MACHINE LEARNING BASED EFFICIENT AUTOMATED NDE METHODS FOR DEFECT DIAGNOSTICS

By

Subrata Mukherjee

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical Engineering—Doctor of Philosophy

2023

ABSTRACT

Non-destructive evaluation (NDE) assumes a crucial role in examining infrastructure across diverse industries that require ongoing monitoring to uphold product quality, sustain aging infrastructure, and ensure operational equipment safety. Recent advancements in data acquisition (sensing technologies) and data evaluation (data analytics) have prompted a significant shift in the NDE community towards fully automated, data-driven inspection routines. The primary goal of this dissertation is to develop intelligent, state-of-the-art NDE-based artificial intelligence (AI) methods. This research involves incorporating NDE theory and insights to adapt modern machine learning (ML) and AI algorithms, enabling efficient decision-making processes with high accuracy and reliability, while minimizing human intervention. Several cost-effective, miniaturized NDE sensing methodologies based on techniques like Magnetic Flux Leakage (MFL), Eddy Current (EC), Capacitive Imaging (CI), and laser profilometry are developed based on the materials being tested. The research focuses on the following key aims: (a) Developing spatially adaptive denoising algorithms to address uncertainties in the data collection process, enabling accurate defect detection and localization, (b) Dynamically updated defect tracking through mixture regression and optimally binned hypothesis testing, (c) Establishing an integrated diagnostic framework for inline inspection and automated defect characterization using a hybrid deep learning setup, (d) Efficient data augmentation and fusion techniques to combine information from multiple heterogeneous sensors, and improving defect diagnostics by twinning/registering experimental and simulated data through transfer learning principles, (e) Implementing sophisticated data compression techniques to achieve cost-effective and fast defect diagnostics without compromising efficiency.

This dissertation develops new-age integrated NDE frameworks for data acquisition and evaluation, providing accurate and reliable defect diagnostics. NDE data used in practical applications often suffers from noise contamination and errors caused by various factors. In chapter 2, a Bayesian decision theory-based approach is established to understand the noise corruption in capacitive sensing data. A solely data driven spatially adaptive denoising algorithm is modelled that does not need any oracle information. In chapter 3, an integrated inline inspection (ILI) and automated defect

detection framework for plastic pipelines using laser profilometry data and a hybrid deep learning approach are established. Chapter 4 proposes an under-sampling compressed scheme for fast and cost-effective scanning, providing accurate diagnosis. Spatially adaptive methods for identifying defect growth using dynamically updated transfer learning are introduced in subsequent chapters 5 and 6, leveraging information from less noisy baseline scan. Bivariate function estimation using mixture regression (TLMR) and transfer learning-based binned hypothesis testing (TLBH) are incorporated for tracking defect growth on later noisy scans. An inverse data-driven approach for estimating electrical parameters of the substrate is proposed in chapter 7, along with registration-aided machine learning models that adaptively utilize information from large synthetic datasets to improve predictions on scarce experimental data. Finally, an integrated multi-modal fusion setup is designed in chapter 8 for enhanced defect detection, incorporating optimal transport (OT) based registration of data from different sensing modalities.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of people who stood by me during my years at Michigan State. I would like to convey my special thanks to all of them.

I am indebted to Professor Yiming Deng for his encouragement, mentorship and unwavering attention. I have gained valuable knowledge and experiences from him, both in academic and personal front. He has been extremely enthusiastic throughout my Ph.D. years that enable me to assimilate new ideas in the field of Nondestructive evaluation and applied machine learning. I am immensely grateful for the freedom that Professor Deng gave me and for his support in working on a plethora of topics. This thesis owes its debt to Professor Lalita Udpa for deeply influencing the research trajectory. Her valuable feedback on my work helped my ideas blossom and come into life. I would also like to convey my profound gratitude to Professors Satish Udpa and Tapabrata Maiti for their help and support during my Ph.D. life. I greatly enjoyed discussions with Professor Satish Udpa and am obligated to him for his time and patience.

My appreciation goes out to all the members of the NDEL research group for creating such an intellectually stimulating environment. We have been each others support throughout the years and I hope that our friendships will continue for many years to come. My stay at East Lansing was very enjoyable because of their awesome company.

Finally, and most importantly, earning a Ph.D. would have been stayed a distant dream if it weren't the never-ending support and unconditional love of my family: my parents, brother, Rini didi and our little Uma. I would like to thank my parents for their sacrifices in raising us, two brothers. It is my brother who always galvanize me and have taught me not to take anything for granted. He has planted a seed of love for learning in me-I dedicate this thesis to him. You all have made this work worthwhile!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Non destructive evaluation its usage contemporary applications

Non-destructive evaluation (NDE) plays a vital role in inspecting infrastructure across various industries where continuous monitoring is necessary to ensure product quality, maintain aging infrastructure, and guarantee the safety of operational equipment. The primary objective of non-destructive testing is to assess the condition of the test subject in a non-invasive manner, facilitating risk management. NDE methods effectively mitigate the risk of damaging critical infrastructure while it remains in place, unlike destructive testing methods that render subjects unfit for further use. Non-destructive testing proves cost-effective when the potential risk of failure outweighs the testing expenses. It finds application in numerous industries where failures could pose threats to human life and property. Figure 1.1 (a) shows the importance of NDE in life cycle of product assessment, while figure 1.1 (b) showcases a wide range of industrial sectors such as aviation, naval, medicine, advanced manufacturing, pipeline, and railway, among others, where NDE inspections are employed on a daily basis. Various NDE-based inspection methodologies have been developed, relying on the sensitivity of materials to electric, magnetic, electromagnetic, mechanical, chemical changes, or a combination thereof. Examples of popular techniques include magnetic flux leakage (MFL), eddy current (EC), radiography testing (RT), thermography testing, ultrasonic testing (UT), acoustic testing, electromagnetic acoustic transducer (EMAT), microwave and millimeter wave testing, laser, capacitive imaging, vibration, leak, and liquid penetrant testing (1; 2; 3; 4; 5; 6; 7; 8; 9; 10; 11; 12).

NDE methods are extensively utilized in the power and transportation industries. The nuclear industry in France allocates approximately 1.5 billion euros per year to maintenance, which accounts for 17% of its total costs whereas maintenance of civilian aircraft contributes to around 20% of the ticket prices. Several thousand casualties have been reported in the past decade due to infrastructure collapses. In the United States, the Pipeline and Hazardous Materials Safety Administration (PHMSA) has reported numerous pipeline failures resulting in billions of dollars in losses over the last decade (13; 14), highlighting the need for proper NDE inspections. Recent news reports

have highlighted the implosion of the submarine "Titan" as another unfortunate incident that could have been prevented with proper NDE inspection. Such incidents serve as a poignant reminder of the critical importance of NDE in ensuring the safety and reliability of various structures and equipment that are in practice. Due to the significant volume of NDE inspections necessary, there is an imperative need to develop monitoring services that offer cost-effective, rapid, accurate, and dependable solutions. The sheer scale of inspections demands efficient and streamlined processes to ensure timely and comprehensive assessments while maintaining affordability. Developing such monitoring services would not only enhance productivity and efficiency but also enable more widespread and frequent inspections, thereby improving overall safety and reducing the risk of failures. By combining speed, accuracy, and reliability, these services can effectively address the demanding requirements of NDE inspections and facilitate proactive risk management in a cost-efficient manner. To meet these contemporary demands, the field of NDE is experiencing a rapid revolution, transitioning from traditional approaches to robust automated data-driven **"smart inspection"** routines known as NDE 4.0, coined by the NDE community.

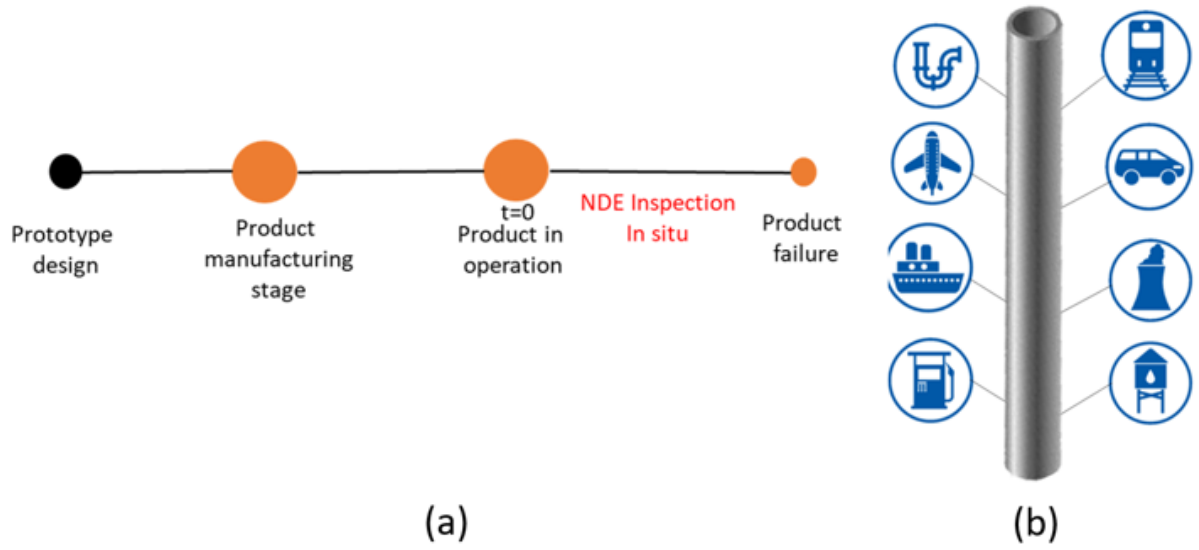

Figure 1.1 Various industrial applications of NDE.

1.2 Machine Learning, the need of automation in NDE & Objective

The advancement in wireless sensing technologies has facilitated the collection of massive volumes of sensing data. Industry 4.0 aims to achieve smart machines, smart production, and smart services, and the deployment of intelligent non-destructive sensor systems has resulted in the generation of large amounts of real-time data. Addressing this challenge, the NDE community urgently requires automated approaches for designing high-fidelity, data-driven models to analyze the reliability of structures with minimal human intervention. The development of new automated detection and tracking algorithms, utilizing Artificial Intelligence (AI) and Machine Learning (ML), is crucial to effectively process and extract insights from these extensive modern data repositories. Manual analysis is time-consuming and costly due to the scale of the collected data. Real-world NDE applications involve numerous uncertainties, such as lift-off, probe drift, variations in data acquisition sampling rates and scanning speed, complex sample structures, limited sensor resolution and sensitivity to the material under test (MUT). Characterizing noise in such a broad spectrum of uncertainties becomes challenging. Consequently, robust online defect detection algorithms that can handle noise effectively are required. In this thesis, several robust algorithms are developed to achieve accurate defect detection under these uncertainties, featuring the key aspect of noise agnosticism. The main objective of this dissertation is to design NDE-based sensing methodologies tailored to the inspection material. Subsequently, various statistical models based on AI and ML are developed using these modern datasets, targeting diverse deployment platforms in NDE as shown in figure 1.2. These models exhibit high precision in identifying defect characteristics. Effective diagnosis through NDE inspection typically involves five stages, which are elaborated upon in the subsequent chapters of this report.

1. Detection: Detection provides a qualitative indication regrading the presence of damage in structure.
2. Localization: This gives information regrading the probable position of the damage.
3. Classification: Classification states the probable types of damage.

4. Assessment: Assessment bestows an estimate of the extent of the damage and perform tracking of the damage.

5. Prediction: Prediction offers information about the safety of the material under test . For example, it can give an estimate of the remaining useful life (RUL).

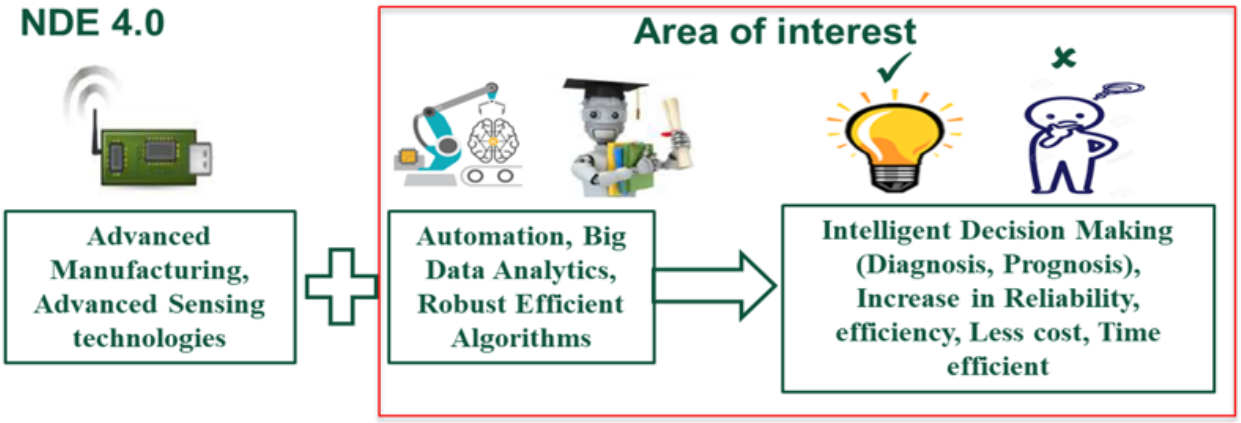

Figure 1.2 (a) Schematic showing the advancement of NDE 4.0 and the area of interest of this dissertation.

In this thesis we developed AI and machine learning based modern NDE methods that can conduct cheap, fast yet accurate and reliable defect diagnostics. The methods greatly improve the efficacy obtained by traditional methods in these new-age applications. The methods developed in this thesis uses the following techniques:

1. Sophisticated data compression for cheap and fast defect diagnostic without any efficiency loss due to compression;

2. Using flexible deep learning networks for automating NDE based defect localization and classification in very complex systems where there is no established parametric NDE models;

3. Efficient data augmentation rules for combining information collected by multiple sensors as well as improving defect diagnostics by augmenting experimental data with simulated data via transfer learning principles.

We briefly elaborate these attributes below.

Sparse signal processing is recognized as a powerful technique for extracting relevant data in various modalities utilized in NDE. NDE signals exhibit a high level of structure, resulting in a significant portion of the signal's energy being sparsely represented. This sparse representation enables data compression in specific domains, leading to a substantial reduction in data volume and facilitating real-time online operations. In addition to temporal structure, spatial correlation within the signals is also significant, making spatial sub-sampling a favorable approach. Sparse reconstruction methods are commonly employed in the field of medical imaging to achieve efficient and accurate non-invasive monitoring. Presently, the NDE community is rapidly adopting sparse exploitation techniques as well (15; 16; 17). Sparse image reconstruction proves beneficial in scenarios requiring fast and time-efficient scanning, as well as situations where complete access to the target being scanned is limited. This dissertation investigates different sub-sampling-based image reconstruction algorithms, exploring their applicability in NDE.

To achieve automated defect detection and localization in applications without explicit parametric setup, the integration of deep learning with NDE knowledge proves valuable. Shallow machine learning (ML) models and deep learning (DL) models, such as Convolutional Neural Networks (CNNs) and U-Nets, are being employed in NDE (18; 19; 20; 21; 22; 23; 24; 25) to learn from sensor signals. However, widespread online application of DL in NDE encounters various obstacles, primarily due to the limited availability of experimental training data. Obtaining a sufficient and representative NDE dataset under realistic operational conditions is often impractical. One solution to this challenge is to effectively learn from synthetic databases and perform intelligent and meaningful data augmentation. However, blindly augmenting data without considering the differences between the training and testing data sources can lead to erroneous results. Thus, the development of AI algorithms in NDE necessitates the inclusion of expert knowledge rather than treating ML models as black boxes. The process of understanding these models is referred to as "Explainable AI" as depicted in figure 1.3. In Chapter 3 and 7 of this thesis, efficient NDE inspection frameworks equipped with AI are proposed, addressing these considerations. In many

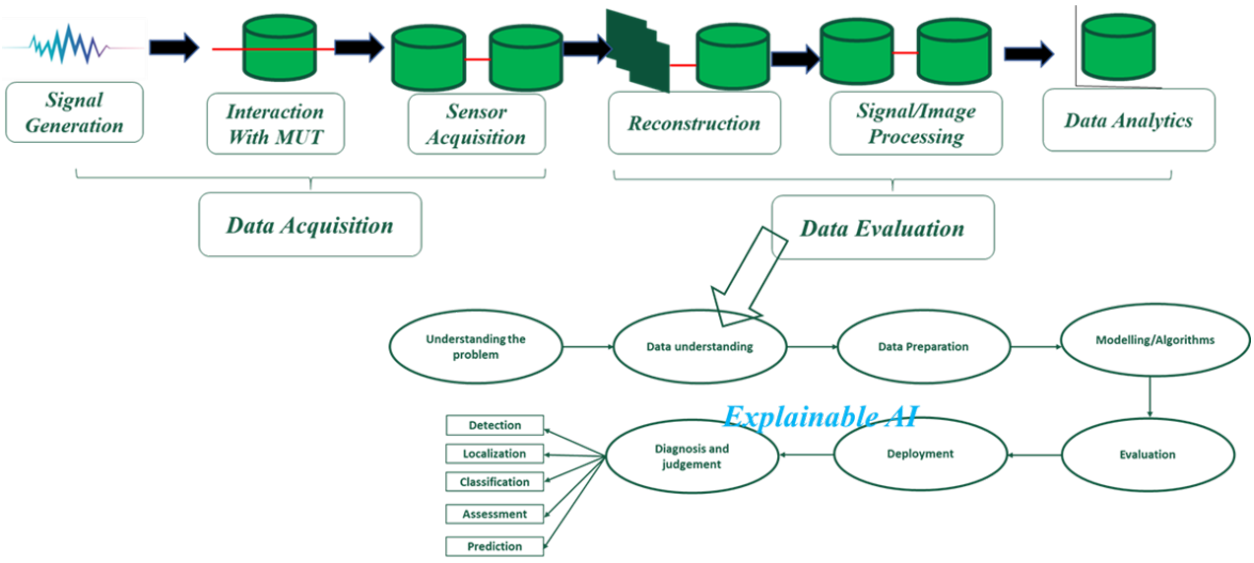

Figure 1.3 Steps involved in explainable, interpretable AI.

cases, a single measurement technique may have limitations in effectively characterizing the health of a component due to constraints in resolution. Different sensing modalities, on the other hand, exhibit sensitivity to various types of degradation, thus offering complementary information that surpasses the capabilities of a single method. The utilization of data fusion techniques holds promise in enhancing non-destructive evaluation (NDE) by increasing the accuracy and reliability of diagnostic outcomes. However, it is important to consider boundary conditions where individual methods may fail, as blindly fusing data in such cases can lead to suboptimal results. Additionally, the presence of heterogeneous data sources introduces the challenge of mis-registration, which the NDE community must address to improve the reconstruction of anomalies.

1.3 Brief description of the chapters

This thesis aims to increase our current understanding and to provide efficient solution in the following areas of modern NDE:

1. Flexible deep learning networks for automating NDE based defect localization and classification in very complex systems where there no established parametric NDE models exists.

2. Developing a host of spatially adaptive denoising algorithms to weed out the various uncertainty induced noises and thereby providing accurate defect detection and localization.

3. Need of optimal binned hypothesis tests to track dynamic defect growth under uncertainties.

4. Sophisticated data compression using principles of compressed sub sampling and kriging for cheap and fast defect diagnostics without any efficacy loss due to compression.

5. Developing heterogeneous sensor fusion and data augmentation algorithms to reduce the uncertainty of individual sensor signal, thereby producing reliable diagnosis.

6. Establishing different registration aided regression frameworks for material characterization that can effectively answer 'what' to transfer and 'how' to transfer from large synthetic (simulated) dataset to scarce experimental data.

The thesis provides accurate diagnostic algorithms that can be classified into two broad sub-groups, (a) spatially adaptive denoising algorithms, (b) different sub sampling and transfer learning based defect tracking schemes. The organization of the subsequent chapters are as follows:

1. In Chapter 2, we develop a high-Q capacitive sensing based robust non-destructive evaluation (NDE) methodology that can be widely used in varied NDE applications. We show that the proposed method can detect defects in a host of robust regimes where uncertainties such as lift-off, probe tilt, fluctuations in sampling rates and step sizes are inherent in the data collection process. We explicitly characterize the corruption in the capacitive sensing data due to various lift-off-based uncertainties. We use a Bayesian decision theoretic approach to rigorously understand the impact of these corruptions on defect identification efficacy. Using an optimally tuned weighted classification loss, we prove that it is theoretically feasible to accurately detect defect location and sizes from capacitive sensing signals collected under the aforementioned uncertainties. The Bayesian decision theoretic study needs prior information for accurate detection that is not available in real NDE inspections. So, we develop a solely data driven algorithm that analyzes the capacitive sensing signals without any prior knowledge of defect or uncertainty types. The developed algorithm is non-parametric and

uses spatially adaptive denoising to weed out uncertainty induced noises. By leveraging the spatial association in the capacitive sensing signals, our algorithm greatly improves on popular non-spatial approaches. Compared to popular thresholding methods and low-rank based denoising approaches we demonstrate superior performance of the proposed method in terms of coverage and false positive metrics for defect identification. Using spatially adaptive denoising we design a robust capacitive sensing method that can detect defects with high precision under various uncertainty regimes.

