Therefore, although counter-intuitive, the calibration procedure, which is typically used to calibrate experimental data to the numerical inverse model, must also be applied to the synthetic test data set. This effectively calibrates data obtained from one modelling engine (the DGM forward solver in this case), with it's own inherent systematic errors (such as mesh error), to the training set data obtained via a different modelling engine (the FEM forward solver in the present case). Figure 5.7 shows the reconstruction of the tumor in two different positions within each size of fibroglandular region superimposed on the true prior. To visualize the tumor, the real part of the recovered contrast, χ, was thresholded at 85% of the maximum reconstructed contrast for the given test example, these thresholded regions are superimposed on the true real permittivity map of the model (Figure 5.7). This result shows that when using prior information obtained from the Stage 1 neural network, the approximate location of a tumor within the breast model can be determined.

As the intent of this work is to demonstrate the ability to distinguish the tumor tissue from the background, and not necessarily to accurately determine the permittivity, CSI was terminated after relatively few (200) iterations. A consequence of this choice is that permittivity reconstruction is consistently lower than expected. Running CSI for more iterations, and applying post-processing to the images using deep learning has been shown to improve the artifact prone, low contrast inversions shown in Figure 5.7 [23].

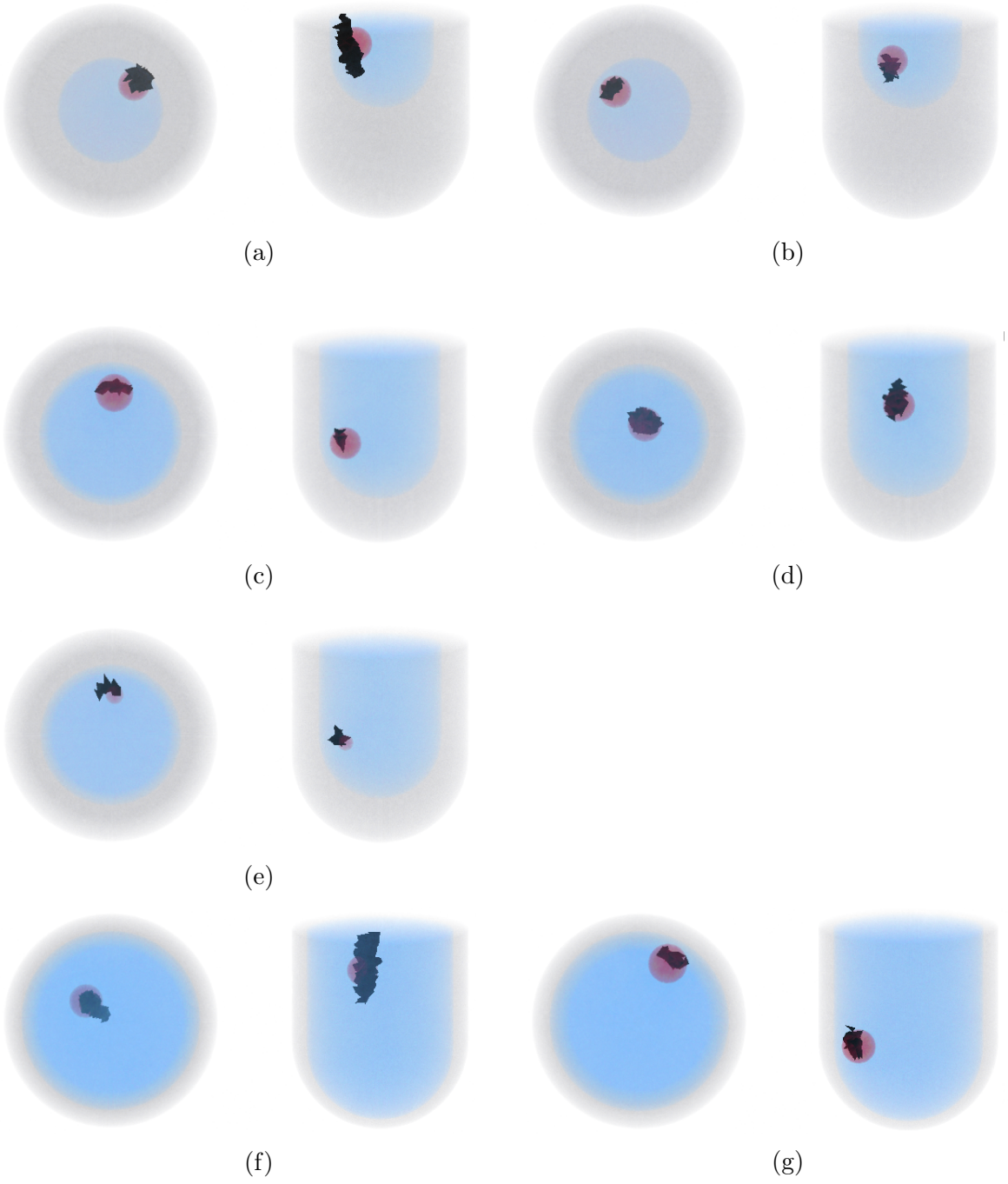


Figure 5.7: CSI reconstructions for two different tumor positions, when prior information has been generated by our neural network. +z projection (left) and +x projection (right) are shown. (a) and (b) show T1 and T2 for the small fibroglandular case; (c), (d), and (e) show T1 (9mm), T2, and the T1 (4.5 mm) for the medium fibroglandular case; (f) and (g) show T1 and T2 for the large fibroglandular case. Adipose (grey), fibroglandular (blue), and tumor (red) show the true geometry of the breast model tissues. The 85% thresholded contrast representing the reconstructed tumor is shown in black.

5.4.5 Monitoring Response to Tumor Treatment

Figure 5.8 shows the imaging results for the T1 tumor (fixed position) with two different tumor sizes, compared with the imaging result for the tumor-free case. The CSI reconstruction of the data for these two different tumor sizes shows a decrease in the size and recovered contrast of the region identified as a tumor with a decrease in tumor size in the forward problem. This result suggests that the proposed workflow may be useful in monitoring tumor response to treatment, where the most recent image of the tumor is used as baseline prior information.

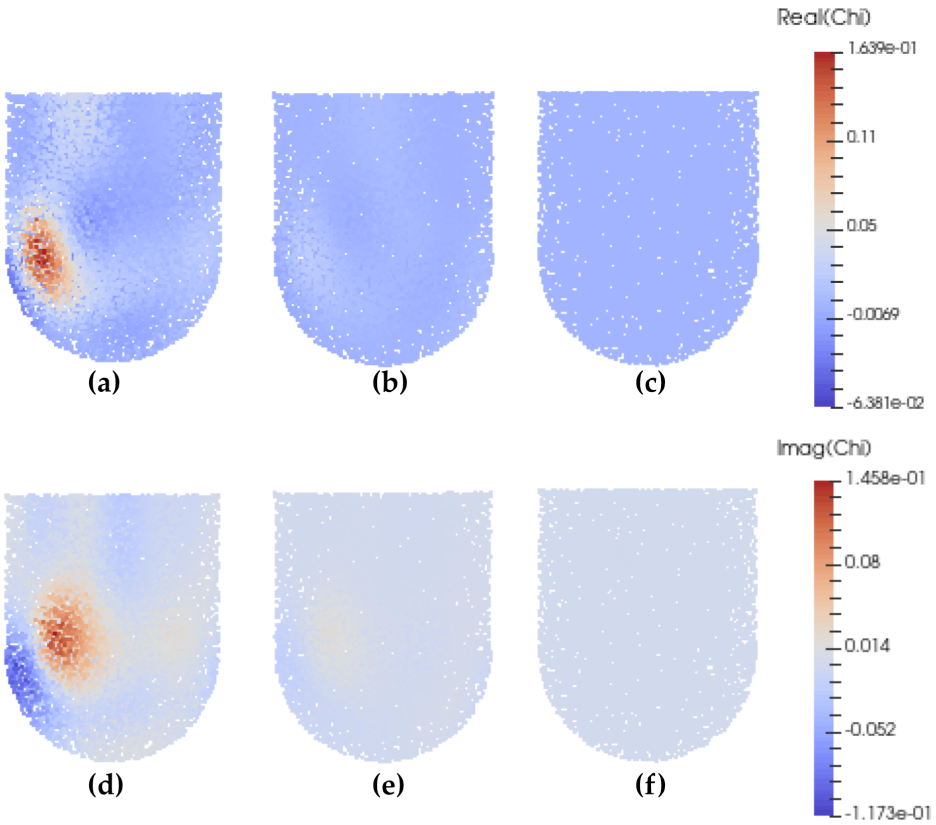
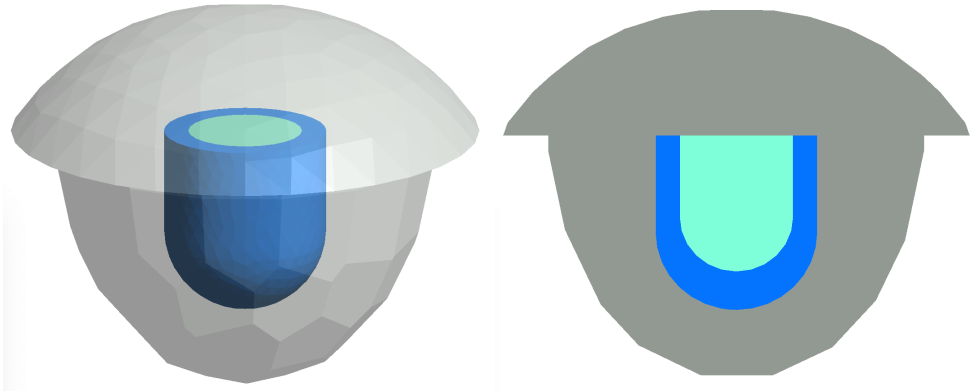


Figure 5.8: Point cloud representation of the CSI reconstruction of the real and imaginary parts of the complex contrast for the fibroglandular region for the T1 test examples with (a)/(d) a 9 mm tumor at position T1, (b)/(e) a 4.5 mm tumor at position T1, and (c)/(f) no tumor.

5.4.6 Preliminary Bulk Parameter Inference Results on Experimental Data

A limited set of experimental data obtained from the system as described in [14] was evaluated using a similar bulk parameter inference network. The experimental breast phantom consisted of a phantom that fits our general description of fibroglandular and adipose tissues (e.g. Figure 5.3), with tissue mimicking fluids filling each region in thin-walled containers (see Fig. 5.1). Two measurements were taken with only adipose and fibroglandular regions and two measurements were taken with a tumor in the fibroglandular region in two different locations.

Forward data was generated for the same parameters and ranges as outlined in Table 5.1. The experimental data was collected with an open top, and thus the flat PEC top model used for the synthetic results presented above is unsuitable. Instead, to model the experimental setup, the imaging system model was modified such that the region above the top surface of the chamber is air, where an absorbing boundary condition (ABC) is applied. In order to keep the model discretization manageable for timely labelled dataset generation, the absorbing boundary condition was left in relative close proximity to the chamber's PEC surfaces limiting the model's accuracy (Figure 5.9).



(a) (b)

Figure 5.9: Synthetic model of the breast within the system with air (grey), adipose (dark blue) and fibroglandular (light blue) shown. The full chamber is shown in (a), and a cross section in (b).

New data was generated according to this model to train a neural network of the same architecture described in Section 5.3.1.2. The new synthetic dataset consists of 25,409 labelled examples (partitioned 85%, 10%, 5% into training, validation, and test

sets, respectively). Prior to being passed as inputs to the neural network, the experimental measurements were calibrated to the synthetic data as scattered fields relative to a measured adipose model background. For experimental data, these background measurements were obtained by filling the phantoms with canola oil. The full process for experimental data collection is described in [14]. Table 5.5 provides preliminary predictions of $\underline{\mathbf{p}}$ for experimental data.

Table 5.5: Fibroglandular region bulk parameter predictions for calibrated experimental examples.

Position	Tumor radius [mm]	Radius [cm]	Height [cm]	ε'	ε''
True values: $\underline{\mathbf{p}}_{true}$		3.40	8.50	20.0	-21.6
No tumor 1	-	3.50	8.39	13.37	-14.24
No tumor 2	-	3.47	8.29	13.38	-14.85
T1	9	3.56	8.56	11.83	-15.38
T2	9	3.55	8.25	11.61	-16.35

These results show that, for the appropriate synthetic training set, the Stage 1 network architecture outlined in this work is capable of recovering the fibroglandular geometry information with reasonable accuracy. In these preliminary results, the neural network undershoots the complex-valued permittivity predictions, which may be due to the accuracy of the absorbing boundary condition used when generating the training set. Further research is ongoing to determine ways to improve the results for experimental data.