2. In chapter 3, an integrated inline inspection and automated defect detection using deep learning framework has been established. Here, we develop a novel NDE methodology for inspecting plastic pipelines used in natural gas distribution. Our proposed method is based on optical imaging and uses laser profiling to collect inline inspection data. In our laser profiling images, deformities due to defects lead to local perturbations in the circular rings produced by the otherwise non-defective inner walls of cylindrical pipes. Using experimental data collected by the proposed methodology, we studied the performance of a gamut of defect classification algorithms including conventional machine learning classifiers as well as deep learning architectures. Special attention was paid to take into account both the spatial and temporal features in our laser-scan image, particularly exploiting the circular shape of non-deformed laser rings. This led to the development of a hierarchical bilinear pooling (HBP) based deep learning framework that can precisely identify local perturbations in our laser-scan data by assimilating information across multiple cross-layer features. Equipping our endoscopic laser profiling system with the HBP based defect classification algorithm, we develop an integrated diagnosis method for inline data collection and automated defect characterization. Based on experimental data collected under varying profiling conditions, we demonstrate superior performance of our proposed integrated method in damage localization and classification.

3. Chapter 4 shows the applicability of sub sampling scheme based on kriging in the field of Nondestructive evaluation. Magnetic flux leakage (MFL) testing is a widely used non-

destructive evaluation (NDE) technique for defect detections within the pipelines, particularly those composed of ferromagnetic materials. Pipeline inspection gauge (PIG) procedure based on line-scans can collect accurate MFL readings for defect detection. However, in real world applications involving large pipe-sectors such extensive scanning techniques are extremely time consuming and costly. In this chapter, we develop a fast and cheap methodology that does not need MFL readings at all the points used in traditional PIG procedures but conducts defect detection with similar accuracy. We consider an under-sampling based scheme that collects MFL at uniformly chosen random scan-points over large lattices instead of extensive PIG scans over all lattice points. Based on readings from the chosen random scan points, we use Kriging to reconstruct MFL readings. Thereafter, we use thresholding-based segmentation on the reconstructed data for detecting defective areas. We demonstrate the applicability of our methodology on synthetic data generated using finite element models as well as on MFL data collected via laboratory experiments. In these experiments spanning a wide range of defect types, our proposed novel MFL based NDE methodology is witnessed to have operating characteristics within the acceptable threshold of PIG based traditional methods and thus provide an extremely cost-effective, fast procedure with competing error rates.

4. Chapter 5 presents a novel spatially adaptive method for identifying defect growth using dynamically updated transfer learning technique on data from magnetic flux leakage (MFL) sensors. The operation of pipeline inspection gauge (PIG) within the pipeline to collect accurate, low noise readings for defect detection is expensive and time-consuming. Running probes within the operational pipeline produces noisy data. In this chapter we consider a less noisy and time-consuming baseline readings within pipelines taken in the beginning. Using the baseline data, our goal is to first automatically detect the defective areas during inspection and thereafter monitor the growth of those defects and raise a flag of caution if they become too large and detrimental. Based on the baseline data, a bivariate function was estimated using a function estimation method based on mixture regression framework to compute posterior probabilities of the defects at each scanning point. Thereafter, it is seen

that applying direct function estimation with noisy field data on subsequent inspections is not effective. We use transfer learning perspectives by leveraging the defect probabilities and location from the previous inspections, and then consequently update those probabilities based on current data by applying a dynamically updated transfer learning technique. The defect growth is dynamically tracked and characterized with high accuracy and sensitivity.

5. In Chapter 6, a transfer learning aided optimally binned hypothesis testing framework has been developed for dynamic defect tracking under noisy conditions. A key characteristic of the proposed dynamic monitoring method is its noise type agnosticism in the low-cost NDE scans. The gathering of low-noise, precise non-destructive evaluation (NDE) readings for defect detection can be both time-consuming and costly. In practical applications, NDE data can be highly noisy due to several unwanted. Given the broad spectrum of potential uncertainties that can contribute to noise accumulation, characterizing the noise in such NDE data presents significant difficulties. As a result, the extraction of defect signals through deconvolution becomes challenging. This chapter introduces an NDE-based dynamic defect tracking framework designed for robust conditions. It incorporates regular periodic monitoring of the material under test (MUT) using a cost-effective, easy-to-implement magnetic flux leakage (MFL) probe. The framework proposes a dynamic setup that includes occasional, highly accurate, expensive MFL scans among regular, low-cost, noisy MFL scans. The objective here is to detect and precisely track defect formation in metallic pipes over time, thereby enabling the efficient alarming and localization of defects before they reach detrimental sizes. The proposed method remains effective even when the underlying noise-producing uncertainties fluctuate significantly over time, a feature achieved through the use of transfer learning. The method transfers accurate information regarding the locations and sizes of existing corrosions from the detailed MFL scan to the subsequent noisy MFL inspections. With this supportive information from defective scan-points, it estimates the degradation of the signal-to-noise ratio (SNR) in the noisy MFL scans. The framework suggests binned hypothesis tests in noisy MFL scans, setting the level of aggregation based on the SNR estimate in the scan

data. This proposed binned hypothesis test optimizes defect coverage and maintains control over the false discovery rate (FDR). Experimental MFL data, gathered under a wide range of uncertainties and defect types, illustrates the highly heterogeneous nature of noise distributions in low-cost MFL scans and the significant variance in SNR deteriorations. The necessity of using binned hypothesis tests in these low-cost MFL data is highlighted by the poor performance of pointwise multiple hypothesis tests. Finally, the high effectiveness of the proposed transfer learning-based binned hypothesis testing method is demonstrated. Although the method is applied to MFL data in this instance, it is also suitable for other NDE applications. The applicability of the algorithms has also been confirmed through validation on experimentally generated Eddy Current (EC) data.

6. Chapter 7 consists of two discrete works. In the first work, an adaptive denoising scheme based on nearest neighbor divergence measure in wavelet transform domain is proposed which is sensitive in detection of coincident multiple defects. It is also cost effective by not classifying most of the harmless cavities as harmful defects. Knowledge of the electrical properties, such as complex permittivity, permeability and loss tangent measurements is rapidly becoming a necessity for Nondestructive Evaluation (NDE) based material characterization. In the next work, we aim to provide a data- driven approach to estimate the wideband dielectric permittivity for a given substrate material based on the frequency responses from microstrip transmission lines fabricated with the material. We demonstrate registration-aided machine learning models that adaptively use information from large simulated datasets to make improved predictions on experimental data where we have acute data scarcity. Machine learning (ML) models are trained using simulation data for several unique combinations of substrate and microstrip line dimensions and is tested on experimental data where the microstrip line are fabricated on eleven different unknown substrates. The S parameters associated with the reflection and transmission coefficients are treated as functional data across the frequency sweeps. As we had very few experimental data, along with complex non-parametric methods, we also consider low-complexity models on the frequency curves.

In this aspect, dimensionality reduction techniques are considered to deal with situations in the experimental data where the number of features obtained from the frequency sweeps are much higher than the number of samples in the experimental data. We compare the efficacy of data-hungry machine learning methods with these low-complexity models. As the source of train and test data are different, registration strategies based on intercept correction are implemented. We illustrate the efficacy of registration-based varied ML techniques for lab generated experimental data and obtained encouraging results. This work is an attempt to bypass material characterization models of electromagnetic (EM)-physics that is based on closed form mathematical equations and have the limitations that they can only be applied in idealized set-ups.

7. In chapter 8, single modality-based Nondestructive Evaluation (NDE) data used in practical applications often suffers from noise contamination and errors caused by various factors like lift-off/standoff distances, probe drift, scanning speed, variation in data acquisition rates, and poor sensor sensitivity. The presence of agnostic noise types poses a challenge in extracting defect signals and understanding damage characteristics using single NDE modalities alone. To address these challenges, in this chapter we present an automated NDE theory-based data fusion framework aimed at enhancing the detection of surface and near-surface defects in magnetizable and conductive specimens. Both experimental and simulation data obtained from Magnetic Flux Leakage (MFL) and Eddy Current (EC) based NDE sensing methods demonstrate the highly heterogeneous nature of noise distributions. Given the heterogeneity of the inspection methods, a screening rule is proposed to determine the conditions under which fusion would be beneficial. An important aspect of the proposed fusion method is registration, which ensures accurate alignment of multi-sensor image data. Two registration methods are proposed in this study to address the issue of different inspection sources. Performing blind fusion without registration leads to erroneous results. The first registration method is translational, where one source image obtained from a particular sensing technique is aligned with an anchored source image from another technique. The second method

is registration based on linear optimal transport (OT) which proves to be effective in the boundary conditions as it provides a mathematical framework to transfer one probability source distribution into another while minimizing the cost of transportation. Finally, the registered source images from the EC and MFL modalities are fused using pixel-based fusion algorithms, including transform domain and spatial domain-based methods. Sparse representation, wavelet decomposition, and bilateral filter-based fusion techniques are implemented in this study. Qualitative and quantitative assessments demonstrate that the registered fusion results exhibit higher accuracy and reliability compared to unregistered fused results and source images. Although the fusion method is applied to MFL and EC data in this paper, it is also suitable for other NDE modalities.

8. Chapter 9 concludes with the contribution of this research in the field of Non destructive evaluation.

CHAPTER 2

ROBUST DEFECT DETECTION UNDER UNCERTAINTIES USING SPATIALLY ADAPTIVE METHODS

2.1 Introduction

Structural integrity of safety-critical infrastructures such as aircrafts, bridges, nuclear power plants, oil and gas transmission pipelines and load-bearing metal structures in highways decrease with time. Nondestructive evaluation (NDE) plays an important role in maintenance of these infrastructures by detecting significant erosions and other defects that crop up with usage. For providing effective diagnosis, NDE methods use several inline inspection techniques to access these structures without causing any further damage in a non-invasive manner (26). The advancement in sensing technology has led to the development of different NDE sensing modalities (27; 28; 29; 30; 31; 11; 32; 33; 10; 34; 9; 8; 35). However, often in practice, it is extremely difficult to set-up new non-destructive inspection procedures abiding the ideal laboratory conditions under which these methods were developed and tested. As a result, in these situations, NDE data are collected under uncertainties such as lift-off, probe tilt, fluctuations in sampling rates and step sizes (36; 37). Contemporary NDE methods based on robotic systems can conduct very fast scans but often suffer from the aforementioned uncertainties due to anomalies in probe positioning. Lift-off is the variation of distance between the probe and material under test (MUT) whereas probe tilt is the angular deviation of probe from perpendicularity. NDE data are often contaminated by the presence of unknown lift-offs due to coarseness of MUT surface, non-conducting coatings or irregular paintings on the MUT (38; 7). While lift-offs and probe tilts are the most common uncertainties, fluctuations in sampling rate and step sizes can also contaminate NDE data due to operator's movement related discrepancies such as changes in frictional force between probe and material surfaces as well as misalignment of probe and MUT coordinates. These uncertainties in the data collection process give rise to different kinds of corruption in the recorded signals. Defect identification algorithms that were developed with sensor data collected under ideal laboratory conditions, fail to provide accurate defect analysis under these uncertainties which corrupts data

with various kinds of noise contaminations. Subsequently, there is an urgent need to develop robust, automated, reliable defect identification methods that can operate with high precision under uncertainties that practitioners encounter in real-world applications (6).

Here, we develop a high-Q capacitive sensing based shortwave NDE method which is low cost, highly flexible and easily implementable as it does not require extensive hardware setup at MHz frequencies. Unlike other popular NDE methods, it can be employed under a wide range of robust conditions for it enjoys the following advantageous properties over the competing NDE methods: (a) unlike capacitive sensing, ultrasonic testing (UT) normally requires a coupling medium and physical contact with MUT (32). Air coupled UT, however, does not need coupling medium. But, it uses long sound pulse for excitation for which accurate measurement of timing is difficult. Moreover, it is not applicable for metal inspections due to high acoustic impedance of metal structures (5); (b) Thermographic methods have low sensitivity unless we use very expensive thermal camera for imaging and are non-robust to temperature variations (4; 1). X-ray imaging is not only expensive but also needs proper screening because of the detrimental ionizing radiations (33); (c) Microwave imaging based NDE detection do not penetrate deep in conducting materials and use high frequency for operation in GHz ranges which require cumbersome data acquisition circuitry (2); (d) Conventional electro-magnetic (EM) based NDE techniques such as magnetic flux leakage (MFL) and eddy-current (EC) methods are mostly effective in inspecting highly conductive or ferromagnetic materials (27; 28; 29; 10; 34; 9; 8; 35).

We concentrate on the highly flexible, easily applicable capacitive sensing method and undertake a disciplined approach to understand how signals collected from defective samples get corrupted as uncertainties such as lift-off, speed and fluctuations in sampling rates are introduced in the data collection process. In next section, we describe the data collection methodology as well as the different robust set-ups considered in this study. Existing NDE literature mainly focused on the effect of lift-offs based uncertainties in EC and MFL methods. Lift-off corrections in these traditional EM techniques have been developed based on novel hardware-based probes (3; 39), new sensor array designs (40) and by usage of signal processing techniques (38; 41). Novel probing

methods introduced in (38) uses reference signals to cancel lift-off induced contaminations in pulse eddy current (PEC) signals whereas dual frequency mode of EC inspection is used in (42) for measuring lift-off and accurate defect sizing. Signal processing techniques based on dynamic trajectories of fast Fourier transform of PEC scans are used to reduce the lift-off effects in (43) and features extracted based on rise time, zero crossing, differential time to peak are used in (44). Unlike traditional EM methods, there has been very limited research on the scope of capacitive imaging (CI) under uncertainties. Specialized capacitive sensors have been designed to tackle lift-off effects (45; 46; 47; 48) but they have limited applications. Here, we develop a very easy-to-implement capacitive sensing probe and provide a systematic study on the impact of uncertainties on signals captured by it. Based on the study we develop a novel spatial adaptation based signal processing algorithm that is exhibited to provide very high defect identification rates. The resulting NDE method is cost effective, less time consuming and can be employed across a wide range of applications for it does not require any domain specific modifications or usage of complex probes or guidance of NDE experts.

In later section, using a disciplined statistical framework we compare the voltage readings recorded by the capacitive sensors under different kinds of uncertainties. We show that the corruption in the voltages due to uncertainties which decreases signal strength for defect defection is not merely due to homogeneous addition of noise over all the inspection readings. We analyze the differences in the capacitive sensing voltage readings collected under various kinds of lift-offs by Gaussian mixtures models. In each case, we document the target-to-background voltage comparisons where the target is the defective area amidst the background of non-defective scan points. Based on this mixture model framework, in Section 2.4 we report the theoretically possible lowest limit on the defect misclassification error rates across varied uncertainties by considering the Bayes estimator. In most NDE applications, it is much more harmful to misclassify defects as non-defects (which is termed as false negatives (FN) or 'missing') than to misclassify non-defects as defects (which is termed as false positives (FP) or 'false alarm'). In this context, we evaluated the Bayes error under the weighted classification loss and reported the minimum achievable FP

rates when the FN rates are constrained below 10% i.e., when we, at worst, underestimate 10% of the defect sizes. Based on experimental data collected under different types of uncertainties we found that the FP rates were quite low even when FN rates are controlled below 10%. This suggests that it is theoretically possible to recover defect sizes under the uncertainties considered here.

Computation of the Bayes rules and errors reported in following sections needed pertinent information on defects that is not available for real applications. Thus, those error rates mark the best theoretically possible errors that any non-spatial decision rule can obtain. In practice, it is very difficult to get a solely data driven procedure that achieves such error rates. In Section 2.5, we propose a data driven classifier of defective scan points. The classifier borrows information across geographically close scan points and is spatially adaptive. Aggregating the data from nearby locations can increase the background-to-signal ratio and increase the defect detection rate under certainty. The proposed approach non-parametrically recovers the structure of the defects without prior knowledge on the uncertainties in the data collection process. We explain the benefits of spatial adaptation and show that the proposed method yields accurate defect identification in highly robust conditions. Figure 2.1 shows the framework of the proposed study. In order to provide a disciplined analysis of robust defect identification using capacitive sensing we further divide the study into the following components:

1. The first stage involves collection of experimental data using common place capacitive sensing methods. In this stage as described in Section IIA, we set-up a cheap capacitive-sensing probe with its parameters tuned to operate optimally in usual non-robust laboratory settings.
2. Thereafter, varied types of uncertainties are systematically introduced in the data collection process through an experimental design (see Table I) that involves 36 experiments on defective samples containing a single defect of varying sizes.
3. We compute the Bayes error in defect misclassification to understand the impact of corruption in the voltage signals recorded by capacitive sensing under uncertainties. The nature of the corruption under different types of uncertainties were characterized.

4. Based on the above characterization a data driven method was developed to adaptively denoise capacitive sensing based voltage signals. The proposed non-parametric method uses spatial filtering and is exhibited to detect defective scan points with very high accuracy.

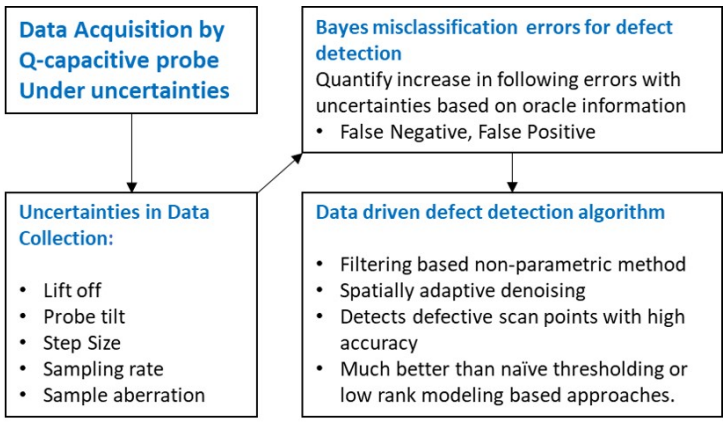

Figure 2.1 Framework used to develop robust capacitive sensing methodology under uncertainty.

2.2 Experimental Data

2.2.1 Capacitive-sensing data collection methodology

We use a capacitive sensing-based probe that consists of a parallel plate capacitor along with a coupled inductor coil. The capacitive probe detects the change in the dielectric constant due to distortion in electrostatic fields in presence of defects in the sample. The voltage recorded by the probe across a grid of scan points produces image that can lead to defect identification (11; 49; 50).

To understand the working principle of the probe, note that the resonant frequency f_R of the sensor is inversely proportional to its net capacitance (C) as $f_R = 1/(2\pi\sqrt{LC})$, where, L is the inductance. We know that $C = \kappa A d^{-1}$, where, A is the area of the probe, d is the distance between the excitation and ground, κ is the dielectric constant of the medium. In presence of defects, κ changes as the medium is altered compared to non-defective scan points. This changes the capacitive loading C of the fringing field which impacts f_R.

The sensor is designed as in (50). It is made up of Roger's 4350 board with dielectric constant of 3.66 and thickness of 1.5 mm. A commercial inductor coil of 100 µH is used to form the LC tank. The experimental scanning setup consists of the following ingredients: (a) LC tank based

capacitive probe with pick up coil (b) Scanning robot arm to move the probe along the sample (c) Power splitter and directional Coupler (d) Data Acquisition System (DAS) and code to produce image (e) Other associated units.

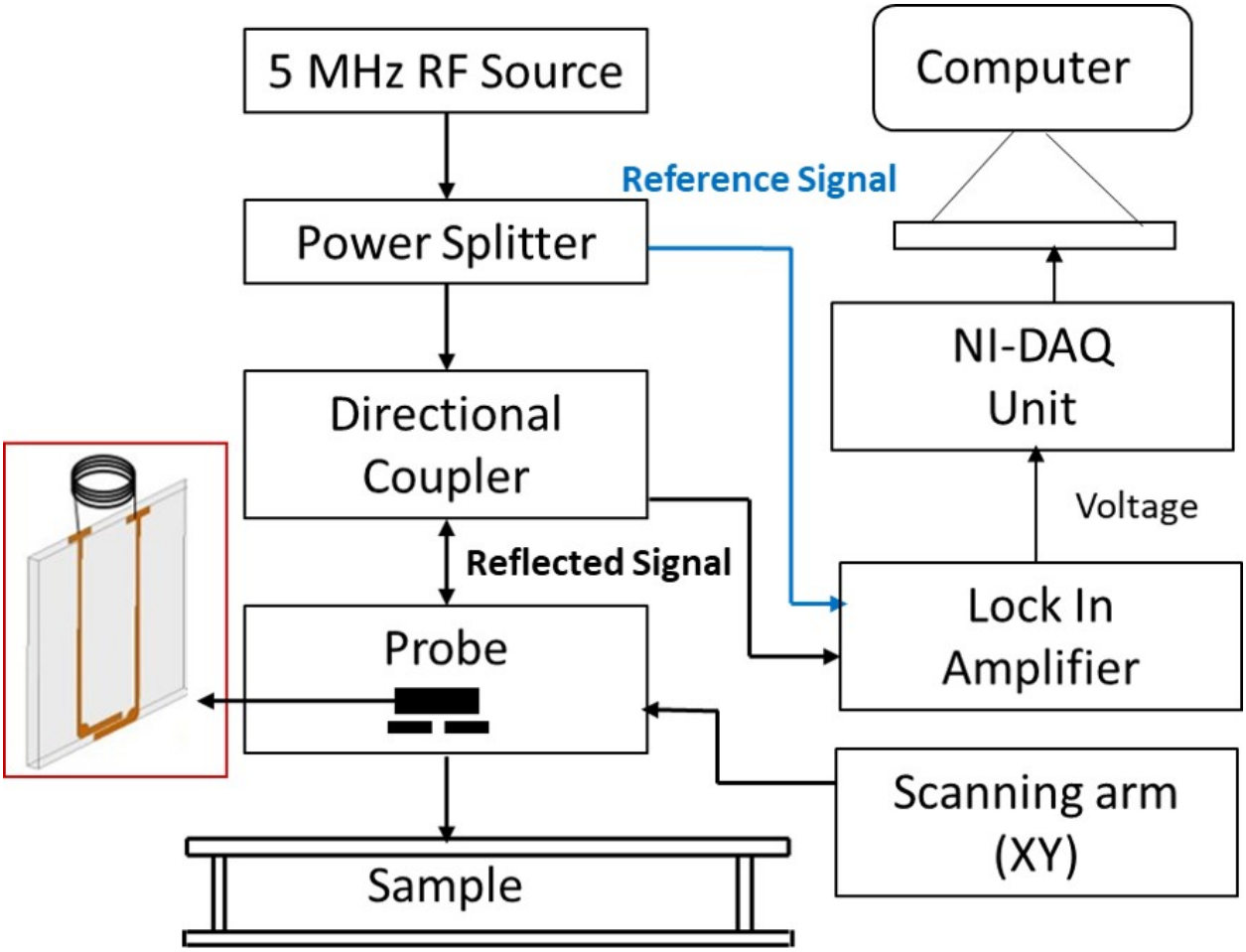

Figure 2.2 Schematic of the entire experimental setup along with the designed capacitive sensor marked in red block.

We used Aerotech AGS1000 programmable XYZ scanner for scanning. The probe is mounted on the scanner. A radio frequency (RF) source is used as input. It is operated at resonant frequency of the probe which was estimated at 5 MHz when the medium was just air with no sample. The probe is read in reflection mode. As shown in Figure 1, it is connected to a power splitter which generates the reference and the measurement signal. The directional coupler is fed with the measurement signal and is connected with the pickup coil of the probe. As in (51; 52), the reflected signal captured by the probe is passed from the directional coupler to the lock-in amplifier which generates voltage signal proportional to the difference in the reference and the measured reflected

signals. This voltage signal thereafter is passed into the DAS consisting of a National Instrument Data Acquisition Card PCIe-6341. The voltage signal is sampled and digitized by a routine in the DAS and the output is recorded. The complete experimental setup used in the laboratory is shown in Figure 2.2.

2.2.2 Experimental design for data generation under uncertainties

To understanding how data collected by capacitive sensors is altered as uncertainties increases in the system, we collect data across different regimes corresponding to varied defect types and capacitive sensing sampling methods. Table 2.1 describes twelve different regimes/experiments. For each experiment, we collect data under three scenarios: (1) capacitive sensing with no lift-off (2) capacitive sensing with moderate lift-off of 3 mm (3) capacitive sensing with increased lift-off of 5 mm. Aside from the lift-off uncertainties, other variations were also incorporated in the capacitive sensing methods. These variations were introduced across the twelve different experiments and they were kept invariant across the three scenarios of the experiments. Table 3.1 shows the details for each experimental set-up. Also, to reflect data collection uncertainties due to non-smoothness of the inspected sample surfaces we used cheap steel samples bought from the shelves of a popular retailer instead of perfectly lab-calibrated expensive samples for collecting our data. Steel samples containing through wall hole (TWH) circular defects of radii 2mm, 4mm and 5mm were used. Figure 2.3 shows one such sample.

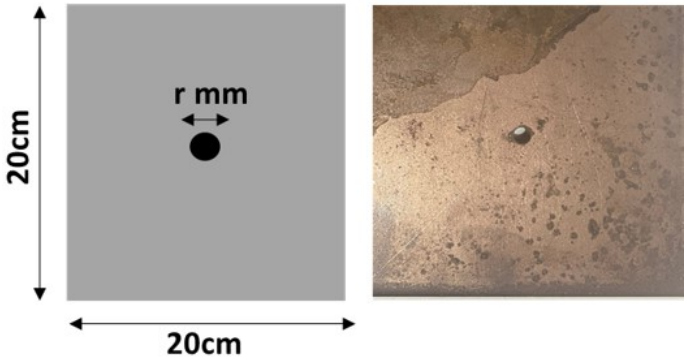

Figure 2.3 Left subplot shows the schematic of steel samples used in the experiments. Samples were of dimension $20cm \times 20cm$ with a circular defect of radius r mm at the center; r equals 2mm for small, 4mm for medium and 5mm for large defects. Right subplot shows one such sample.

Figure 2.4 shows the capacitive sensing images across the three scenarios of experiment 1.

Table 2.1 Description of the twelve different experiments for which voltages based on capacitive sensing was recorded from damaged cheap steel samples under three different kinds of lift-off and the resulting voltage readings were analyzed in details to understand the impact of lift-off.

Experiment No	Step Size (mm)	Sampling rate (s/mm)	Time taken (minutes)	Data Dimension	Defect Size	Remarks
I	1mm	100	5.17	8000*80	Medium	Medium time
II	1mm	500	5.17	40000*80	Medium	Medium time
III	2mm	100	2.5	8000*40	Medium	Minimum time
IV	2mm	500	2.5	40000*40	Medium	Minimum time
V	1mm	100	5.17	8000*80	Small	Medium time
VI	1mm	500	5.17	40000*80	Small	Medium time
VII	0.5mm	500	10	40000*160	Small	Maximum time
VIII	2mm	100	2.5	8000*40	Small	Minimum time
IX	2mm	500	2.5	40000*40	Small	Minimum time
X	1mm	100	5.17	8000*80	Large	Medium time
XI	1mm	500	5.17	40000*80	Large	Medium time
XII	0.5mm	500	10	40000*160	Large	Maximum time

Note that, due to coarseness of the samples in use, voltage readings collected under no lift-off has some fluctuations in the readings at the non-defective scan points. The leftmost plot of Figure 2.4 depicts this variation in the voltage readings as non-defective scan points at the bottom right corner show higher voltage than the non-defective scan points in the top left. The defective points are however clearly detectable under no lift-off with much lower voltage than the voltage at almost all non-defective scan points. With increase in lift-off, the voltage across all the scan points decreases as shown in the legends of middle and right most plots of figure 2.4. While defective scan points still have voltages on the lower end, the difference in voltage between defective and non-defective scan points is much more decreased particularly under increased lift-off. This make classifying defective scan points based on the voltage readings much difficult as lift-off increases.

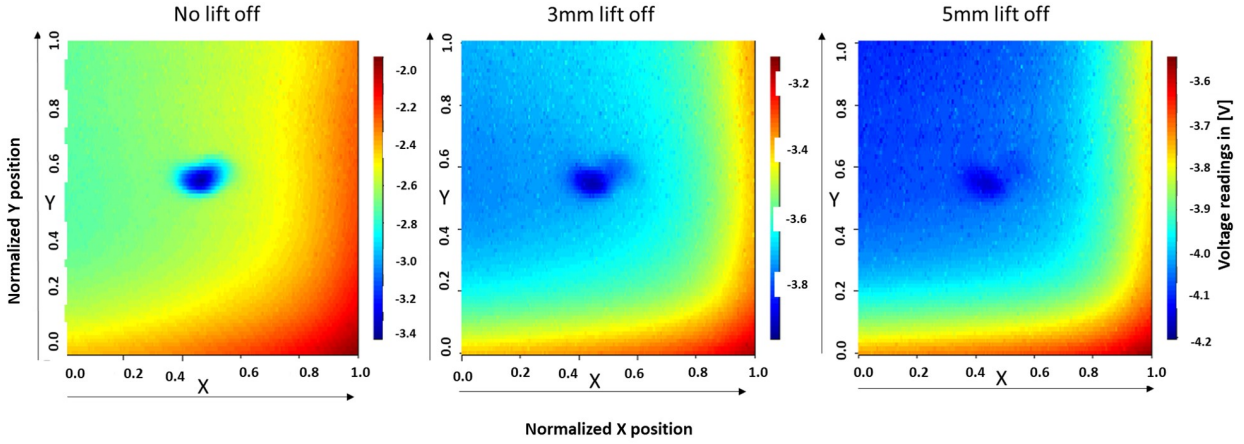

Figure 2.4 Image of voltage readings from the three different scenarios of Experiment 1. The Y and X axes are normalized in $[0, 1]$.

The aforementioned patterns in the three scenarios of Figure 2.4 is broadly seen also across the other experiments. Figures A.1-A.3 in the appendix shows the images of the voltages for all the other cases. Comparing the leftmost (no lift-off scenario) images in figures A.1-A.3, we observe that the voltage magnitudes vary across experiments as uncertainties change across regimes, but the general pattern of diminishing voltages due to increased lift-off was consistent across the experiments. These figures demonstrate that across all the considered regimes which encompass three different defect sizes, as lift off increases we see significant decrease in the gap between the averages of the voltage readings from the defective and the non-defective scan points. The reduction in this gap decreases the signal strength in classifying defective scan points. In the following section, we validate this phenomenon by considering a rigorous statistical framework and quantitatively study the impact of lift off on this gap.

2.3 A framework for studying capacitive sensing signals under uncertainties

We first define a few notations. For the j^{th} scenario of the i^{th} experiment consider observing readings $Y_{ij} = \{y_l(i,j) : l \in \Lambda_i\}$ over grid Λ_i.The grids vary over experiments i when the step sizes are different. We next check if the change in the readings due to lift-offs can be explained by independent corruption in signal intensity across the grid. For this purpose, we consider the additive noise model: $y_l(i,j) = y_l(i,1) + \epsilon_l(i,j)$ for $j = 2, 3$. For any fixed $1 \leq i \leq 12$ and $j = 1, 2$ we test the null hypothesis that $\epsilon_l(i,j)$ are independent across $l \in \Lambda_i$. All the twenty-four p-values are less than 10^{-5} and so, across all the scenarios, the null hypothesis that changes in the signal intensity are independent across locations were conspicuously rejected. This suggests that the corruption in the signals due to uncertainties in the data generation process is not homogeneous across all the scan-points inspected in the sample. We characterize the heterogeneity in the corruptions to the capacitive sensor signals in the following paragraph.

In Figures 2.4 and A.1-A.3, we witness that the lift-off based corruptions in the voltages recorded at the defective scan points is different from the corruptions in the background non-defective scan points. To explain this difference in corruption, we consider a bivariate normal mixture model that uses different voltage distributions for the non-defective and the defective scan

points. Noting that for any experiment i the defective scan points are invariant across scenarios j, let $\Theta_i = \{\theta_l(i) : l \in \Lambda_i\}$ denote the defective scan points in experiment i, i.e., $\theta_l(i) = 1$ if l is a defective scan point and $\theta_l(i) = 0$ otherwise. For experiment i and scenario j, consider the following conditional model on the capacitive sensor readings: for the non-defective background points consider the readings to be independent and identically distributed (i.i.d.) from a normal distribution:

$$[y_l(i,j)|\theta_l(i) = 0] = \boldsymbol{N}(\Delta_0(i,j), \sigma_0^2(i,j)) \tag{2.1}$$

and the readings from the defective scan points also are i.i.d. from a normal distribution with possibly different location and scale than the background:

$$[y_l(i,j)|\theta_l(i) = 1] = \boldsymbol{N}(\Delta_1(i,j), \sigma_1^2(i,j)) \tag{2.2}$$

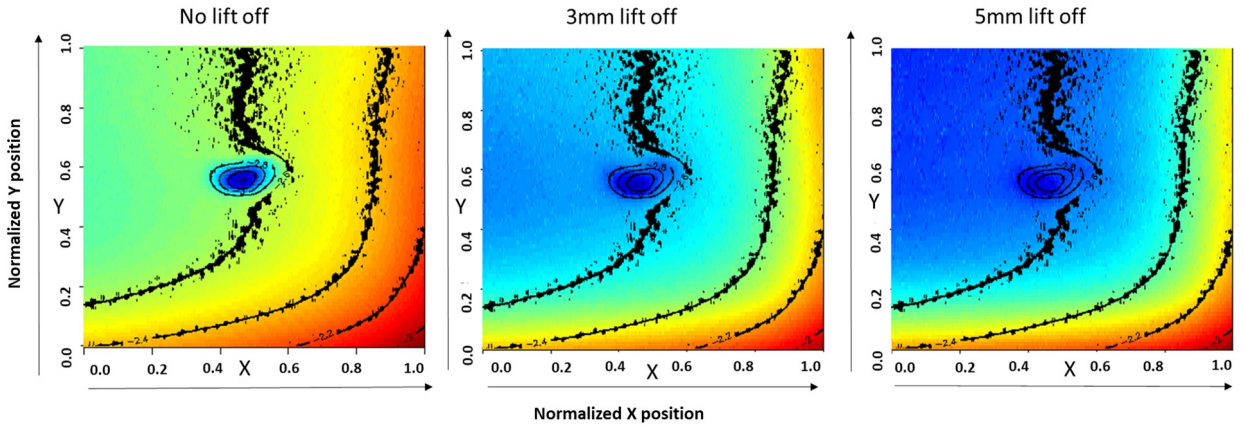

Figure 2.5 Contour plot of left most image is superimposed on each of the images in Figure 1. The threshold of -2.8 is selected based on the contour plot. All scan points in the leftmost image with voltage below -2.8 are donated as defective. The same defective points are marked in the middle and rightmost plot corresponding to 3mm and 5 mm liftoffs based on Experiment 1. The Y and X axes are normalized in $[0, 1]$.

We further assume that conditioned on Θ_i, $y_l(i,j)$ are independent. In practice, we do not observe Θ_i. But, for the experiments in Table 2.1 we know the ground truth regarding the defect size and location. Using the knowledge that defects in all the experiments in Table 2.1 were circular and placed at the center of the grid, we used contour plot based thresholding on the capacitive sensor readings under no lift to estimate Θ_i by $\Theta_i^{\mathtt{or}}$. Note that, we use the suffix `or` in $\Theta_i^{\mathtt{or}}$ to remind that it is not a solely data driven estimator of the support of the defective points but uses oracle information on defects. Figure 3.5 shows how $\Theta_i^{\mathtt{or}}$ is estimated for experiment 1 based on the readings from

case (1). As the readings in case (1) do not have corruptions due to lift-off Θ_i^{or} is well estimated. All the scan points in the grid which has voltage lower than -2.8 were classified as defects and those with voltage higher than -2.8 were classified as non-defects. As we had symmetric regular shaped singular defects in all the experiments in Table 2.1, the simple thresholding scheme used here to calculate Θ_i^{or} was adequate in correctly pin pointing Θ_i. Figures A.4- A.6 in the appendix shows the contour plots for the rest of the experiments.

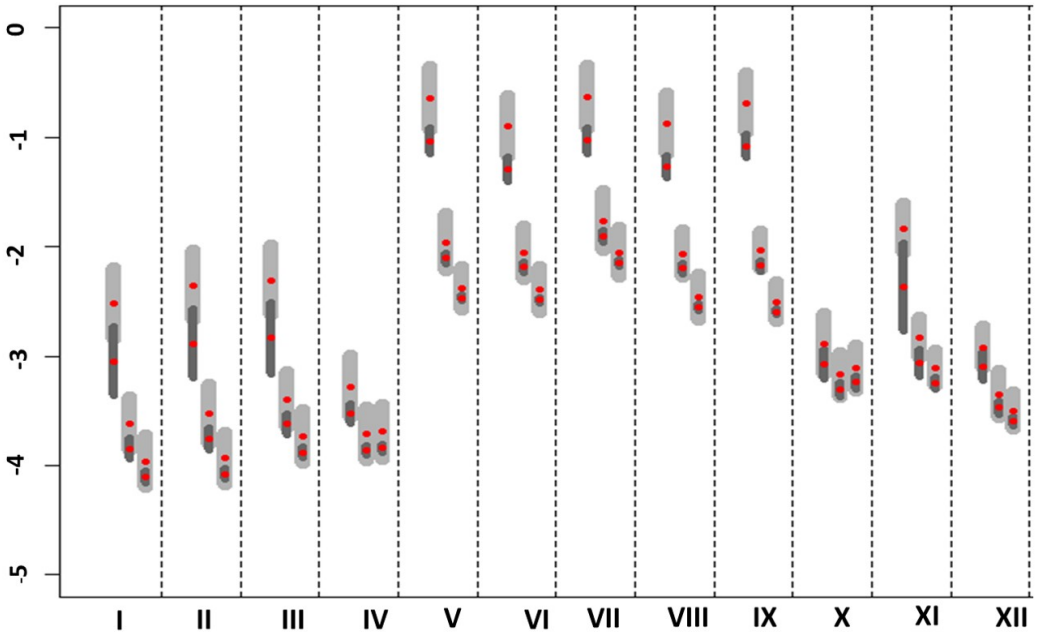

Figure 2.6 95% confidence intervals (CIs) for each of the 3 scenarios (no lift-off, 3mm lift-off and 5mm lift-off respectively) of the 12 experiments documented in table 2.2. The CIs of the non-defective scan-point readings are in light grey whereas the CIs of defective scan-point readings are in dark grey. The means of each distributions are marked in red.

In Table 2.2, based on Θ_i^{or}, we report the means Δ_0, Δ_1 and standard deviations σ_0, σ_1 for the defective and the background readings. Using these means and standard deviations, figure 2.6 shows the 95% confidence interval for the voltage readings at the defective and non-defective scan points. From Table 2.2 and Figure 2.6, we see the following general pattern across the twelve experiments: (a) the difference in means of the voltage between the defective and non-defective area decreases with the increase in uncertainties in the data collection process. The defective scan points always have a lower average voltage; (b) the standard deviation of the signals from the defective areas however decreases with lift-off suggesting that though the mean difference in voltage readings

Table 2.2 The mean and standard deviation of the voltage readings are reported.

Experiment	Defect Proportion	Scenario	Δ_0	σ_0	Δ_1	σ_1
I	1.67%	1	-2.516	0.164	-3.046	0.165
		2	-3.618	0.126	-3.845	0.052
		3	-3.961	0.120	-4.106	0.029
II	1.62%	1	-2.354	0.161	-2.887	0.162
		2	-3.528	0.140	-3.758	0.051
		3	-3.933	0.122	-4.078	0.028
III	1.65%	1	-2.308	0.162	-2.836	0.170
		2	-3.403	0.130	-3.624	0.053
		3	-3.738	0.124	-3.880	0.028
IV	1.60%	1	-3.278	0.142	-3.523	0.047
		2	-3.711	0.124	-3.863	0.024
		3	-3.694	0.125	-3.844	0.023
V	0.51%	1	-0.643	0.150	-1.033	0.060
		2	-1.959	0.133	-2.107	0.028
		3	-2.377	0.099	-2.469	0.016
VI	0.51%	1	-0.899	0.140	-1.291	0.060
		2	-2.051	0.124	-2.188	0.028
		3	-2.393	0.105	-2.489	0.016
VII	0.51%	1	-0.634	0.148	-1.029	0.061
		2	-1.761	0.137	-1.902	0.029
		3	-2.052	0.116	-2.149	0.019
VIII	0.50%	1	-0.869	0.142	-1.268	0.054
		2	-2.061	0.110	-2.199	0.027
		3	-2.462	0.105	-2.551	0.017
IX	0.51%	1	-0.685	0.139	-1.081	0.055
		2	-2.033	0.088	-2.177	0.027
		3	-2.503	0.090	-2.598	0.015
X	5.14%	1	-2.886	0.142	-3.070	0.070
		2	-3.171	0.104	-3.305	0.035
		3	-3.110	0.106	-3.238	0.032
XI	2.50%	1	-1.834	0.116	-2.367	0.207
		2	-2.832	0.095	-3.067	0.065
		3	-3.107	0.084	-3.248	0.032
XII	5.08%	1	-2.919	0.096	-3.095	0.069
		2	-3.347	0.108	-3.473	0.032
		3	-3.498	0.086	-3.596	0.026

between defects and non-defects decreases with lift-off, the readings at the defective point have lower variability and thus are more concentrated with lift-off. Figure 2.7 shows the distributions of the voltage readings at the defective and non-defective scan points for three cases of experiment 1. Figure A.7 in the appendix shows the voltage reading of the other eleven experiments. Figures 2.7 and A.7 also illustrate the above discussed patterns for changes in voltage readings with lift-off.