5.4.7 Multi-Frequency Neural Network Performance

This section is based on [63], accepted to URSI GASS 2021.[2].

The impact of including multi-frequency data in the machine learning workflow for microwave breast imaging was evaluated using the same ABC open-top model that was used to build the synthetic dataset to provide preliminary experimental results. The results for a single-frequency network at 1.1 GHz using this altered model are given in Table 5.6. As in previous sections, all synthetic test examples have -80 dB random white Gaussian noise.

These results provide comparable performance to the single-frequency PEC flat-top

[2]K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, I. Jeffrey, "Recovery of prior information for breast microwave imaging using neural networks," *Accepted to URSI GASS 2021*, 2021

Table 5.6: Fibroglandular region bulk parameter predictions from the single-frequency neural network. ©2021 IEEE.

Position	Tumor [mm]	Radius [cm]	Height [cm]	ε'	ε''
Small fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		2.90	5.35	20.00	-21.60
No tumor	-	2.93	5.32	19.76	-21.46
T1	9	3.10	5.32	12.00	-23.84
T2	9	2.96	5.30	19.07	-21.53
Medium fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		3.40	8.50	20.00	-21.60
No tumor	-	3.40	8.43	19.50	-21.75
T1	9	3.41	8.46	19.20	-21.76
T1	4.5	3.40	8.48	19.72	-21.65
T2	9	3.41	8.46	19.56	-21.89
Large fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		4.05	9.75	20.00	-21.60
No tumor	-	4.03	9.76	20.09	-21.91
T1	9	4.03	9.77	20.11	-21.91
T2	9	4.03	9.77	19.88	-22.00

model used previously (Table 5.3). This suggests that the machine learning workflow is capable of obtaining bulk parameters from noisy synthetic test data and is robust to the presence of tumors within the fibroglandular region as long as the appropriate synthetic dataset is used for network training. This is encouraging for future evaluation of experimental results, as it suggests that there is flexibility in the synthetic model, and that improvements to the model may help to improve performance on experimental data.

Table 5.7 shows the results from a five-frequency multi-frequency neural network using 1.1, 1.2, 1.3, 1.4, and 1.5 GHz data. Comparing Table 5.6 with Table 5.7 shows that the addition of multiple frequencies improves the parameter predictions for the small and medium fibroglandular region cases. While the error for the large fibroglandular region case is slightly higher for multi-frequency networks than for single-frequency networks, the results are still reasonable and based on the results in Section 5.4.4, are of acceptable accuracy to be used as prior information for full inversion.

Table 5.7: Fibroglandular parameter predictions from the multi-frequency neural network. ©2021 IEEE.

Position	Tumor [mm]	Radius [cm]	Height [cm]	ε'	ε''
Small fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		2.90	5.35	20.0	-21.6
No tumor	-	2.90	5.30	20.34	-21.50
T1	9	2.91	5.30	19.97	-22.04
T2	9	2.90	5.30	20.21	-21.64
Medium fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		3.40	8.50	20.0	-21.6
No tumor	-	3.41	8.50	20.04	-21.62
T1	9	3.41	8.50	19.82	-21.59
T1	4.5	3.41	8.50	19.90	-21.60
T2	9	3.41	8.50	19.96	-21.90
Large fibroglandular case:					
True values: $\underline{\mathbf{p}}_{true}$		4.05	9.75	20.0	-21.6
No tumor	-	4.06	9.70	19.70	-21.77
T1	9	4.06	9.70	19.70	-21.77
T2	9	4.06	9.70	19.70	-21.85

5.5 Chapter Summary

Application of the bulk parametric inversion neural network to the microwave breast imaging problem further demonstrates how the workflow presented herein can be generalized to different electromagnetic imaging systems. By presenting a two-stage workflow that combines a neural network (stage 1) and 3D image reconstruction using CSI (stage 2), it has been shown that the bulk parameters obtained from the neural network are of sufficient quality to support full-phase inversion of the electromagnetic imaging data.

This chapter focused primarily on single-frequency data from noisy synthetic test examples to prove out the two-stage workflow and demonstrated that from this data tumors can be detected and localized within a model of a human breast. Bulk parameters were inferred from calibrated experimental data, showing encouraging results for the geometry and reasonable (but lower than expected) predictions for the complex-valued permittivity. In addition, predictions on noisy synthetic examples from a multi-frequency network were presented and suggest that incorporating data from multiple frequencies may, in some cases, improve bulk parameter inference results.