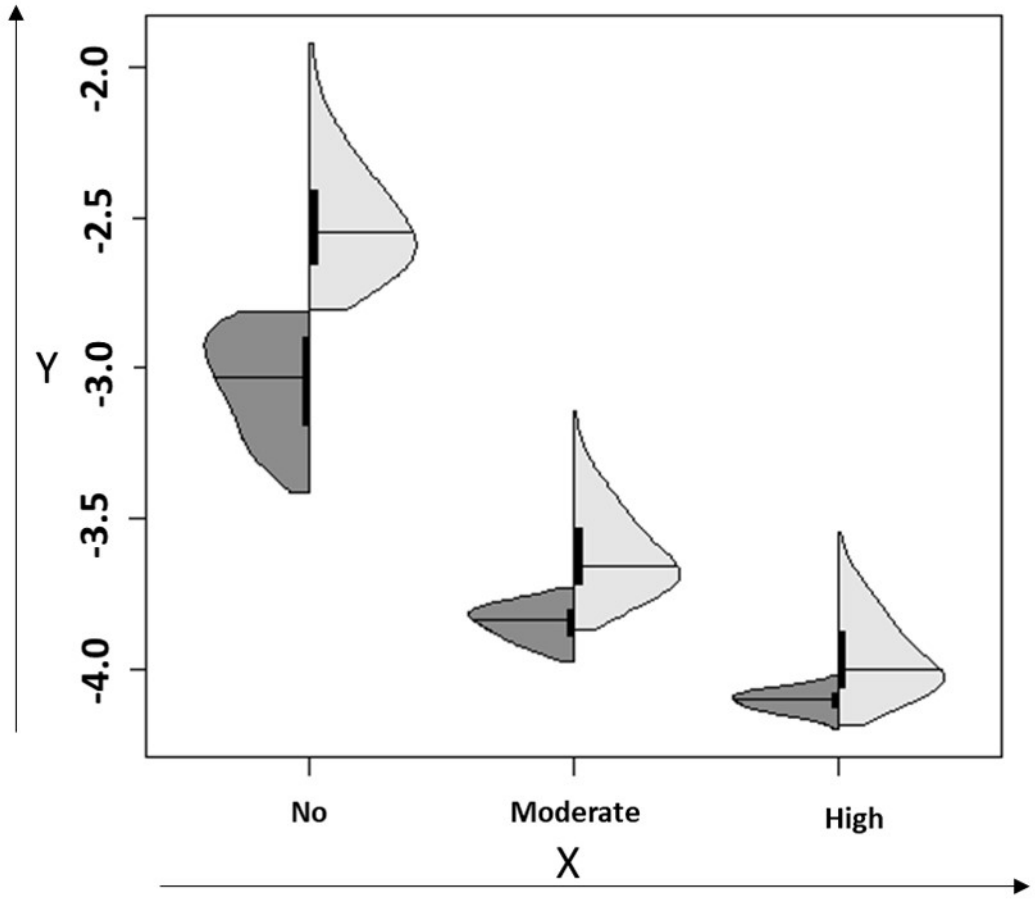

Figure 2.7 Violin plots showing the distributions of voltage readings for the three cases of experiment I (a) no lift-off (b) moderate lift-off of 3mm (c) high lift-off of 5mm. The distributions of the readings from the defective scan points are shown in dark grey and the distributions of the non-defective areas are shown in light grey.

2.4 Bayes errors for defect defection under uncertainties

In practice, we only observe Y_{ij} and not Θ_{ij}. For scenario j of experiment i, the marginal distribution of Y_{ij} is:

$$y_l(i,j) = (1-\pi_i)\, N\big(\Delta_0(i,j),\, \sigma_0^2(i,j)\big) + \pi_i N\big(\Delta_1(i,j),\, \sigma_1^2(i,j)\big) \text{ for } l \in \Lambda_i, \tag{2.3}$$

where, $\pi_i = P(\theta_l(i) = 1)$ is the proportional size of the defect in the grid of inspected points. For any i and j, if we had known $\pi_i, \Delta_0(i,j), \Delta_1(i,j), \sigma_0(i,j), \sigma_1(i,j)$,then based on observing $y_l(i,j)$ we can estimate the support of defective scan points as a point-wise classification problem. In Table 2.3 (last three columns), using the oracle values of the above parameters from Table 2.2 we report the Bayes misclassification error rate (BME) for point-wise classification. We also report the defect misdetection or false negative (FN) rate and the false positive rate (FP) of misclassifying non-defects as defects. BME is the sum of FP and FN.

As the proportions of defective scan points were quite small in all the experiments in Table 2.1, the FN rate (column 8 of Table 2.3) were quite high as compared to FP rate (column 7 of Table 2.3). However, in these NDE applications FN are more harmful than FP and so, we next consider minimizing the FP rate when the FN rate is controlled below 10%.

With the FN rate of defective scan points controlled below 10%, we can detect at least 90% of the defective scan points which would be adequate to identify defect shapes provided the FP rate is also low. In order to minimize FP rate keeping FN rate controlled below 10%, we consider minimizing the weighted classification loss as in the compound decision theory set-up of (53), where, $w_{ij} \geq 0$ is the relative weight of a false positive. For the j^{th} case of the i^{th} experiment, the weighted classification loss for a data driven estimator $\hat{\Theta}(i,j)$ of the support of the defective points Θ_i is:

$$L(\Theta_i, \hat{\Theta}(i,j)) = \frac{1}{|\Lambda_i|}\Big[\sum_i I(\theta_l(i) = 0)\hat{\Theta}(i,j)w_{ij} + I(\theta_l(i) = 1)\{1 - \hat{\Theta}(i,j)\}\Big]. \tag{2.4}$$

When $w_{ij} = 1$ we get back the usual classification loss. The Bayes estimator $\hat{\Theta}_w(i,j)$ that minimizes the weighted classification loss is: for location $l \in \Lambda_i$ set $\hat{\Theta}_w(i,j)[l] = 1$ if

$$\sigma_1^2(y_l - \Delta_0)^2 - \sigma_0^2(y_l - \Delta_1)^2 \geq 2\sigma_0^2\sigma_1^2 \log\left(\frac{w_{ij}\sigma_1(1-\pi)}{\sigma_0\pi}\right) \tag{2.5}$$

and 0 otherwise. Note that for presentational ease, we have dropped the suffix (i,j) from the mean and standard deviation parameters in the above expression. The detailed calculations are provided in the appendix. We calculate the FP and FN rates of this weighted Bayes estimator $\hat{\Theta}_w(i,j))$ as the weights w_{ij} are varied over a wide range of values in $(0,1]$.

Table 2.3 False positive and false negative rate for classifying defective scan points based on the bivariate mixture model with parameters in Table 2.2 is reported across all the experiments. Weighted classification loss reduces the loss of FPs by the reported weights resulting in the FN rates to be lower than the unweighted classification loss. The Bayes misclassification error is reported for both the losses (column 9 for unweighted and column 5 for weighted).

Experiment	Scenario	Weighted Classification Loss				Unweighted Classification Loss		
		FP	FN	BME	Weight	FP	FN	BME
I	1	3.0%	9.0%	3.1%	0.0400	0.2%	36.5%	0.8%
	2	10.0%	9.9%	10.0%	0.0400	0.0%	100.0%	1.7%
	3	15.3%	8.5%	15.2%	0.0333	0.0%	100.0%	1.7%
II	1	2.4%	9.6%	2.5%	0.0500	0.2%	34.4%	0.7%
	2	12.4%	8.7%	12.3%	0.0333	0.0%	100.0%	1.6%
	3	15.1%	8.4%	15.0%	0.0333	0.0%	100.0%	1.6%
III	1	3.0%	9.5%	3.1%	0.0400	0.2%	36.4%	0.8%
	2	12.2%	9.3%	12.2%	0.0333	0.0%	100.0%	1.7%
	3	15.7%	8.7%	15.6%	0.0333	0.0%	100.0%	1.7%
IV	1	10.0%	8.1%	9.9%	0.0400	0.0%	100.0%	1.6%
	2	12.3%	8.3%	12.3%	0.0400	0.0%	100.0%	1.6%
	3	12.3%	8.2%	12.2%	0.0400	0.0%	100.0%	1.6%
V	1	2.1%	8.5%	2.1%	0.0400	0.0%	100.0%	0.5%
	2	15.4%	8.2%	15.4%	0.0100	0.0%	100.0%	0.5%
	3	15.0%	7.6%	15.0%	0.0100	0.0%	100.0%	0.5%
VI	1	1.4%	8.1%	1.5%	0.0500	0.1%	70.4%	0.5%
	2	16.3%	9.0%	16.2%	0.0100	0.0%	100.0%	0.5%
	3	14.6%	7.2%	14.6%	0.0100	0.0%	100.0%	0.5%
VII	1	1.7%	9.3%	1.7%	0.0500	0.0%	100.0%	0.5%
	2	16.5%	9.0%	16.5%	0.0100	0.0%	100.0%	0.5%
	3	15.6%	7.9%	15.6%	0.0100	0.0%	100.0%	0.5%
VIII	1	1.1%	8.1%	1.2%	0.0667	0.1%	68.2%	0.5%
	2	15.2%	8.5%	15.2%	0.0100	0.0%	100.0%	0.5%
	3	15.3%	7.9%	15.3%	0.0100	0.0%	100.0%	0.5%
IX	1	1.1%	8.0%	1.1%	0.0667	0.1%	61.6%	0.5%
	2	10.4%	8.8%	10.4%	0.0133	0.0%	100.0%	0.5%
	3	12.9%	9.8%	12.9%	0.0133	0.0%	100.0%	0.5%
X	1	27.1%	8.4%	26.2%	0.0500	0.0%	100.0%	5.1%
	2	20.6%	6.8%	19.9%	0.0667	0.0%	100.0%	5.1%
	3	20.6%	6.6%	19.9%	0.0667	0.0%	100.0%	5.1%
XI	1	1.4%	9.0%	1.6%	0.0667	0.1%	19.1%	0.6%
	2	6.0%	8.7%	6.1%	0.0500	0.5%	54.6%	1.9%
	3	12.1%	8.6%	12.0%	0.0500	0.0%	100.0%	2.5%
XII	1	20.3%	8.1%	19.7%	0.0400	0.8%	80.0%	4.8%
	2	20.9%	6.7%	20.2%	0.0667	0.0%	100.0%	5.1%
	3	21.8%	7.1%	21.0%	0.0667	0.0%	100.0%	5.1%

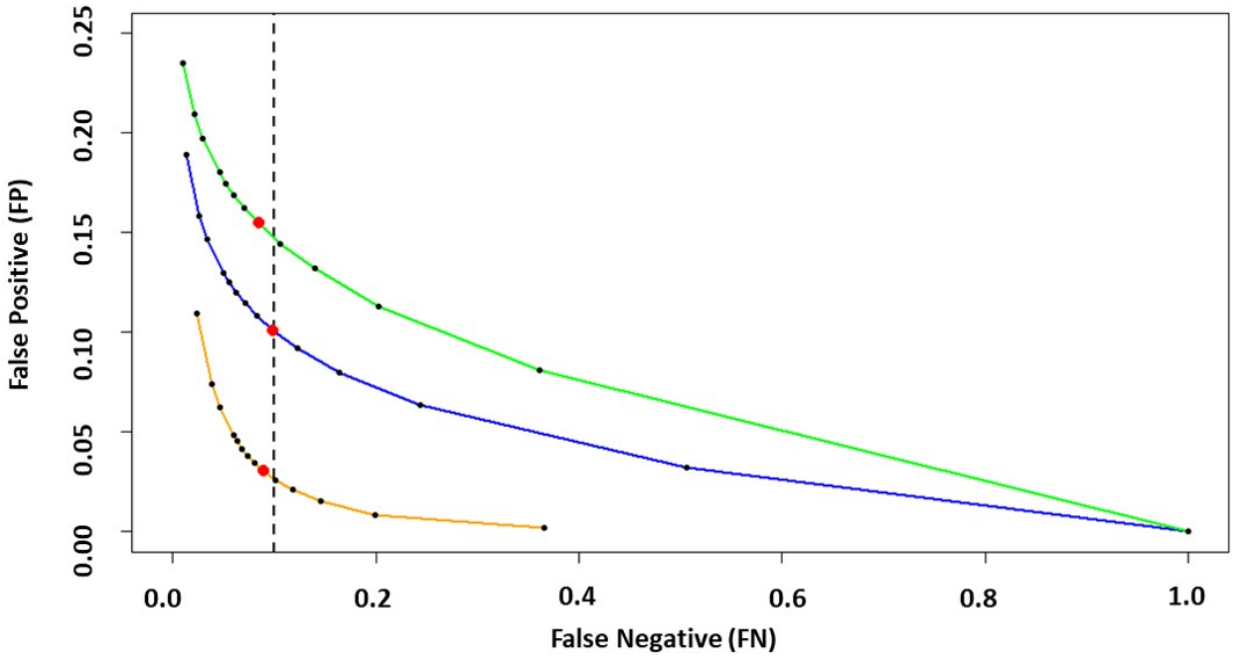

Figure 2.8 The False Positive (FP) vs False Negative (FN) curves for the three scenarios (orange, blue and green for cases I, II and III respectively) of experiment I as the weights are varied in the weighted classification loss. For each curve, the right extreme point denotes the FP and FN proportions for the unweighted loss whereas the left extreme point corresponds to weight $1/200$. The points marked in red are those corresponding to the optimal weights which minimize FP controlling FN below 10%. They are reported in the first three rows of Table 2.3.

We consider weights in $S = \{1, 1/5, 1/10, \ldots, 1/50\} \cup \{1/75, 1/100, 1/200\}$. In Table 2.3, we report the optimal weights w_{ij} in S that minimizes the FP rate keeping the corresponding FN rate below 10%. Note that, in all the cases the sum of the FP and FN rates for the optimal weighted Bayes classifier is more than that of the BME of the unweighted case which minimizes the unweighted misclassification error. However, the optimal weighted Bayes classifier controls the defect misdetection rate below 10% of the defective scan points. From Table III, we witness that FP rates of the optimal weighted Bayes classifier increases with lift-off but are usually well controlled below 15%. In experiments X and XII, the FP rates are a little higher. However, they also have high FP rates under no lift-off which suggest that better FP vs FN trade-offs can be obtained by slightly increasing the FN rates in those two experiments. Note that the classified defective scan points are clustered together at the defect location. If the wrongly classified non-defective scan points are randomly distributed over the grid they can be easily weeded off and reclassified as long as their

proportion is controlled based on local density of defective scans. In Section 6.4, we undertake such an approach that leverages the spatial connectedness of the defective scan points. The low FN and FP error rates for the weighted Bayes estimator in Table 2.3 (columns 3 and 4) shows that capacitive sensing-based voltage readings contains significant information to detect defect under lift-offs considered in cases 2 and 3 of Table 2.1. In Figure 2.8, the False Positive (FP) vs False Negative (FN) curves for the three scenarios of experiment 1 are shown as the weights are varied in the weighted classification loss. For each curve, the right extreme point denotes the FP and FN proportions for the unweighted loss whereas the left extreme point corresponds to weight $1/200$. The points that corresponds to optimal weights that minimize FP controlling FN below 10% are marked in red. Across all defect sizes considered in Table I, we validate that there is a consistent increase in Bayes misclassification error rates with increase of lift-offs. Note that, as the data used in this chapter was solely collected by laboratory experiments, the voltage signals captured by our capacitive sensors had very high signal strength when there were no uncertainties. All our validation results are based on ground truth that is recovered from signals without uncertainties.

2.5 A data driven algorithm for defect detection

The errors reported in Table 2.3 mark the theoretically possible lower limit on the error rates that any decision rule that conduct point-wise classification without borrowing spatial information can produce. To obtain the results in Table 2.3, we have used knowledge of Θ_i^{or} which is not available for real world NDE applications. In this section, we develop a solely data based estimator which does not use any prior knowledge on defect size, shapes or locations. Achieving the error rates in Table 2.3 with solely data driven estimators is not always possible as it would require highly accurate estimates of $\pi_i, \Delta_0(i,j), \Delta_1(i,j), \sigma_0(i,j)$ and $\sigma_1(i,j)$. Estimating $\pi_i, \Delta_1(i,j)$ and $\sigma_0(i,j)$ in presence of noise is extremely difficult and lead to significant estimation error (54; 55; 56; 57). To mitigate these estimation problems, we develop a spatially adaptive procedure to estimate Θ_i.

First a naïve thresholding estimator $\hat{\Theta}_T[l] = 1\{y_l < c_\alpha\}$ is considered, that classifies the l^{th} scan point as defect if its voltage is below the α^{th} quantile of voltage readings over the grid. In Table 2.4, we report the performance of $\hat{\Theta}_T$ for a range of α values. As α increases Coverage rates

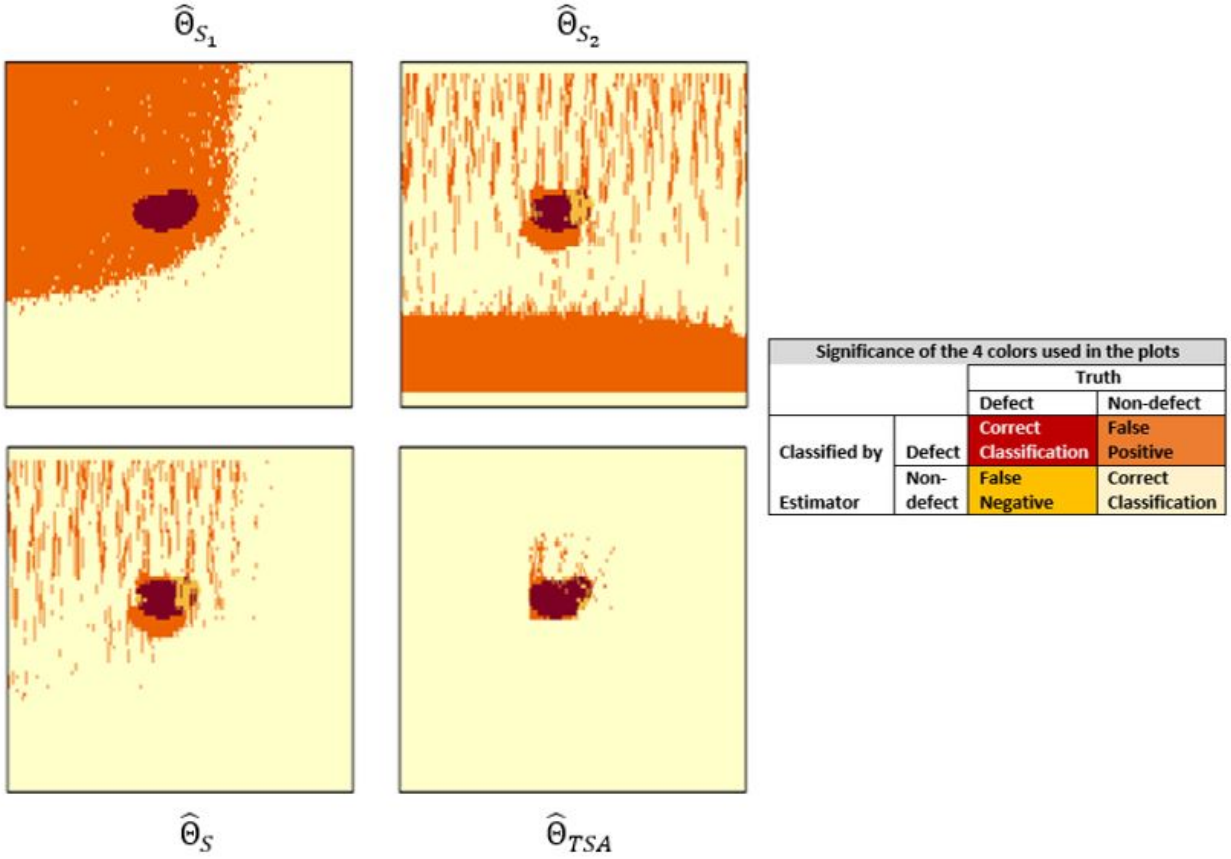

Figure 2.9 For case 3 of experiment I, we present the efficiency of the intermediate estimators that are involved in our proposed spatially adaptive classifier $\hat{\Theta}_{TSA}$ of the defective scan points. Each plot represent the confusion matrix of the corresponding classifier with brown showing the true defect scan points that the classifier had correctly classified, light yellow showing the true non-defective scan points it correctly classified, orange displaying false positives (which are non defective scan points wrongly classified as defects by the classifier) and dark yellow showing false negatives (which are defective scan points wrongly classified as non-defects). The axes in all the plots display scaled normalized locations in $[0, 1]$. The four plots from top left to bottom right, show the gradual filtering process that is involved in the intermediate steps of the proposed $\hat{\Theta}_{TSA}$ estimator. Here, $\beta_1 = \beta_2 = 0.4$ and $\alpha = 0.025$.

of the defect (defined as 1-FN) increase at the cost of increased FP. We observe that while $\hat{\Theta}_T$ works very well for case 1 when there is no lift-off, its performance drastically deteriorates as signals are corrupted due to increase in lift-off uncertainty. We next consider denoising the signals collected under lift-off uncertainty by using the adaptive shrinkage estimator proposed in (58). For denoising we use the denoiseR package of (59) that estimates a low-rank signal assuming Gaussian noise by minimizing the risk estimation criterion in (60; 58). On the denoised data we again tried the elementwise thresholding described above. We observed that under lift-off its performance is not

much better than naïve thresholding of voltage signals (See Table 2.4). Next, we consider a spatial adaptive estimator that is based on local variances in the capacitive imaging data. For voltage readings collected over any grid Λ consider a symmetric neighborhood N_l around the l^{th} scan point in the grid. N_l is a set of grid points that includes the scan point and its neighbors. Consider the variance v_l associated with N_l:

$$v_l = \frac{1}{|N_l| - 1} \sum_{s \in N_l} (y_s - \bar{y}_l)^2, \text{ where, } \bar{y}_l = \frac{1}{N_l} \sum_{s \in N_l} y_s. \tag{2.6}$$

We compute variance over each point in the grid. For N_l we consider a square window with 25 scan points centered around the l^{th} scan point. We use this local variance filter on the voltage readings and estimate the support of defective grid points as:

$$\hat{\Theta}_S[l] = 1\{y_l < c_{\beta_1} \text{and } v_l > c_{\beta_2}\}, \tag{2.7}$$

where, $c_{\beta_1} = \text{quantile}(y_l : l \in \Lambda; \beta_1)$ and $c_{\beta_2} = \text{quantile}(v_l : l \in \Lambda; 1 - \beta_2)$ are the quantiles of the voltage readings and local variances. We universally set the values of β_1 and β_2 at 40%. Note that, as we use large values of the tuning parameters β_1 and β_2 the estimator is not very sensitive to minor changes in the values. Figure 2.9 shows $\hat{\Theta}_S$ for case 3 of experiment I. Consider the decomposition

$$\hat{\Theta}_S = \hat{\Theta}_{S_1} * \hat{\Theta}_{S_2},$$

where, $\hat{\Theta}_{S_1}[l] = 1\{y_l < c_{\beta_1}\}$ and $\hat{\Theta}_{S_2}[l] = 1\{v_l > c_{\beta_2}\}$. Figure 2.9 (top left plot) shows $\hat{\Theta}_{S_1}$ contains almost all defective locations along with a large proportion of non-defective scan points. These non-defective scan points are not uniformly distributed over the grid which would have made our task of filtering them easier but are concentrated in the top left quadrant. This is a result of lift-off uncertainty compounded with probe tilting. Applying the filter $\hat{\Theta}_{S_2}$ based on the local variances (Figure 2.9, top right plot) helps in cleaning these non-defective scan points in $\hat{\Theta}_{S_1}$ keeping most of the defective scan points. The resultant estimator $\hat{\Theta}_S$ (see figure 9, bottom left image) still contains sporadic non-defective scan points. As these non-defective scan points are not concentrated unlike in $\hat{\Theta}_{S_1}$ we can easily weed them off by using a uniform local filter. For this purpose, we consider the averaging filter FA which divides the grid Λ into 100 rectangles and finds the average of signals

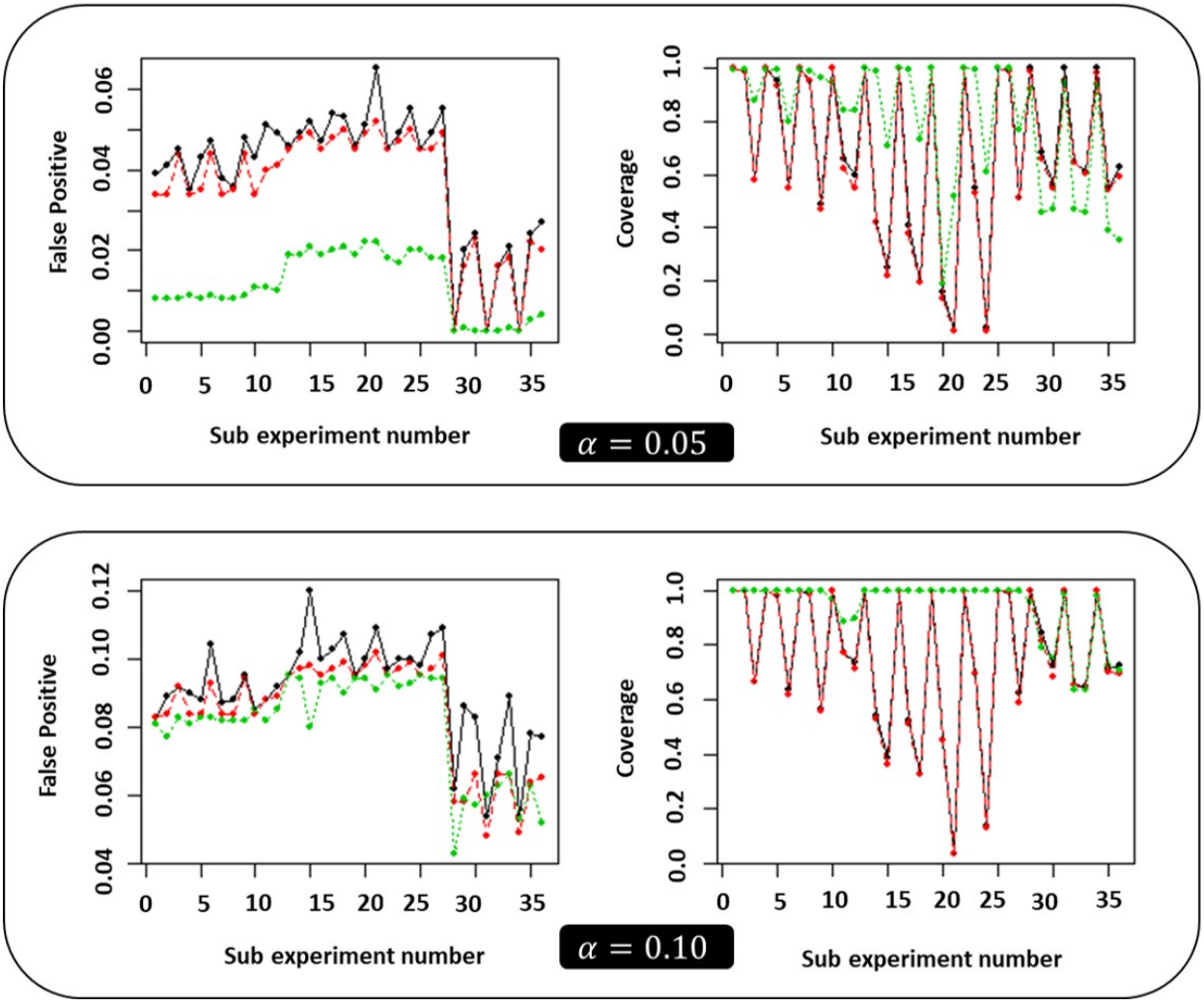

Figure 2.10 We plot the False Positive (FP) proportions and Coverage rates across the 36 sub-experiments (plotted maintaining the order of rows reported in Table 2.2). We present the performance of threshold estimators on raw voltage data (in black), on low-rank model based denoised signals (in red) and by our proposed spatially adaptive estimator (in green) when the threshold parameter was set at 0.05 (top panel) and 0.10 (bottom panel).

over each of those rectangles. We apply FA on $\hat{\Theta}_S$. This smooth (figure 2.9, bottom right plot) out the sporadic non-defective misclassified scan points keeping the concentrated mass of correctly detected scan points intact. Finally, using these as the local weights on the original voltage readings we consider the spatially adaptive estimator: $\hat{\Theta}_{SA} = F_A(\hat{\Theta}_S) * \tilde{Y}$, where, $\tilde{Y} = \max(Y) - Y$.

We conduct a binary classification of grid points as defects by thresholding $\hat{\Theta}_{SA}$ and considering $\hat{\Theta}_{TSA}[l] = 1\{\hat{\Theta}_{SA}[l] > t_{1-\alpha}\}$ where $t_{1-\alpha} = \text{quantile}(\hat{\Theta}_{SA}[l] : l \in \Lambda; \alpha)$. In Table 2.4, we report the FP and coverage rates of the proposed estimator. As shown in recent works (57) , using spatially adaptive local weights can tremendously help in recognizing underlying patterns from very noisy imaging data. Here, thresholding the spatially adaptive estimator $F_A(\hat{\Theta}_{S1} * \hat{\Theta}_{S2})$ instead of the original signal or low rank model based denoised signal, we observe that we can reasonably filter out the effects of probe tilt and lift-off based uncertainties.

Table 2.4 and figure 2.10 compares the efficacy of the proposed method with naïve thresholding and low rank signal denoising. They show that using the aforementioned spatially adaptive estimator based on local features of the imaging data, it is possible to attain appreciable coverage and low false positive rate in detecting defective scan points under the varied uncertainties considered here. From Table 2.4 and Figure 2.10, we observe that thresholding based on raw voltage signals is effective under no lift-off but provides very low coverage as lift-off increases. Low rank model based denoising also does not help. It reduces false positive rate a bit but suffers from insufficient coverage. The proposed spatial adaptive estimator however can provide considerably high coverage with very tight control on the false positive rates. As such with false positive rates controlled at only 2%, the spatially adaptive estimator provides sufficient coverage under lift-off in all experiments barring VII, X, XI and XII. When the limit on the false positive proportion is raised to 10%, the spatially adaptive estimator provides sufficient coverage across all the 36 regimes considered. It is to be noted that there is discernable difference in the performance of the estimator (as shown in Figure 2.10) between regimes 1 to 30 and 31 to 36. This is due to the fact that regimes 31 to 36 contain defects of considerably different characteristics than those in the previous regimes. Overall, the results in Table 2.4 and Figure 2.10 show that the spatially adaptive estimator based on local

features of the imaging data can attain appreciable coverage and low false positive rate in detecting defective scan points on the grid under the different lift-off uncertainties considered in this chapter.

In equation 2.6, we considered fixed sized rectangular neighborhoods. They can be easily replaced by Gaussian kernel filters that weighs each scan points in the lattice inversely proportional to the exponent of its distance from the scan point under study. In this context, the proposed filters are similar to signal processing techniques that uses Gaussian filters. However, we found that the local variances in these data sets contain important information and so, we filtered the variances. In this attribute, we differ from traditional filtering where filtering is done based on signal intensity or amplitude and not variances.

2.5.1 Multiple coexisting defects

All the experiments considered in Table 2.4 contained a single defect. Next, we consider two experiments (MI and MII) where we had multiple coexisting defects. Each experiment had three scenarios as before pertaining to no, 3mm and 5mm lift off along with other embedded uncertainties. Figure 2.11 shows the structure of the defects. There were four and twelve defects of varying shapes and sizes in experiments MI and MII respectively.

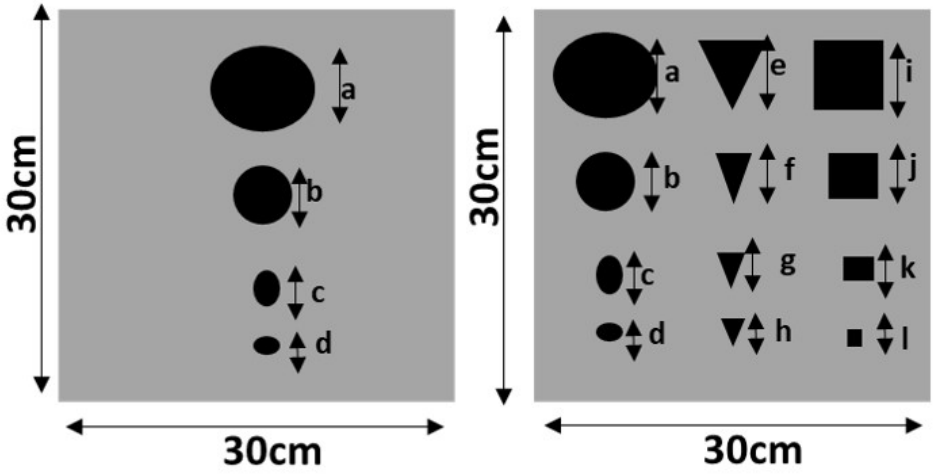

Figure 2.11 Schematic shows the defects used in experiments MI (left) and MII (right). MI had four circular defects a, b, c and d of diameters 48 mm, 20 mm, 8 mm and 1mm respectively. In MII along with these four defects, there are four equilateral triangle defects (marked as e, f, g , h) with sides 42, 28, 10 and 2 mm respectively and square defects (marked as i, j, k, l) whose length of sides are 30, 20, 6 and 3 mm respectively.

Table 2.4 False Positive proportion and Coverage rate for threshold estimators on raw voltage data, on low-rank model based denoised signals and by our proposed spatially adaptive estimator when the threshold parameter α was set at 0.05 and 0.10 respectively are reported. Results for the 36 sub-experiments of Table 2.2 are presented in order across rows.

Threshold at 0.05th quantile						Threshold at 0.10th quantile					
Raw Voltage		Low Rank Denoised		Spatial Denoising		Raw Voltage		Low Rank Denoised		Spatial Denoising	
FP	Coverage	FP	Coverage	FP	Coverage	FP	Coverage	FP	Coverage	FP	Coverage
0.039	1.000	0.034	1.000	0.008	0.996	0.083	1.000	0.083	1.000	0.081	1.000
0.041	0.990	0.034	0.987	0.008	0.993	0.089	0.997	0.084	0.996	0.077	1.000
0.045	0.582	0.044	0.582	0.008	0.879	0.092	0.663	0.092	0.663	0.083	0.996
0.035	1.000	0.034	1.000	0.009	0.998	0.090	1.000	0.084	1.000	0.081	1.000
0.043	0.952	0.035	0.936	0.008	0.994	0.088	0.983	0.084	0.979	0.083	1.000
0.047	0.552	0.044	0.549	0.009	0.799	0.104	0.639	0.093	0.616	0.083	1.000
0.038	1.000	0.034	1.000	0.008	0.995	0.087	1.000	0.084	1.000	0.082	1.000
0.036	0.956	0.035	0.956	0.008	0.992	0.088	0.986	0.084	0.986	0.082	1.000
0.048	0.485	0.044	0.472	0.009	0.968	0.095	0.562	0.094	0.556	0.082	1.000
0.043	1.000	0.034	1.000	0.011	0.949	0.085	1.000	0.084	1.000	0.085	0.972
0.051	0.662	0.040	0.624	0.011	0.842	0.088	0.772	0.088	0.772	0.082	0.884
0.049	0.601	0.041	0.548	0.010	0.843	0.092	0.740	0.089	0.712	0.085	0.898
0.046	1.000	0.045	1.000	0.019	0.999	0.095	1.000	0.095	1.000	0.095	0.999
0.049	0.423	0.048	0.418	0.019	0.989	0.102	0.540	0.097	0.529	0.094	0.999
0.052	0.250	0.049	0.221	0.021	0.709	0.120	0.386	0.098	0.362	0.080	0.999
0.047	1.000	0.045	1.000	0.019	1.000	0.100	1.000	0.095	1.000	0.093	1.000
0.054	0.406	0.048	0.380	0.020	0.997	0.103	0.523	0.097	0.509	0.094	1.000
0.053	0.196	0.050	0.196	0.021	0.732	0.107	0.329	0.099	0.324	0.090	1.000
0.046	1.000	0.045	1.000	0.019	1.000	0.095	1.000	0.095	1.000	0.094	1.000
0.051	0.155	0.049	0.135	0.022	0.187	0.100	0.449	0.098	0.449	0.094	0.999
0.065	0.011	0.052	0.009	0.022	0.517	0.109	0.038	0.102	0.037	0.091	1.000
0.045	1.000	0.045	1.000	0.018	1.000	0.097	1.000	0.095	1.000	0.095	1.000
0.049	0.547	0.047	0.530	0.017	0.993	0.100	0.694	0.097	0.693	0.092	1.000
0.055	0.022	0.050	0.013	0.020	0.609	0.100	0.139	0.099	0.132	0.093	1.000
0.045	1.000	0.045	1.000	0.020	1.000	0.098	1.000	0.095	1.000	0.095	1.000
0.049	0.989	0.045	0.989	0.018	0.999	0.107	0.995	0.097	0.995	0.094	1.000
0.055	0.515	0.049	0.515	0.018	0.771	0.109	0.627	0.101	0.591	0.094	1.000
0.000	1.000	0.000	0.987	0.000	0.925	0.062	1.000	0.058	1.000	0.043	0.956
0.020	0.682	0.016	0.660	0.001	0.458	0.086	0.842	0.058	0.817	0.059	0.792
0.024	0.560	0.023	0.549	0.000	0.470	0.083	0.726	0.066	0.685	0.057	0.752
0.000	1.000	0.000	0.956	0.000	0.948	0.054	1.000	0.048	1.000	0.060	0.987
0.016	0.650	0.016	0.649	0.000	0.472	0.071	0.655	0.066	0.655	0.063	0.639
0.021	0.611	0.018	0.606	0.001	0.460	0.089	0.648	0.066	0.645	0.066	0.639
0.000	1.000	0.000	0.985	0.000	0.942	0.054	1.000	0.049	1.000	0.053	0.984
0.024	0.548	0.022	0.546	0.003	0.390	0.078	0.711	0.064	0.700	0.063	0.723
0.027	0.629	0.020	0.592	0.004	0.356	0.077	0.726	0.065	0.695	0.052	0.710

Figure 2.12 shows the image plots of the voltage readings collected by our capacitive sensing probe.

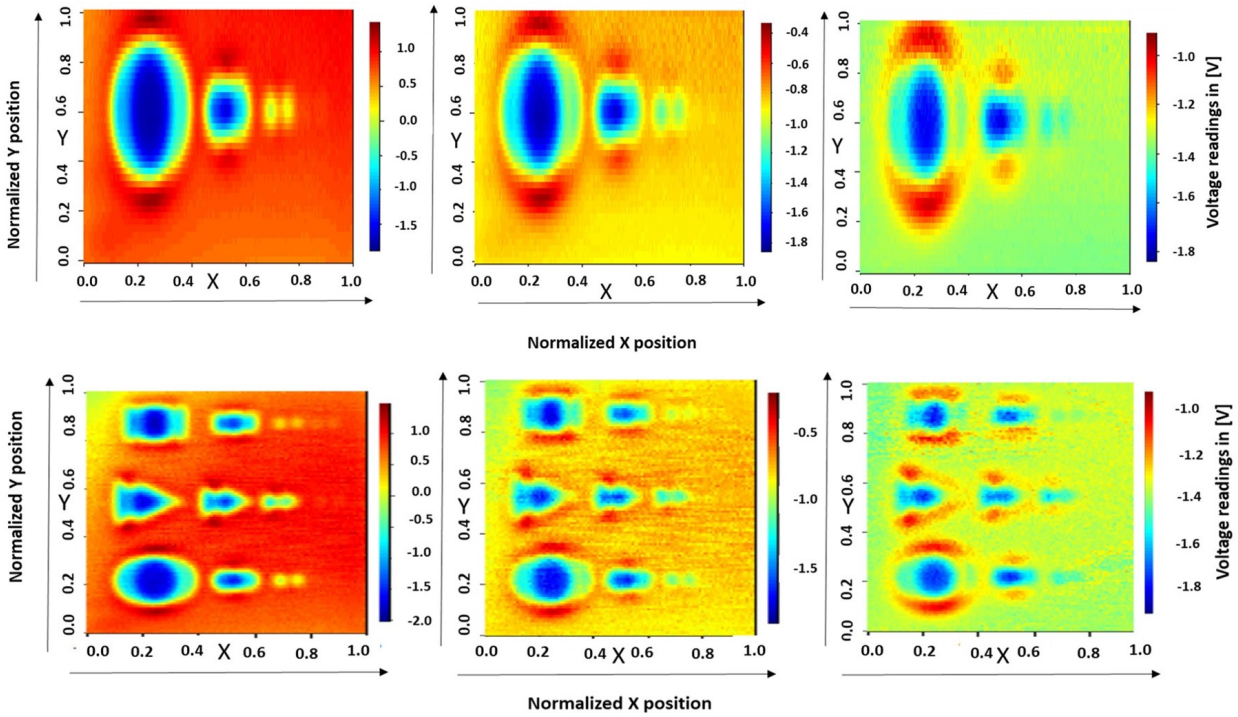

Figure 2.12 Plot of voltage readings from the three different scenarios of Experiment MI (top) and MII (bottom) respectively. The Y and X axes are normalized in [0,1].