Chapter 6

Conclusions and Future Work

This thesis has presented a generalized workflow for a synthetically-trained machine-learning-based method enabling rapid parametric inversion of experimental measurements from different electromagnetic imaging systems. For the applications explored herein, the parameters are chosen to represent the bulk geometry and bulk complex-valued permittivity of a target; these parameters could be modified and/or expanded in the future as needed. The inferred values for these parameters can later be used as prior information to enable calibration and facilitate full-phase inversion of the electromagnetic data. The method's utility has been demonstrated for both single- and multi-frequency parametric inversion for two different applications at significantly different scales and corresponding frequencies: stored grain monitoring (10s of meters and 60-110 MHz), and microwave breast imaging (10s of centimeters and 1-1.5 GHz).

The method requires the creation of a large synthetic training set unique to the imaging system, which represents a significant initial computational cost. However, this initial cost quickly offset by the significant reduction in time required to perform the parametric inversion, making this supervised machine learning approach a cost-effective long-term monitoring solution.

In the stored grain monitoring case, it has been shown that synthetically trained networks can be used to perform a simple (4-parameter) bulk parametric inversion on experimental data from uncooperative systems where calibration is not possible. A primary limitation of this simple inversion is that restricting the parameterization of the bin to only four parameters is an oversimplification that does not account for any variation in the permittivity of the grain throughout the volume, and is limited to simple grain-air

interfaces. Expansion of the parameter set to model more complex grain fill geometries (using two additional parameters) was investigated and shown to provide reasonable results for noisy synthetic data.

The microwave breast imaging work focused on parametric inversion of noisy synthetic data but takes the workflow a step further, demonstrating how the results of the inversion from the neural network can be used as the necessary prior information to facilitate image reconstruction using CSI. These reconstructed images can be used to detect and localize tumors and monitor tumor response to treatment. Preliminary results were presented for calibrated experimental data, which showed reasonable geometry predictions and some robustness to the presence of tumors, but yielded consistently low permittivity estimates.

This work has demonstrated that synthetically trained neural networks can be a useful tool for connecting computational models of electromagnetic imaging systems to experimental data obtained from these systems. Such networks provide a way of obtaining the necessary prior information and calibrating the experimental data to the computational model of the system to support full inversion (demonstrated herein using CSI) and reconstruct a 3D image of the ROI. The developed workflow makes simplifying assumptions about the geometry and permittivity profiles of the materials of interest. Even with these assumptions, the results presented for experimental data, particularly in the case of stored grain monitoring where the imaging system is in an uncontrolled outdoor environment, have shown that this is a feasible method of analysing experimental data at industrial scales. This work could be extended and improved by accounting for some of these assumptions, which may further improve the performance of the networks and increase the practicality of the method for industrial applications.

Suggestions for Future Work

- The extended parameter set demonstrated for stored grain monitoring serves as a proof-of-concept to support future investigation into other more complex parameter sets. Such parameters sets need not be restricted to modelling complex grain-air interfaces, but could be used to infer features such as different permittivity distributions within the grain, or additional information about the model of a human breast.

- To date, the human breast model is relatively simple and assumes uniform tissue

regions with smooth boundaries. The method in this work still needs to be proven for more realistic breast models.

- Investigation into the ability to extract bulk information from these neural networks when there exists larger inhomogeneities in the region should be considered; such inhomogeneities may include gradients in moisture content, or large regions of high moisture content in stored grain, and larger or multiple tumors within the breast fibroglandular regions.

- This research would benefit from investigation into the permittivity estimates obtained from the parametric inversion model to understand why the predictions, although reasonable, tend to be higher than expected for stored grain and lower than expected for the breast fibroglandular tissue. Additional studies of the networks' robustness to regions of high permittivity, and how the presence of such regions affects the inferred parameters should also be considered.

- Independent of the application, an interesting next step for this research would be to evaluate a full end-to-end machine learning workflow, where the output of the parametric inversion neural network is used by a deep-learning model designed to replace the CSI step, such as those presented in [23, 43].