In Table 2.5 we report the false positive proportion and defect detection coverage rate that was achieved by applying our proposed spatial denoising algorithm on the capacitive sensing data. We observe that our proposed method not only outperforms the aforementioned competing algorithms but also obtains high coverage and low false positive rate in detecting the multiple, disconnected defect clusters.

Table 2.5 False positive and coverage rates for coexisting defects. α was set at 0.10.

Experiment	Scenario	Raw Voltage		Low Rank Denoised		Spatial Denoising	
		FP	Coverage	FP	Coverage	FP	Coverage
MI	1	0.034	1.000	0.033	1.000	0.036	1.000
	2	0.034	0.999	0.034	0.999	0.052	0.979
	3	0.055	0.822	0.054	0.819	0.076	0.887
MII	1	0.057	1.000	0.056	1.000	0.055	1.000
	2	0.057	0.999	0.056	0.999	0.055	1.000
	3	0.073	0.870	0.069	0.866	0.080	0.930

2.6 Conclusion

In this chapter a robust NDE based defect detection methodology is developed using short waves based capacitive sensing. This method is cheap as it doesnot need complex expensive sensors. It does not need any expert supervision and can produce instantaneous defect identification under highly robust conditions by using spatially adaptive denoising. Moreover, the apparatus involved for data collection is light and flexible and can be implemented across a wide range of NDE applications.

We demonstrate high efficacy of the proposed NDE methodology in non-smooth samples under varied lift-off uncertainties. We provide a Gaussian mixture model based disciplined analysis to check when defective scan points in the sample can be detected with high accuracy and with low false positive misclassification rates for non-defective scan points. This would lead to identification of defect characteristics such as location, size and shape with desired accuracy. The results are validated on experimental data on samples containing single defect of three different sizes as well as on samples containing multiple coexisting defects. We show that the proposed spatially adaptive denoising algorithm leverages the spatial contiguity of defective scan points and can well identify defects based on capacitive sensing data that is convoluted with the uncertainties due to lift-off, probe tilt and low sampling rates.

Portions of this chapter appeared in the publication (12).

CHAPTER 3

INLINE PIPELINE INSPECTION USING HYBRID DEEP LEARNING AIDED ENDOSCOPIC LASER PROFILING

3.1 Introduction

Structural integrity of pipelines decreases with usage and thus regular monitoring is needed for safe operations. Leaks and breakages in pipelines left undiagnosed can cause significant loss of property and lives (61; 62). It is very important to detect defects in a timely manner. In this chapter we have conducted NDE inspection on non-metallic pipelines which constitute a significant percentage of our current natural gas distribution and domestic pipeline systems. These pipelines are mostly made of plastic materials such as high density polyethylene (HDPE), medium density polyethylene (MDPE) and polyvinyl chloride (PVC). They are extensively used for their high strength to weight ratios and low manufacturing and installation costs (63). Aging plays a major role in degradation of these pipelines and can lead to catastrophic failure without proper monitoring. Existing non-destructive evaluation (NDE) methods such as magnetic flux leakage (MFL), Eddy current, ultrasonic testing and electro-magnetic acoustic transducer (EMAT) (28; 29; 27; 30; 64) which are very popular for inspecting metallic pipelines can not be used in these polymer-based pipelines due to the dielectric nature of their wall materials. While MFL, Eddy current and EMAT need highly conductive or ferro-magnetic pipe walls, ultrasound techniques need coupling medium which is not achievable in practice. A novel NDE method using optical imaging is deveoped here that does not need any coupling medium or conductive pipe walls. This method uses laser ring scanning and profiling to collect inline pipeline inspection data which is then subsequently processed by pre-trained deep learning networks to produce accurate defect identification.

A laser ring scanning the inner pipe wall can provide a clear and accurate profile of pipeline using triangulation methodology (65; 66). The proposed non-destructive inspection apparatus uses a camera to capture the high-quality reflected images from the inner pipe wall. Existing NDE works using such endoscopic profiles used manual interventions to introspect the recorded images. Such offline mode of assessment is highly labor intensive and time consuming. Modern NDE

methods need to routinely inspect massive pipeline systems and provide real time feedback and so, need to incorporate automated, instantaneous defect identification routines (67). Recently, machine learning algorithms have been successfully applied for instantaneous defect detection in pipelines using different NDE based imaging methods (68; 69). However, these algorithms are highly sensitive to the data collection procedures and cannot be robustly appended for post-processing in other related NDE techniques. Also, while locating defects is the important first step, contemporary NDE systems need to differentiate between the different defects and provide detailed diagnosis, thereby accurately recognizing the obtrusive defects from the benign types (70; 71; 72). Classifying defect types with high accuracy based on laser images is a significantly more challenging problem than merely detecting aberrations in the pipe inner walls. Here, an integrated framework for probing and analysis that uses deep learning classifiers to provide accurate diagnosis of different defect types based on laser ring images collected from endoscopic scanning is developed.

Developing the software for defect classification in this framework is challenging as the differences in the laser-scan images across different defect types is very subtle and localized. The difficulty is also compounded for the following two reasons: (a) images from endoscopic scanning conducted at different speeds can have varying characteristics (b) there is imbalance among different defect types; for example, defects resulting from material losses are more frequent than dents and holes in the pipeline. A robust identification system is developed that can learn from experimental data and can generalize well across diverse test sets. For that purpose, I have analyzed and compared the predictive accuracy for a host of state-of-the-art machine learning algorithms. An extensive bag of methods which along with several deep learning architecture also included logistic regression (LR), linear discriminant analysis (LDA), non-parametric nearest neighbor classifier (KNN), support vector machines (SVMs) with different kernel choices, random forest (RF) and classification and regression trees (CART) are considered here. By analyzing such a wide array of methods, we demonstrate the predictive accuracy gains that complex learning methods can produce over simpler methods. In particular, by calibrating network complexity in deep learning set-ups based on out-of-sample error rates, an efficient deep learning architecture using crossed

hierarchical bilinear pooling (HBP) is developed that yield 98.33% accuracy rates. The optimally calibrated HBP method was arrived by gradually increasing modeling flexibility from the simplest 2D convoluted neural network (CNN) to transfer learning aided CNN (CNN-TL) to hybrid network comprising of CNN and long short-term memory (CNN-LSTM) recurrent neural networks. In laser-ring scanned images, deformities due to defects lead to local perturbations in the circular rings produced by the otherwise non-defective inner walls of cylindrical pipes. HBP exploits inter layer feature interactions to identify the local features embedded within the images (73) and can detect different types of edge deformities in laser-ring scans with high precisions.

The key contributions of this chapter are as follows:

1. An integrated framework has been developed for automatic online inspection for defect detection and defect type identification in the inner walls of plastic pipes using laser profilometry. Distortions caused by defects in the laser rings are collected by the camera in the endoscopic scanning and is recorded in the video format as the camera traverse through the pipe. These videos are fragmented into images. These images are fed to a pre-trained HBP deep learning network to detect different defect types along the pipe walls.
2. For pipelines in usage, light intensity and scanner speeds in endoscopic methods can greatly vary across different scans in the same pipeline (65). To reflect such heterogeneous characteristics in real inspection data, we vary the scan speeds in our experimental set-ups. Our proposed HBP based NDE method can produce highly accurate diagnosis in such datasets with varying profilometrics.
3. Fine grained visualization is achieved by extensively tuning a hybrid cross layer HBP that uses Resnet50 (74) as the baseline model. In the process, we document predictive performance by a gamut of machine learning classifiers clearly demonstrating the analytical perspective that can lead to significant gains. In particular, by carefully curating a series of frames along the laser-ring's circular path via circular hough transformation (CHT) (66) we develop a hybrid neural network based on CNN-LSTM that take into account both the spatial and temporal features present within the image. We show that such CNN-LSTM outperforms basic CNNs

but is dominated by HBP which is specifically geared for fine grained visualization and edge detection by accounting for local variations.

The rest of the chapter is organized as follows. In section 3.2, we describe the hardware developed for laser ring imaging and the data collected by using it in laboratory experiments. In section 3.3, we describe the different deep learning architectures including the HBP which we prescribe for laser-scan image data analysis in our integrated NDE framework for online inspection of plastic pipes. In section 3.4, we provide the misclassification errors from the optimally tuned deep learning models along with those from conventional machine learning models. We not only exhibit superior performance of the HBP based laser profilometry method but also demonstrate its very low out-of-sample error rates in identifying defect types.

3.2 Endoscopic Laser Profiling: Hardware and Data Generation

Laser profilometry forms an effective tool for NDE inspection of plastic pipelines. The hardware consists of a laser projector which act as the source and projects defined patterns of light inside the pipe. The light reflected on the pipe walls is captured by the camera sensor. CMOS (Complementary Metal Oxide Semi-conductor) camera is the most popular method for visualization in laser profilometry due to its cheap cost as compared to CCTV (closed circuit television). A laser diode (75) is used as the light source and the reflected light forms a circular ring which is read by the camera. The precision of the projector is important as in presence of defects there will be distortions in the laser ring.

Our laser ring profiling methodology is based on the method of triangulation (76; 77) and the parallax between the camera and illumination. A signal proportional to the object position is generated by the reflected laser light. In our apparatus, we place the illumination source (laser) and the detector (CMOS) on the same optical axis as shown in the schematics in Figure 3.1. This results in constant illumination over the inner circumference of the pipe and thus all the points along the circumference are measured with same intensity. Figure 3.2(a) shows the entire experimental setup which comprises of a laser sensor, CMOS camera for data acquisition, motor to move the scanner inside the pipes, pipes of varying diameters and materials used for scanning. Figure 3.2(b) shows

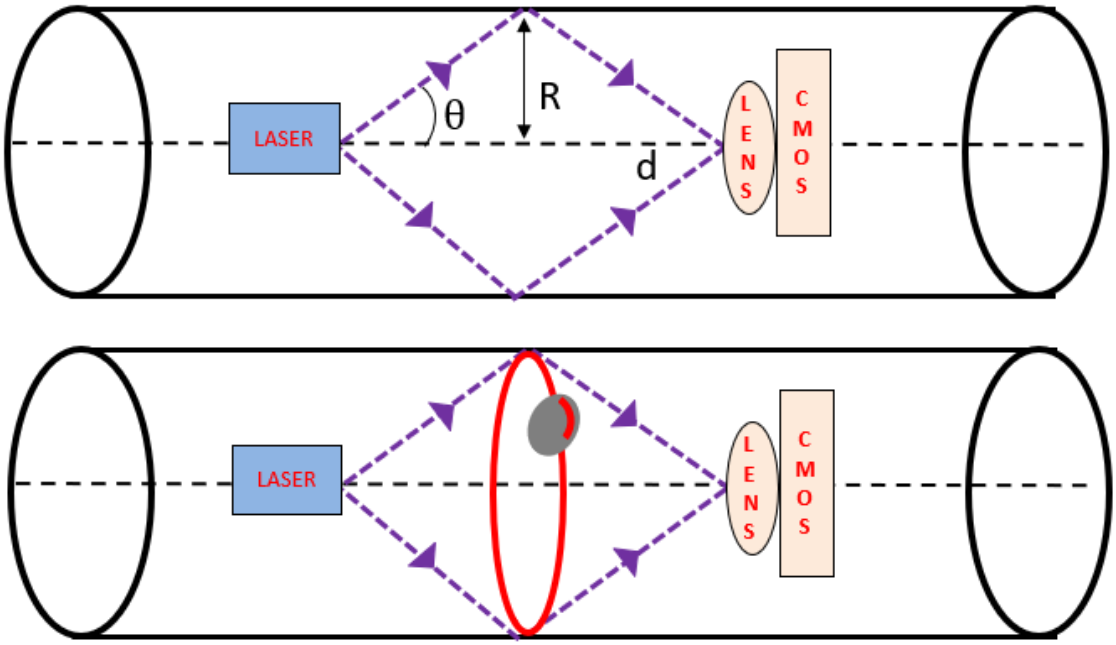

Figure 3.1 (a) In top schematic illustrating the hardware used in laser-profiling showing laser projector (source) and CMOS camera (detector) stationed along the same optical axis. R is the inner diameter of the pipe wall, d is the distance between source and detector and theta is the angle of projection from the source (b) bottom schematic shows deformation in the red laser ring in presence of defect in gray.

Figure 3.2 (a) Entire experimental setup (on left) (b) portion of the scanner with the camera and laser projector facing each other (right image).

the scanner arm with the laser source and the camera sensor facing each other.

The developed sensor enjoys the desirable property that it can be easily adapted to work across

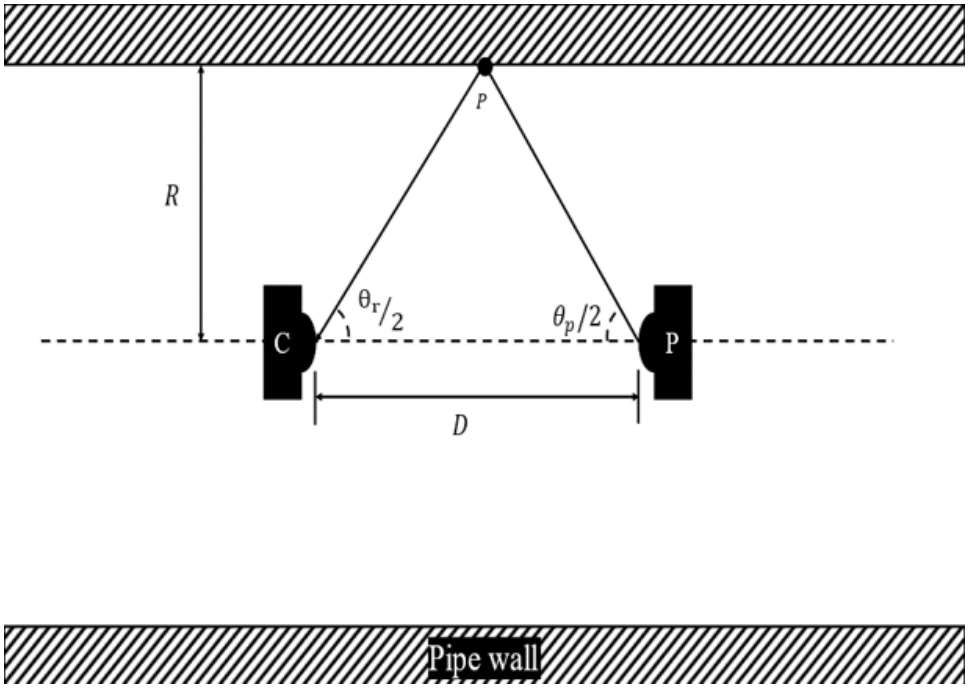

Figure 3.3 Schematic showing laser projector and camera setup.

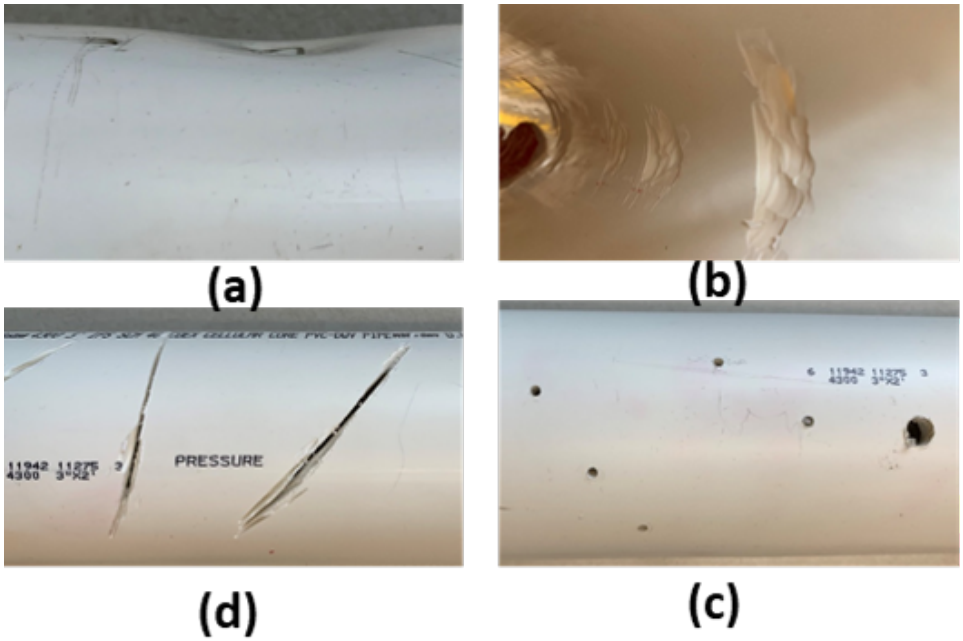

Figure 3.4 Different defects: (a) dent,(b) material loss, (c) holes, (d) slits are presented.

pipes of highly varying diameters. The adaptation would just involve increasing the length of the metal rod that connects the camera and projector holders with increase in the diameter of the pipe. If the change in diameter is not large the adjustment is even simpler and can be done by merely

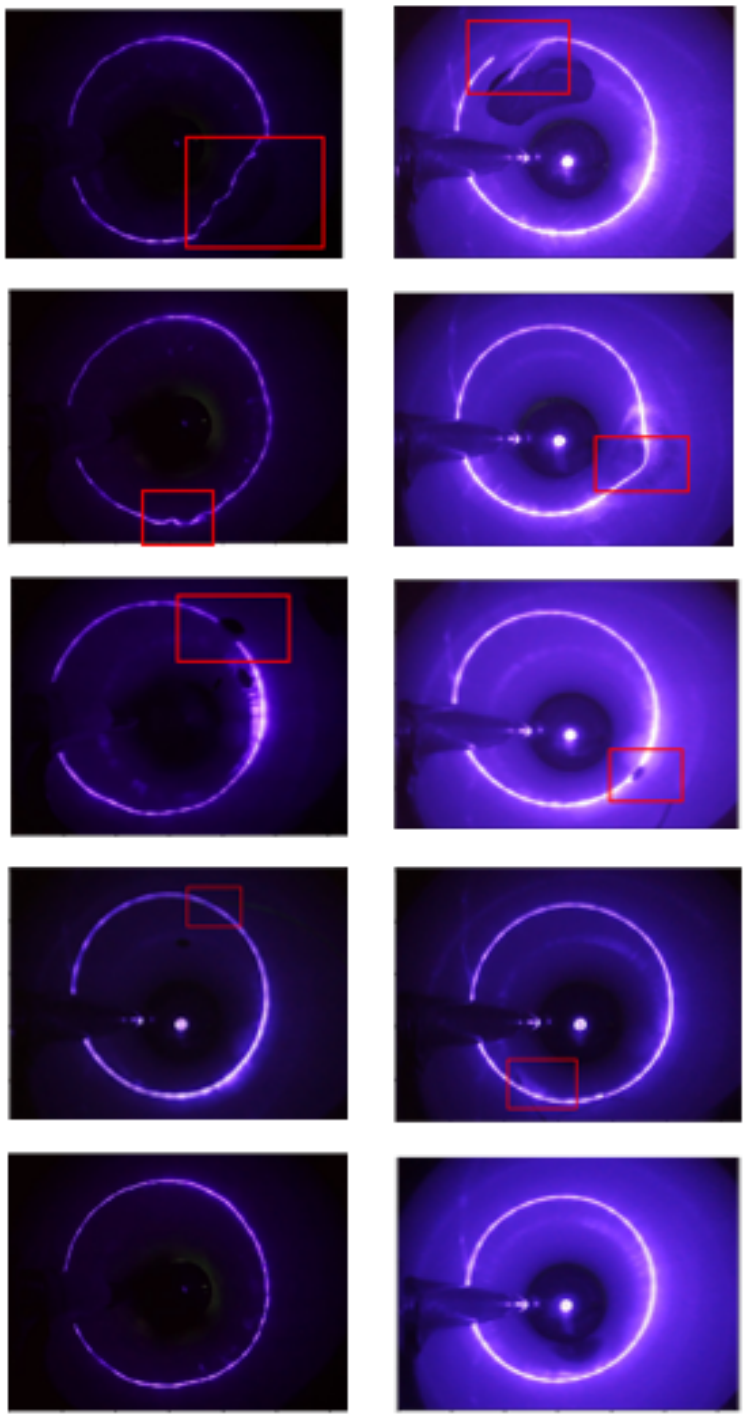

Figure 3.5 Distortion of the laser ring due to various defects. From top to bottom, we respectively have laser-ring scans as recorded by the CMOS in pipe walls with dent, material loss, hole, slit and no deformities. For display only, the defective area is manually gated by a red box. The laser intensity used while collecting on the left images were much lower than those on right.

adjusting the position of the nuts on the metal rod. To understand the reason behind the adjustment, consider that the projection angle of the projector used in our sensor is θ_p and its camera's field of

view angle is θ_c. Also, assume that the camera and projector are separated by a distance D and the sensor was placed along the axis of the cylindrical pipe pointing along the z-axis. Let R be the radius of the pipe. For a point P on the pipe surface, we have $\tan\tilde{\theta}_p = R/P_z$, where, $\tilde{\theta}_p = \theta_p/2$ is half projection angle of the laser projector, and P_z represent the Z-coordinates of the point. The angle θ_r of the received camera ray satisfies: $\tan\theta_r = R/(D - P_z)$. This implies,

$$\theta_r = \arctan\left(\frac{R}{D - R\cot(\theta_p/2)}\right).$$

To guarantee that the projected ring is within the camera field of view as shown in Figure 3.2, we need to satisfy the condition $\theta_c > 2\theta_r$. When the projector projection angle, the camera field of view and the radius of the inspected pipe are fixed, the above condition can be satisfied by changing D. Thus, our sensors can be used across pipes of varying diameters by changing the length of the metal rod that connects camera and projector holders or by simply adjusting the position of the nuts on the metal rods if the change in diameter is not large.

Using the set-up in figure 3.2(a), experimental data is collected from pipes with three varying inner diameters (7.5cm, 10 cm, 15 cm). Four different types of defects that usually occur in plastic pipelines as (a) dent (b) hole (c) material loss, and (d) slit are considered here. These defect types trigger highly different alarm levels for pipeline maintenance with holes being the most detrimental type and need immediate maintenance. Figure 3.4 shows pipe samples containing various types of defects.

Defect of varying types and lengths produce different kinds of distortions in the laser rings at different positions on the circumference. Figure 3.5 shows the distortions (in red) due to different defects. The discontinuity in the laser ring as seen also in images from non-defective pipe sectors (see Figure 3.5 bottom images), is due to the endoscopic motor rod that hosts the laser and the camera. In images containing defects, such discontinuity is also accompanied with perturbations in the circular ring. Three different laser intensities are used in data collection.

3.3 Defect classification using deep learning

In this section we analyze a series of deep learning models starting from 2D CNN and thereafter gradually increasing modeling complexity by incorporating spatial features in the data. Significant

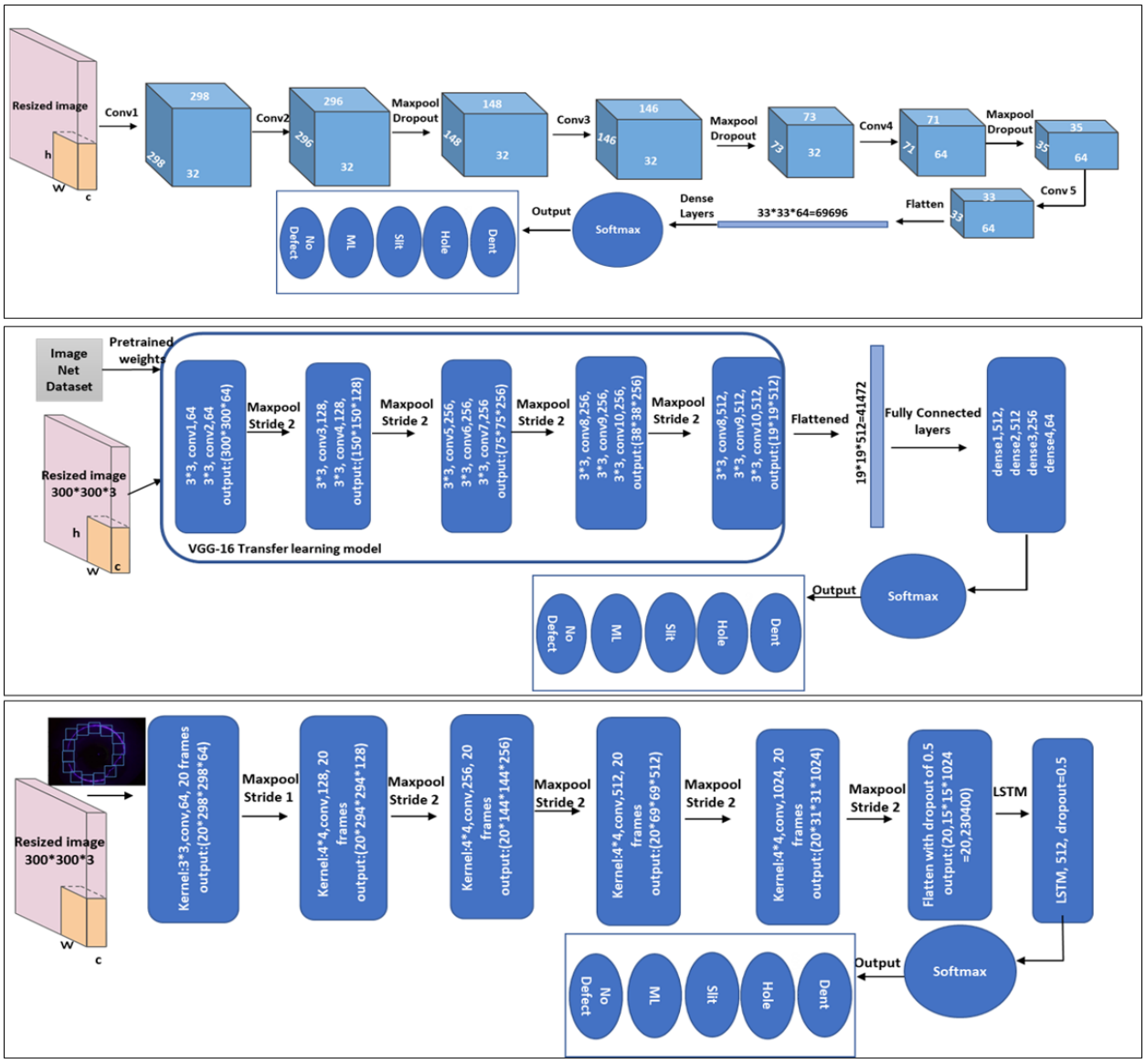

Figure 3.6 From top to bottom, network architectures of (a) CNN, (b)transfer learning aided CNN and (c) Structure adaptive CNN-LSTM are presented.

effort was spent to optimally calibrate each of the framework and the optimal models in each framework is presented and their predictive accuracy is compared. The laser-scan images are resized to 300×300 dimensional inputs for all the models. All the models except HBP were executed using Keras and Tensorflow and HBP was implemented using pytorch in the Google Colab cloud based configuration. All the models are run for 100 epochs with batch size 16.

3.3.1 2D CNN and transfer learning aided CNN

2D CNN deals with data in grid formats (78; 79) learning features from lower layers to higher layers by constructing a deep neural hierarchical architecture (80). Tuning the network over several iterations lead to the usage of 5 convolutional layers followed by 4 fully connected dense layers. Convolution layers are followed by max pooling layers to reduce dimension. In max pooling layers, dropout is also used to introduce regularization. Figure 3.6 shows the detailed architecture which is described below.

The first convolution layer on the 300×300 dimensional image input has 32 convolutional 3×3 kernels with stride 1. The output dimension $298 \times 298 \times 32$ reduces to $296 \times 296 \times 32$ on application of the second convolution layer. Thereafter, a 2×2 maximum pooling window is used outputting $148 \times 148 \times 32$ array. To reduce regularization and to prevent overfitting, a dropout layer with dropout parameter set to 0.25 is introduced before a third convolutional layer reduces the output dimension to $146 \times 146 \times 32$. Similarly, after the application of another maximum pooling and dropout, the output dimension is $73 \times 73 \times 32$. A fourth convolutional layer containing 64 convolutional kernels is then applied and the output dimension is $71 \times 71 \times 64$. After the application of another maximum pooling and dropout the output dimension becomes $35 \times 35 \times 64$ which is further reduces to $33 \times 33 \times 64$ by the fifth convolutional layer. The output is flattened and the dimension becomes $33 \times 33 \times 64 = 69696$, and four dense layers are then used. Categorical cross entropy is used as the loss function and classification of defects is done using SOFTMAX activation in the last layer. The convolutional and dense layers use the Rectified Linear units (RELU) as the non-linear activation function. It helps with the vanishing gradient problem and have faster convergence rates than that of Tanh and Sigmoid functionals (81; 82). Adam is used as

the optimizer in the model.

Next, 2D CNN model with VGG-16 as the transfer learning model is developed. Transfer learning (TL) constitute a set of extremely powerful information assimilation technique for retaining knowledge gained from one problem and applying it to a different but related problem (83; 84; 85). Here, information learnt by VGG-16 model (86) on ImageNet dataset which contains around 3.2 million images (87) is used for our laser-scan imaging data. Figure 3.5b shows the schematic of the network architecture. VGG-16 contains 13 convolutional layers followed by 3 fully connected layers. We used pretrained weights for the 13 convolutional layers of VGG-16 while retraining the fully connected layers. The feature vector thus obtained is flattened and retrained on 4 fully connected dense layers which use RELU activation and dropout of 0.3. RMSprop with a learning rate of 0.0001 is used as the optimizer.

3.3.2 Laser-scan structure adaptive CNN-LSTM model

The goal here is to leverage structural information in the laser-scan images which contain circular rings amidst the pipe background. We use CHT (66) to identify the region of interest containing the laser ring neglecting the other non-significant areas in the images. Each circular ring is segmented into 20 overlapping rectangular frames using the `OpenCV` (88) python package as follows:

- Image is converted to gray scale and threshold filters are applied as the laser intensity varied over different images.
- The laser ring center (c_1, c_2 and radius r is estimated by using circular hough transform (CHT) function. Then, the center of the rectangle falling on the arc is derived for ith frame as $(c_1 + r\cos(2\pi i), c_2 + r\sin(2\pi i))$ for $i = 1, \ldots, 20$.

Figure 3.7 shows the 20 frames thus obtained for an image. Each rectangular frame contains a part of the circular arc arranged in the form of sequences. To maintain the spatial proximity order among the frames, we develop a time distributed sequential architecture (see Figure 3.6c) which contains 5 time distributed convolutional layers followed by a LSTM layer. The LSTM layer is implemented using keras (89), a deep learning library that uses Tensorflow at the backend. Figure 3.8 shows the time distributed layer setup where each frame is subjected to a convolutional layer followed

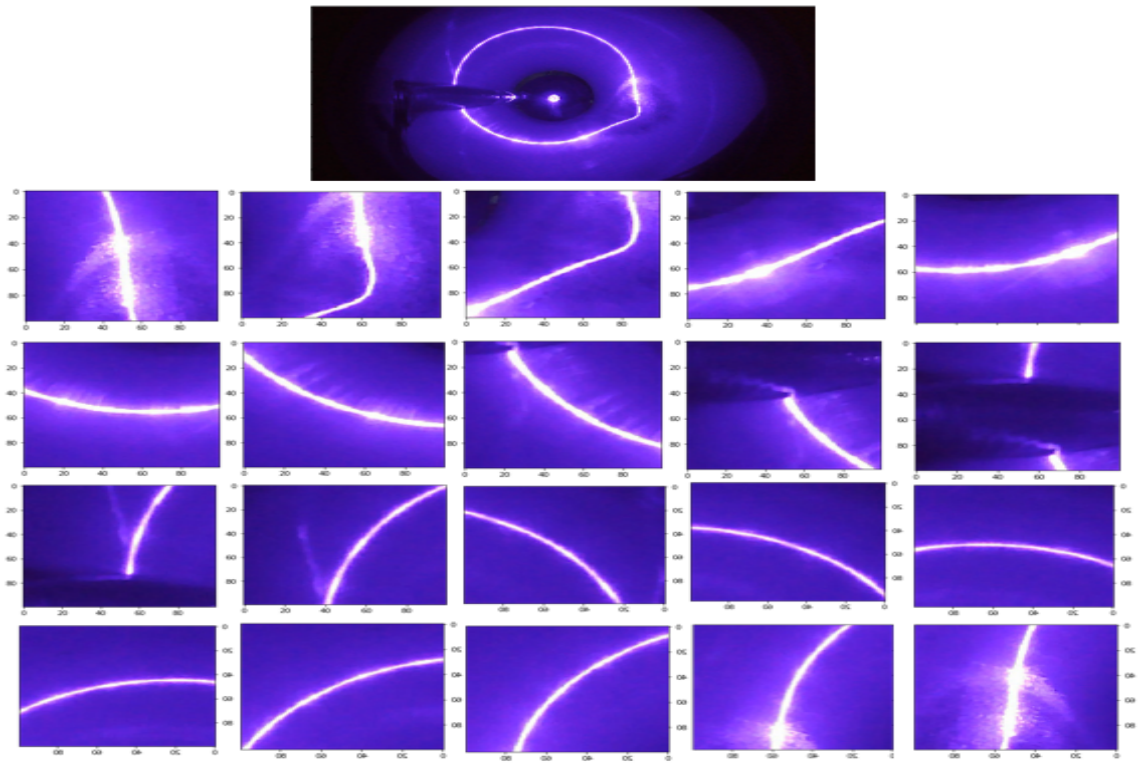

Figure 3.7 Laser ring segmentation using CHT in structure adaptive CNN-LSTM.

by a maxpool layer. Each set of convolution and max pool layer is repeated 5 times. The frames segmented on the basis of CHT are shown in block I and the time distributed CNN and maxpool operations in block II of Figure 3.8. After the final convolution layer, the output features are flattened and a dropout of 0.5 is applied. The convolved features of each frame in the sequence are then represented as a 1 dimensional vector. The sequence of these vectors is then feed into a LSTM layer parameterized with an output dimension of 512 and the return-sequence is set to false. The yellow subplot of figure 3.8 shows block III which is the structure of a LSTM cell. It uses four functions f, i, g, O_p to calculate c_t and h_t which are the cell states and hidden states respectively at iteration t:

$$c_t = f \circ c_{t-1} + i \circ g \text{ and } h_t = O_p \circ \tanh(c_t), \text{ where,}$$

$$\begin{pmatrix} i \\ f \\ O_p \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} \boldsymbol{W} \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$

and W is the weight matrix, σ is the sigmoid function and x_t is the input at the t^{th} iteration. The

output is finally obtained by flattening. Stochastic Gradient Descent (SGD) with initial learning rate of 0.01 and having 10000 decay steps with 0.9 rate is used as the optimizer. The convolution and LSTM layers use RELU as the nonlinear activation and as before, SOFTMAX is used to classify into various classes.

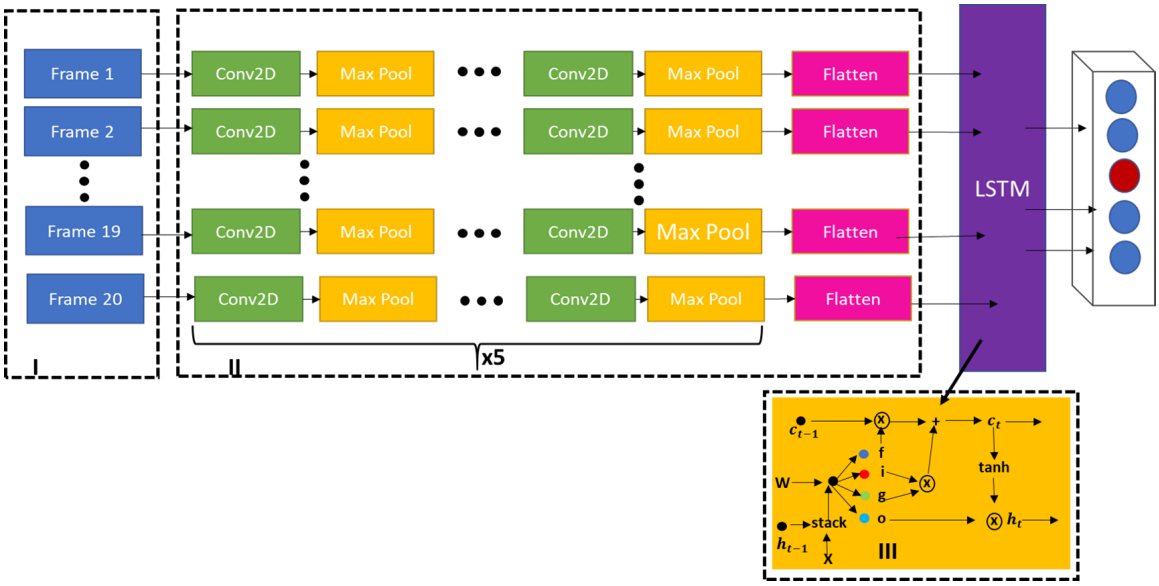

Figure 3.8 Schematic of the CNN-LSTM hybrid framework with fragmented sequential frames produced by CHT, the convolutional and the LSTM framework in blocks I, II and III respectively.

3.3.3 Hierarchical cross-layer bilinear pooling

Different categories of defects have different kinds of deviations from the circular laser-ring corresponding to non-defects. We conduct fine grained image recognition to capture these subtle laser-ring deviations among the different defect types. We concentrate on local features to accurately detect the subtle differences that can distinguish different defect sub-categories. We implement fine grained recognition using hierarchical cross layer bilinear pooling (73) which fuses inter-layer features enabling the model to focus more on the crucial part of the object.

There exists other methods for fine grained recognition (90; 91) where at first the foreground object is localized by bounding boxes or by annotations as done manually in figure 3.5 and then discriminative features are extracted for further classification. These methods being time-consuming and expensive, cannot be used for our objective. Simple bilinear pooling (92; 93) that uses activations of the last convolutional layer as representation of the images often misses minute

local attributes. Moreover, there is loss of information in the propagation of CNN. Integrating multiple cross layer features is an important aspect of HBP where interaction of part features is obtained from multiple layers instead of a single convolutional layer. Before the final classification task is conducted, cross layer bilinear features are concatenated to make full use of the intermediate convolutional layers.

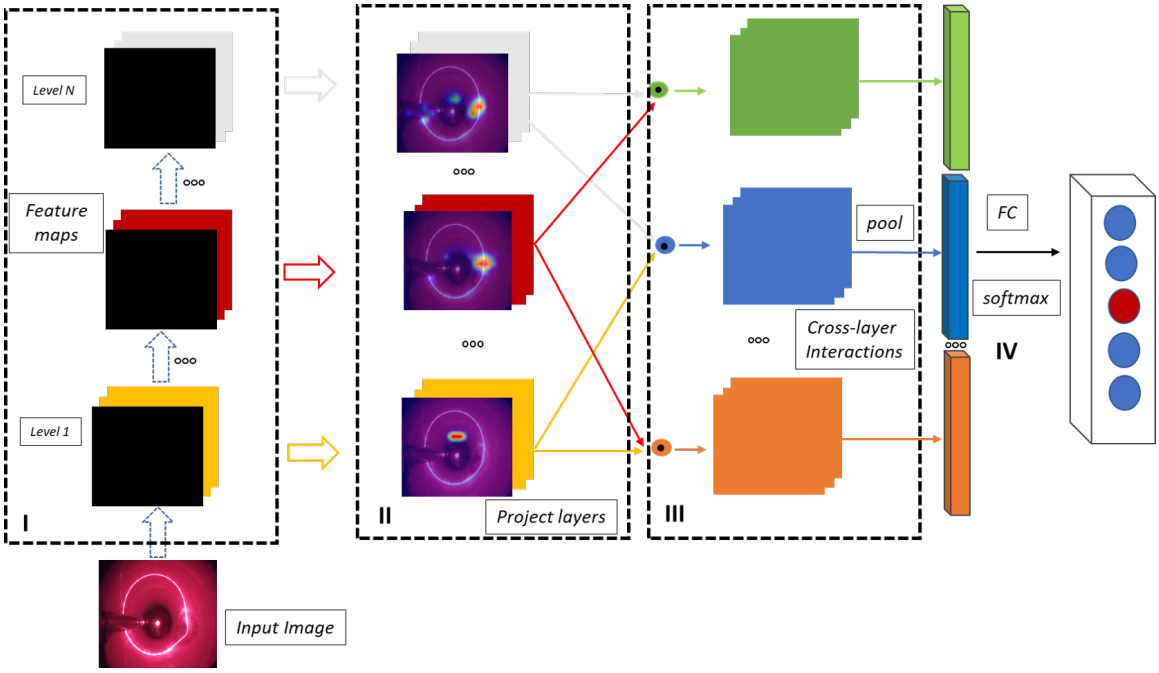

Figure 3.9 Working principle of HBP involving feature extraction by convolutional layers (leftmost block I), projections to higher dimensions in block II and integration by element wise crossing as in block III (left to right in figure).