- Semi-supervised learning techniques could be investigated as a way of augmenting the labelled synthetic dataset. If successful, this could further reduce the upfront cost associated with training set generation using a forward solver.

- The use of Generative models (GANs) for creating new labelled training sets based on experimental data sets could be explored as an alternative to the forward-solver-based labelled synthetic data generation used in this work.

- For model simplicity, this work chooses to neglect the dispersive properties of the permittivity of target materials. An interesting future direction would be to attempt to build a dispersive model that accounts for this and investigate whether or not the new model improves the accuracy of the predictions.

Bibliography

[1] S. I. Kabanikhin, *Inverse and ill-posed problems: theory and applications*. Walter De Gruyter, 2011, vol. 55.

[2] M. Oristaglio and H. Blok, "Wavefield imaging and inversion in electromagnetics and acoustics," *Course Notes*, 1995.

[3] M. T. Bevacqua, G. G. Bellizzi, T. Isernia, and L. Crocco, "A method for effective permittivity and conductivity mapping of biological scenarios via segmented contrast source inversion," *Progress In Electromagnetics Research*, vol. 164, pp. 1–15, 2019.

[4] N. Abdollahi, D. Kurrant, P. Mojabi, M. Omer, E. Fear, and J. LoVetri, "Incorporation of ultrasonic prior information for improving quantitative microwave imaging of breast," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 4, pp. 98–110, 2019.

[5] A. Asnaashari, R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, "Regularized seismic full waveform inversion with prior model information," *Geophysics*, vol. 78, no. 2, pp. R25–R36, 2013.

[6] R. Guo, X. Song, M. Li, F. Yang, S. Xu, and A. Abubakar, "Supervised descent learning technique for 2-d microwave imaging," *IEEE Transactions on Antennas and Propagation*, vol. 67, no. 5, pp. 3550–3554, 2019.

[7] L. M. Neira, B. D. Van Veen, and S. C. Hagness, "High-resolution microwave breast imaging using a 3-d inverse scattering algorithm with a variable-strength spatial prior constraint," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 11, pp. 6002–6014, 2017.

[8] C. Gilmore, I. Jeffrey, M. Asefi, N. T. Geddert, K. G. Brown, and J. LoVetri, "Phaseless parametric inversion for system calibration and obtaining prior information," *IEEE Access*, vol. 7, pp. 128 735–128 745, 2019.

[9] X. Chen, Z. Wei, M. Li, and P. Rocca, "A review of deep learning approaches for inverse scattering problems (invited review)," *Progress In Electromagnetics Research*, vol. 167, pp. 67–81, 2020.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[11] C. Gilmore, M. Asefi, J. Paliwal, and J. LoVetri, "Industrial scale electromagnetic grain bin monitoring," *Computers and Electronics in Agriculture*, vol. 136, pp. 210–220, 2017.

[12] R. Benny, T. A. Anjit, and P. Mythili, "An overview of microwave imaging for breast tumor detection," *Progress In Electromagnetics Research*, vol. 87, pp. 61–91, 2020.

[13] K. Nemez, M. Asefi, A. Baran, and J. LoVetri, "A faceted magnetic field probe resonant chamber for 3d breast mwi: A synthetic study," in *2016 17th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*. IEEE, 2016, pp. 1–3.

[14] M. Asefi, A. Baran, and J. LoVetri, "An experimental phantom study for air-based quasi-resonant microwave breast imaging," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 9, pp. 3946–3954, 2019.

[15] M. S. Zhdanov, *Geophysical inverse theory and regularization problems*. Elsevier, 2002, vol. 36.

[16] J. C. Curlander and R. N. McDonough, *Synthetic aperture radar*. Wiley, New York, 1991, vol. 11.

[17] A. Zakaria, C. Gilmore, and J. LoVetri, "Finite-element contrast source inversion method for microwave imaging," *Inverse Problems*, vol. 26, no. 11, p. 115010, 2010.

[18] J. D. Shea, P. Kosmas, S. C. Hagness, and B. D. Van Veen, "Three-dimensional microwave imaging of realistic numerical breast phantoms via a multiple-frequency

inverse scattering technique," *Medical physics (Lancaster)*, vol. 37, no. 8, pp. 4210–4226, 2010.

[19] P. M. Van Den Berg and R. E. Kleinman, "A contrast source inversion method," *Inverse problems*, vol. 13, no. 6, p. 1607, 1997.

[20] J. LoVetri, M. A. Asefi, C. Gilmore, and I. Jeffrey, "Innovations in electromagnetic imaging technology: The stored-grain-monitoring case," *IEEE Antennas and Propagation Magazine*, 2020.

[21] A. Kraszewski and S. Nelson, "Composite model of the complex permittivity of cereal grain," *Journal of Agricultural Engineering Research*, vol. 43, pp. 211–219, 1989.

[22] M. Lazebnik, D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, T. Ogilvie, A. Magliocco, T. M. Breslin *et al.*, "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries," *Physics in medicine & biology*, vol. 52, no. 20, p. 6093, 2007.

[23] V. Khoshdel, M. Asefi, A. Ashraf, and J. LoVetri, "Full 3d microwave breast imaging using a deep-learning technique," *Journal of Imaging*, vol. 6, no. 8, p. 80, Aug 2020. [Online]. Available: http://dx.doi.org/10.3390/jimaging6080080

[24] A. H. Golnabi, P. M. Meaney, and K. D. Paulsen, "3d microwave tomography of the breast using prior anatomical information," *Medical physics*, vol. 43, no. 4, pp. 1933–1944, 2016.

[25] D. Kurrant, A. Baran, J. LoVetri, and E. Fear, "Integrating prior information into microwave tomography part 1: Impact of detail on image quality," *Medical Physics*, vol. 44, no. 12, pp. 6461–6481, 2017. [Online]. Available: https://aapm.onlinelibrary.wiley.com/doi/abs/10.1002/mp.12585

[26] D. Kurrant, E. Fear, A. Baran, and J. LoVetri, "Integrating prior information into microwave tomography part 2: Impact of errors in prior information on microwave tomography image quality," *Medical physics (Lancaster)*, vol. 44, no. 12, pp. 6482–6503, 2017.

[27] A. H. Golnabi, P. M. Meaney, and K. D. Paulsen, "Tomographic microwave imaging with incorporated prior spatial information," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 5, pp. 2129–2136, 2013.

[28] N. Abdollahi, I. Jeffrey, and J. LoVetri, "Improved tumor detection via quantitative microwave breast imaging using eigenfunction-based prior," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1194–1202, 2020.

[29] M. Hughson, I. Jeffrey, and J. LoVetri, "Ultrasound and microwave imaging with prior property dependencies," in *2019 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)*. IEEE, 2019, pp. 1–4.

[30] R. Obermeier and J. A. Martinez-Lorenzo, "Compressive sensing unmixing algorithm for breast cancer detection," *IET Microwaves, Antennas & Propagation*, vol. 12, no. 4, pp. 533–541, 2018.

[31] X. Chen, *Computational methods for electromagnetic inverse scattering*. Wiley Online Library, 2018.

[32] A. Zakaria, I. Jeffrey, J. LoVetri, and A. Zakaria, "Full-vectorial parallel finite-element contrast source inversion method," *Progress In Electromagnetics Research*, vol. 142, pp. 463–483, 2013.

[33] K. G. Brown, N. Geddert, M. Asefi, J. LoVetri, and I. Jeffrey, "Hybridizable discontinuous galerkin method contrast source inversion of 2-d and 3-d dielectric and magnetic targets," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 5, pp. 1766–1777, 2019.

[34] C. Geuzaine and J.-F. Remacle, "Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities," *International journal for numerical methods in engineering*, vol. 79, no. 11, pp. 1309–1331, 2009.

[35] N. Geddert, *An electromagnetic hybridizable discontinuous Galerkin method forward solver with high-order geometry for inverse problems*, 2020.

[36] K. Nemez, A. Baran, M. Asefi, and J. LoVetri, "Modeling error and calibration techniques for a faceted metallic chamber for magnetic field microwave imaging,"

IEEE Transactions on Microwave Theory and Techniques, vol. 65, no. 11, pp. 4347–4356, 2017.

[37] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The computer journal*, vol. 7, no. 4, pp. 308–313, 1965.

[38] M. Ostadrahimi, P. Mojabi, C. Gilmore, A. Zakaria, S. Noghanian, S. Pistorius, and J. LoVetri, "Analysis of incident field modeling and incident/scattered field calibration techniques in microwave tomography," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 900–903, 2011.

[39] I. T. Rekanos, "Neural-network-based inverse-scattering technique for online microwave medical imaging," *IEEE transactions on magnetics*, vol. 38, no. 2, pp. 1061–1064, 2002.

[40] Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 1849–1860, 2018.

[41] L. Li, L. Wang, F. Teixeira, L. Che, and T. Cui, "Deepnis: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Transactions on Antennas and Propagation*, vol. PP, pp. 1–1, 12 2018.

[42] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Transactions on Antennas and Propagation*, 2020.

[43] V. Khoshdel, M. Asefi, A. Ashraf, and J. LoVetri, "A multi-branch deep convolutional fusion architecture for 3d microwave inverse scattering: stored grain application," *Neural Computing and Applications*, 2021. [Online]. Available: https://doi.org/10.1007/s00521-021-05970-3

[44] R. Kruk, *Deep learning techniques applied to the remote sensing of soil moisture and sea ice type*, 2021.

[45] V. Khoshdel, *Using deep learning approaches for microwave imaging*, 2021.

[46] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014.

[47] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," *arXiv preprint arXiv:1711.04340*, 2017.

[48] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Icml*, 2010.

[49] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv preprint arXiv:1811.03378*, 2018.

[50] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, vol. 30, no. 1. Citeseer, 2013, p. 3.

[51] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[52] K. Edwards, N. Geddert, K. Krakalovich, R. Kruk, M. Asefi, J. Lovetri, C. Gilmore, and I. Jeffrey, "Stored grain inventory management using neural-network-based parametric electromagnetic inversion," *IEEE Access*, vol. 8, pp. 207 182–207 192, 2020.

[53] M. Asefi, G. Faucher, and J. LoVetri, "Surface-current measurements as data for electromagnetic imaging within metallic enclosures," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 11, pp. 4039–4047, 2016.

[54] C. Gilmore, M. Asefi, K. Nemez, J. Paliwal, and J. LoVetri, "Three dimensional radio-frequency electromagnetic imaging of an in-bin grain conditioning process," *Computers and Electronics in Agriculture*, vol. 167, p. 105059, 2019.

[55] M. Asefi, C. Gilmore, I. Jeffrey, J. LoVetri, and J. Paliwal, "Detection and continuous monitoring of localised high-moisture regions in a full-scale grain storage bin using electromagnetic imaging," *Biosystems Engineering*, vol. 163, pp. 37–49, 2017.

[56] K. Edwards, K. Krakalovich, R. Kruk, V. Khoshdel, J. LoVetri, C. Gilmore, and I. Jeffrey, "The implementation of neural networks for phaseless parametric inver-

sion," in *2020 XXXIIIrd General Assembly and Scientific Symposium of the International Union of Radio Science*. IEEE, 2020, pp. 1–3.

[57] A. P. Turner, M. D. Montross, J. J. Jackson, S. G. McNeill, M. E. Casada, J. M. Boac, R. Bhadra, R. G. Maghirang, and S. A. Thompson, "Error analysis of stored grain inventory determination," *Transactions of the ASABE*, vol. 59, no. 3, pp. 1061–1072, 2016.

[58] K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, and I. Jeffrey, "A machine learning workflow for tumour detection in breasts using 3d microwave imaging," *Electronics*, vol. 10, no. 6, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/6/674

[59] N. K. Nikolova, "Microwave imaging for breast cancer," *IEEE microwave magazine*, vol. 12, no. 7, pp. 78–94, 2011.

[60] M. Omer, P. Mojabi, D. Kurrant, J. LoVetri, and E. Fear, "Proof-of-concept of the incorporation of ultrasound-derived structural information into microwave radar imaging," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 3, pp. 129–139, 2018.

[61] R. C. Conceição, J. J. Mohr, M. O'Halloran *et al.*, *An introduction to microwave imaging for breast cancer detection*. Springer, 2016.

[62] C. Gilmore, A. Zakaria, S. Pistorius, and J. LoVetri, "Microwave imaging of human forearms: Pilot study and image enhancement," *International journal of biomedical imaging*, vol. 2013, 2013.

[63] K. Edwards, V. Khoshdel, M. Asefi, J. LoVetri, C. Gilmore, and I. Jeffrey, "Recovery of prior information for breast microwave imaging using neural networks," in *Accepted to URSI GASS 2021*, 2021.