Figure 3.9 describes the working principle behind HBP. In the figure, block I shows local features being extracted from an input image by different levels of convolutional layers. Feature map represents the output of a single channel in a convolution layer whereas the activation map denotes the feature map output of all the channels. Let I be an image and F be the output feature map obtained after being filtered by a convolution layer; $F \in \mathbb{R}^{h \times w \times c}$ where h is the height, w the width and c represent the number of channels. We write a c dimensional object at any spatial location on F by $\boldsymbol{f} = [f_1, \ldots, f_c]^T$. The output Y_i is a bilinear model given as a quadratic form involving $\boldsymbol{f}$ and a $c \times c$ unknown projection matrix W_i, i.e., for $i = 1, \ldots, n$:

$$y_i := \boldsymbol{f}^T W_i \boldsymbol{f} = U_i^T \boldsymbol{f} \circ V_i^T \boldsymbol{f} \tag{3.1}$$

By QR decomposition W_i can be factorized into two rank one vectors $U_i, V_i \in \mathbb{R}^c$ which provides the above identity involving Hadamard product between two rank one reorientations of $\boldsymbol{f}$. To get the n dimensional output $\boldsymbol{y}$, $W = [W_1, ..., W_n]^T \in R^{nc^2}$ need to be learnt from the available data. Writing in vector format for the output vector $\boldsymbol{Y} \in R^n$ we have,

$$\boldsymbol{Y} = U^T \boldsymbol{f} \circ V^T \boldsymbol{f}, \tag{3.2}$$

where U, V are $c \times n$ matrices. Instead of using U and V, we use $c \times p$ matrices $\tilde{U}$ and $\tilde{V}$ to project the features into much higher dimension p with $p \gg n$. p is a hyper-parameter that decides the size of projection layers; in figure 3.9, these are shown in block II. Here, by independent linear mapping, the feature attributes from different layers are expanded into higher dimension space to capture the detailed components of different object parts. In Figure 3.10 we present the activation maps of a projection layer for an arbitrary defect by three different mapping methods (a) gradient weighted class activation (cam_{gb}) (b) class activation mapping (cam) (c) gradient weight (gb) (94; 95). All the different visualization methods show that after projection the anomalies in the laser-ring due to the defects become more prominent in the projection layer.

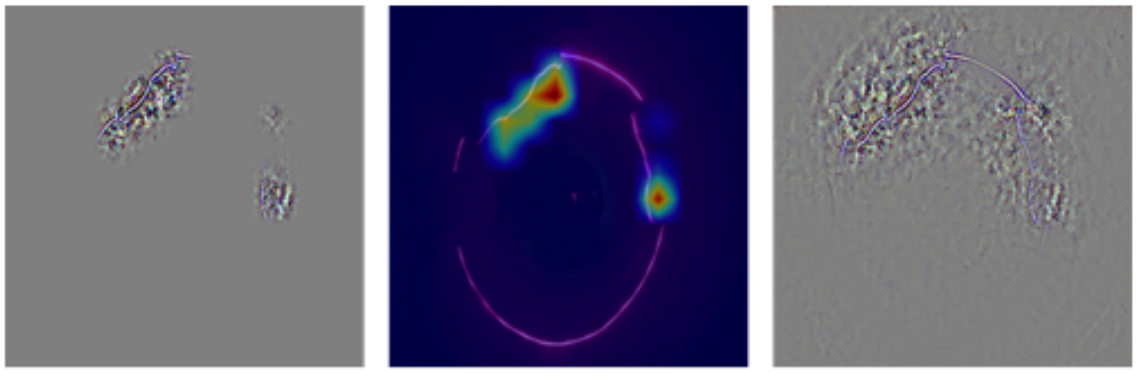

Figure 3.10 Activation maps of a projection layer from the laser-ring data with an arbitrary defect; from left to rights visuals based on gradient weighted class activation, class activation mapping and gradient weight methods.

Unlike the above equation, the output based on the projected layer involves the $p \times n$ classification matrix P as

$$\boldsymbol{Y} = P^T(\tilde{U}^T \boldsymbol{f} \circ \tilde{V}^T \boldsymbol{f}) \ . \tag{3.3}$$

In cross layer bilinear pooling which is represented in block III of figure 3.9 the features obtained from the projection layers are element wise multiplied to obtain the inter layer interactions. Let

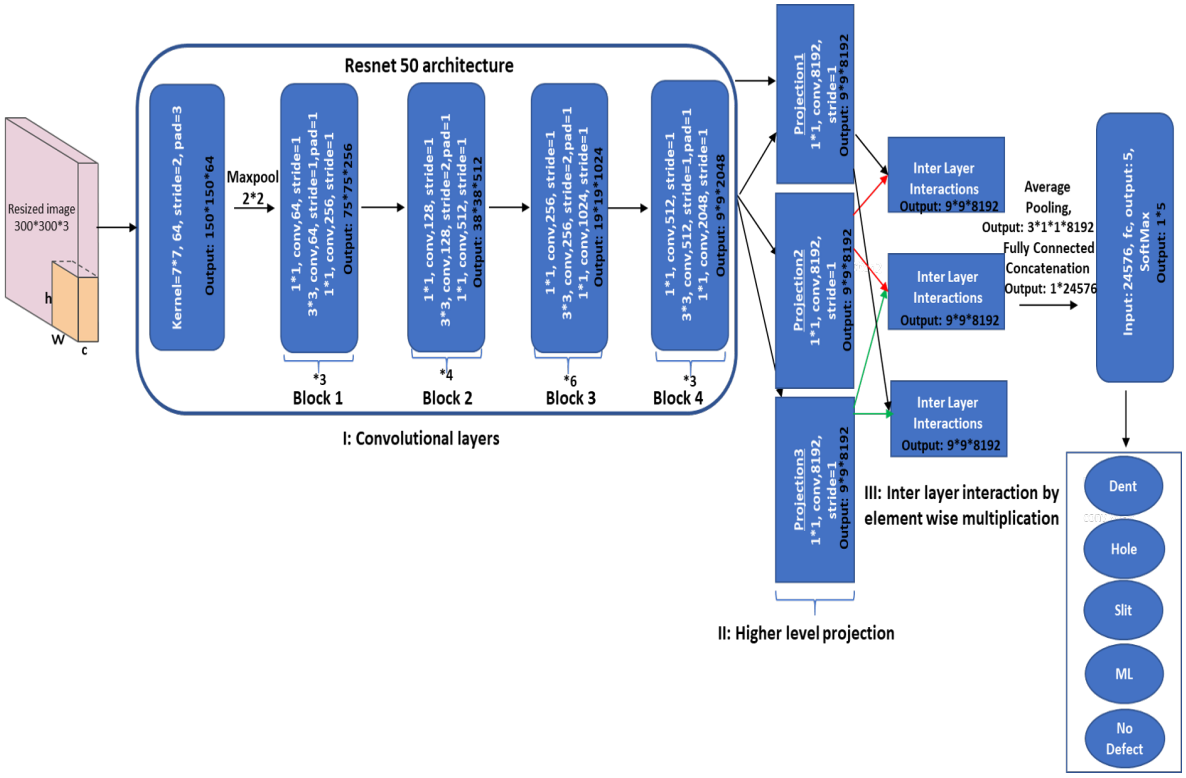

Figure 3.11 Schematic of the different layers in the developed hierarchical bilinear pooling (HBP) network.

$\boldsymbol{f}[j]$ denote the feature set from the jth convolution layer. If there were b convolution layers used, for each $i = 1, \ldots, n$ consider the vectors,

$$S_{c,1}[i] = \{\tilde{U}_i^T f[j_1] \circ \tilde{V}_i^T f[j_2] : 1 \leq j_1, j_2 \leq b\}$$
$$S_{c,2}[i] = \{\tilde{V}_i^T f[j_1] \circ \tilde{U}_i^T f[j_2] : 1 \leq j_1 \neq j_2 \leq b\} \tag{3.4}$$

where $f[j_1], f[j_2]$ represent local features at the same spatial location from convolutional layers j_1 and j_2 respectively . The column $S_c[i] = S_{1c}[i] \cup S_{2c}[i]$ is a concatenation of $b(2b-1)$ cross linear features each of length p. It has been constructed in a hierarchical fashion from convolution layers. The resultant size of $S_c[i]$ is $bp(2b-1)$ and is much larger than the projected dimension p of each convoluted layer. Consider the matrix S_c formed by stacking the columns $\{S_c[i] : 1 \leq i \leq n\}$. For estimating $\boldsymbol{Y}$, HBP uses a linear model on S_c:

$$\boldsymbol{Y} = P_c^T S_c \ . \tag{3.5}$$

The size of the projection matrix P_c is $bp(2b-1) \times n$. We need to learn P_c as well as $\{\tilde{U}_i, \tilde{V}_i : 1 \leq i \leq n\}$. For that purpose, we use Stochastic Gradient Descent (SGD) with initial learning rate as 0.01 with momentum as 0.9 and weight decay as 0.00001. The learning rate decreases by a factor of 10 for every 32 epochs.

We use pretrained weights from the convolutional layers of `Resnet50` and obtain three convolution layers each of dimension 2048 from a 300×300 image. The detailed architecture is presented in figure 3.11. `Resnet50` (87) is a 50 layer deep learning network where the first layer contains a 7×7 kernel of 64 convolutional with stride of 2 and padding 3. Block1 contains 3 layers repeating 3 times. Block 2 contains $3 \times 4 = 12$ layers stacked, block 3 contains $6 \times 3 = 18$ and block 4 contains $3 \times 3 = 9$ layers and the final layer is a fully connected layer in ResNet architecture. The fully-connected layers from the original `ResNet-50` architecture has been removed and several layers of high projection, inter layer interactions, average pooling are developed. The added layers are at first trained keeping the other layers fixed and then the entire network is fine tuned to update parameters in all the layers. The output dimensions from each block is reported in figure 3.11. Features obtained from the different convolutional layers are linearly mapped for higher dimensional projection. Then, inter-layer interaction is implemented by element wise multiplication of the projected layers. After integrating features of different parts, average pooling is performed to squeeze the high dimensional features. Finally, all the features are concatenated together to obtain the final classification scores through fully-connected layers. We use three different choices for projecting these convolution layers to higher dimensions, viz., $p = 4096, 8192, 16384$. We present the results from all the above discussed models in the following section. The HBP saturates for $p = 8192$ which is subsequently used for testing accuracy of our proposed procedure.

3.4 Results and Performance Analysis

As we consider four different defect types we have a classification problem with five different categories with no deformities being the default fifth category. We use 2160 training images and 539 images for testing the out-of-sample performance of the different defect identification algorithms. There were 539 and 123 laser-ring images with no defects in the training and test data

respectively. The data set was collected at three different laser intensities and in pipes with three different diameter lengths. In addition to the deep learning frameworks of the previous section, we evaluate and report classification accuracy for conventional machine learning algorithms. We used *scikit-learn* package to implement these classifiers. Before applying these classifiers meaningful features were extracted from the images.

The feature extraction was done with scale invariant feature transforms (SIFTs) and performing vector quantization following the bag of words (BOW) model as prescribed in (96). For SVM various kernel choices such as RBF, Sigmoid, Linear, Polynomial were used. For random forest, the number of estimators was used as a tuning parameter; it was optimally calibrated to 85. For K-NN, $K = 3$ was the optimal calibration.

Table 3.1 Train and Test set accuracy of different classifiers.

Classifiers	Type	Train Accuracy	Test Accuracy
Deep Learning	CNN	99.57%	83.85%
	CNN + Transfer Learning	93.77%	89.80%
	Structure adaptive CNN + LSTM	91.62%	85.53%
	HBP	100.00%	98.33%
SVM	RBF Kernel	94.81%	72.17%
	Sigmoid Kernel	75.46%	67.90%
	Linear Kernel	47.92%	47.68%
	Polynomial Kernel	93.15%	64.38%
Random Forest		100.00%	68.65%
Decision Tree		100.00%	62.34%
LDA		61.62%	55.47%
KNN		79.44%	57.88%
Regression	Linear	71.29%	0.04%
	Logistic	47.82%	47.68%

Table 3.1 shows the accuracy based on different classifiers. It is evident from the table that the conventional machine learning models perform poorly on the test data in comparison to that of the deep learning models. Random forest and decision trees over-fitted the data. SVM with RBF kernel has the highest test accuracy among the conventional methods. Though its prediction accuracy is

around 25% better than logistic regression - the most popular non-linear classification tool, it is still 10% worse than the naivest deep learning method. Figure 3.12 shows the model accuracy and loss for the deep learning networks described in table 3.1. The plots shows good convergence of the optimizing algorithms used for estimating the four different networks.

CNN had test accuracy of 83.85%. Incorporating structural adaptivity via CHT and LSTM did not significantly enhance performance. However, using pretrained weights from large image databases helped and we witnessed a 6% increase in predictive accuracy by using transfer learning. Using HBP we obtained a test accuracy of 98.33%. The test loss was 0.0067 and the train loss was 0.0002. The high classification accuracy of HBP shows that laser profilometry equipped with HBP can be a powerful NDE method for defect detection and characterization in plastic pipelines. Figure 3.12 shows that the HBP attains the highest test accuracy when projected into higher dimension of $p = 8192$.

We compare the deep learning methods based on Receiver Operating Characteristics (ROC) and Precision Recall (PR) curves. Both ROC and PR curves are effective measures for multi-classification problems. In ROC the true positive rate or recall is compared against false positive rate (FPR):

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive } + \text{False Negative}}$$
$$\text{FPR} = \frac{\text{False Positive}}{\text{False Positive } + \text{True Negative}} .$$

Recall is a measure of the sensitivity as lower value of recall signifies that the algorithm is missing the defects. In NDE problems, lower value of recall is unacceptable as missing defects can be detrimental. Precision defined as follows:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} ,$$

provides a measure of the cost of the procedure for low precision means that the procedure is incorrectly classifying non-defects as defects. While this is not as detrimental as missing defects, replacing too many non-damaged pipe sectors increases maintenance costs and is not sustainable for regular operations. Thus, for efficient NDE procedures, high precision and high recall rates

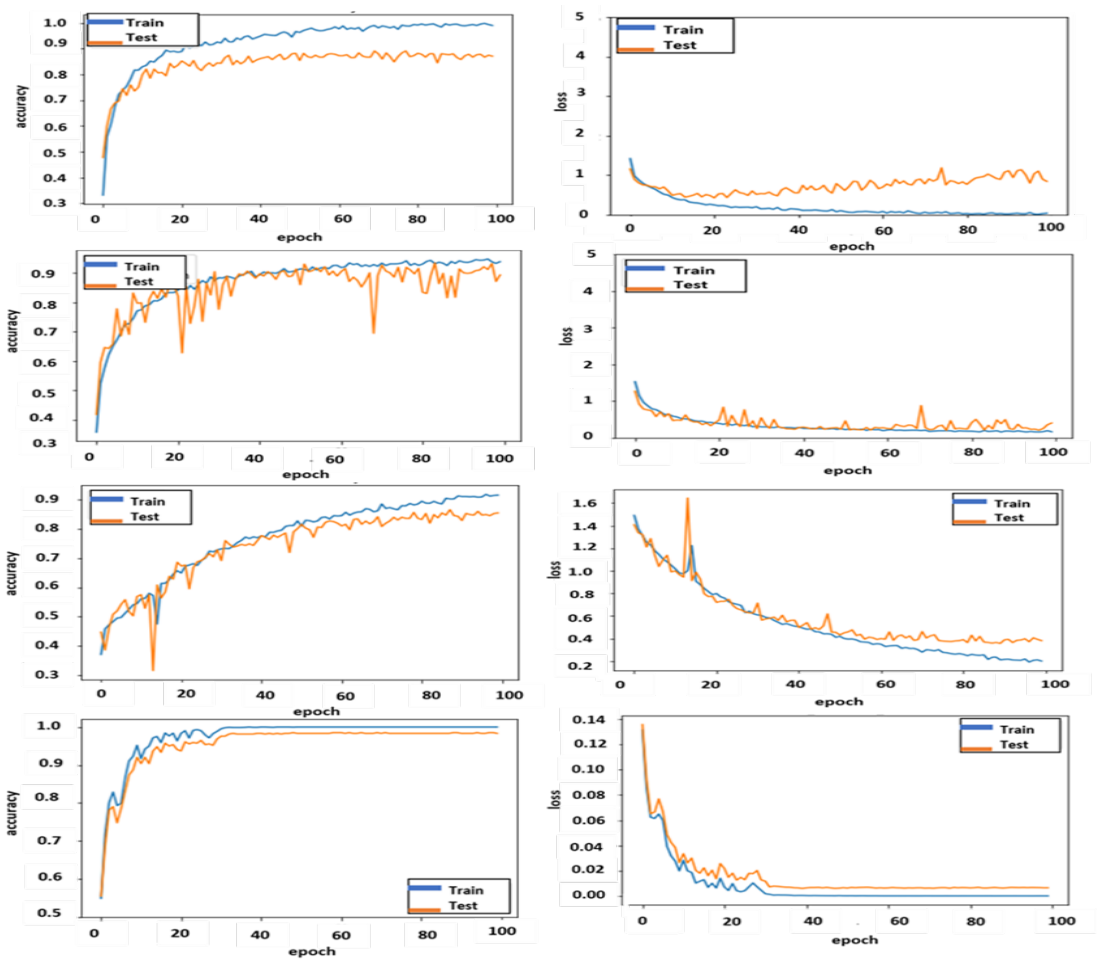

Figure 3.12 Model accuracy (left) and model loss (right) of the different deep learning frameworks on laser-scan data. From top to bottom we have CNN, CNN + TL, structure adaptive CNN + LTSM, HBP.

are desirable. For an ideal model, the area under curves (AUC) should be 1. In PR curves, precision is plotted against recall at different cut-off thresholds (18; 19). ROC curves are helpful for objective decision making when the observations among different classes are balanced. PR curves are popularly used in imbalanced scenarios (97; 98). Figure 3.13 provides the comparison of the effectiveness of the CNN, CNN with Transfer learning and HBP models based on ROC and PR curves. The AUCs for each defect types and non-defects are also reported for the three models. In figure 3.14, we observe that HBP not only provides the best performance across all the different defect types but also provides very high precision and recall rates.

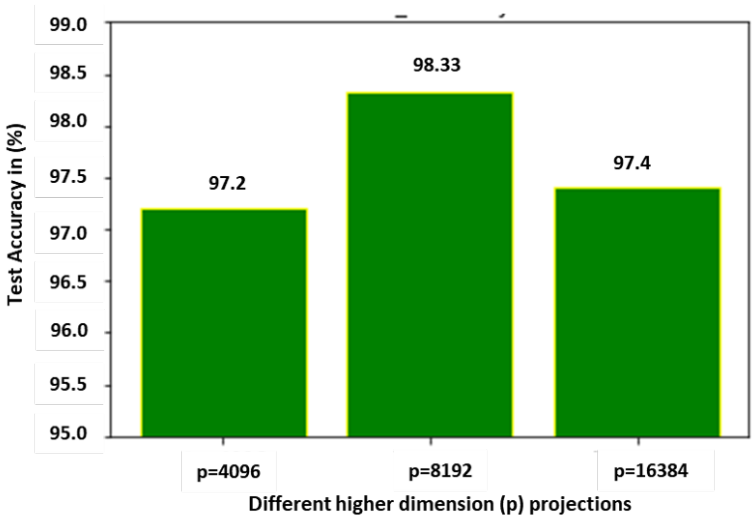

Figure 3.13 Bar plot showing the accuracy of HBP for higher dimension projections.

3.5 Conclusion

A laser profilometry based defect detection and classification procedure that uses crossed layered hierarchical bilinear pooling (HBP) deep learning architecture and other hybrid deep learning frameworks for analyzing the collected images is developed. By allowing inner layer feature interactions, HBP allows fine grained visualization and can successfully detect and distinguish minute deformities that are caused by different kinds of defects in circular laser ring images captured by our proposed procedure. Due to this fine grained visualization ability, the proposed procedure significantly out-performed a gamut of state-of-the-art conventional machine learning tools and deep learning architectures as shown in section 3.4. The proposed endoscopic methodology provides an automatic, cheap defect localization and classification tool that can be used for inspection of large polymer-based pipeline systems. The above discussed laser scanning and analysis framework is highly flexible in detecting defects across pipelines of varying radius and light intensities and can be used in a host of online NDE methods which involve defect detection based on image analysis of polymer materials.

Portions of this chapter appeared in the publication (99).

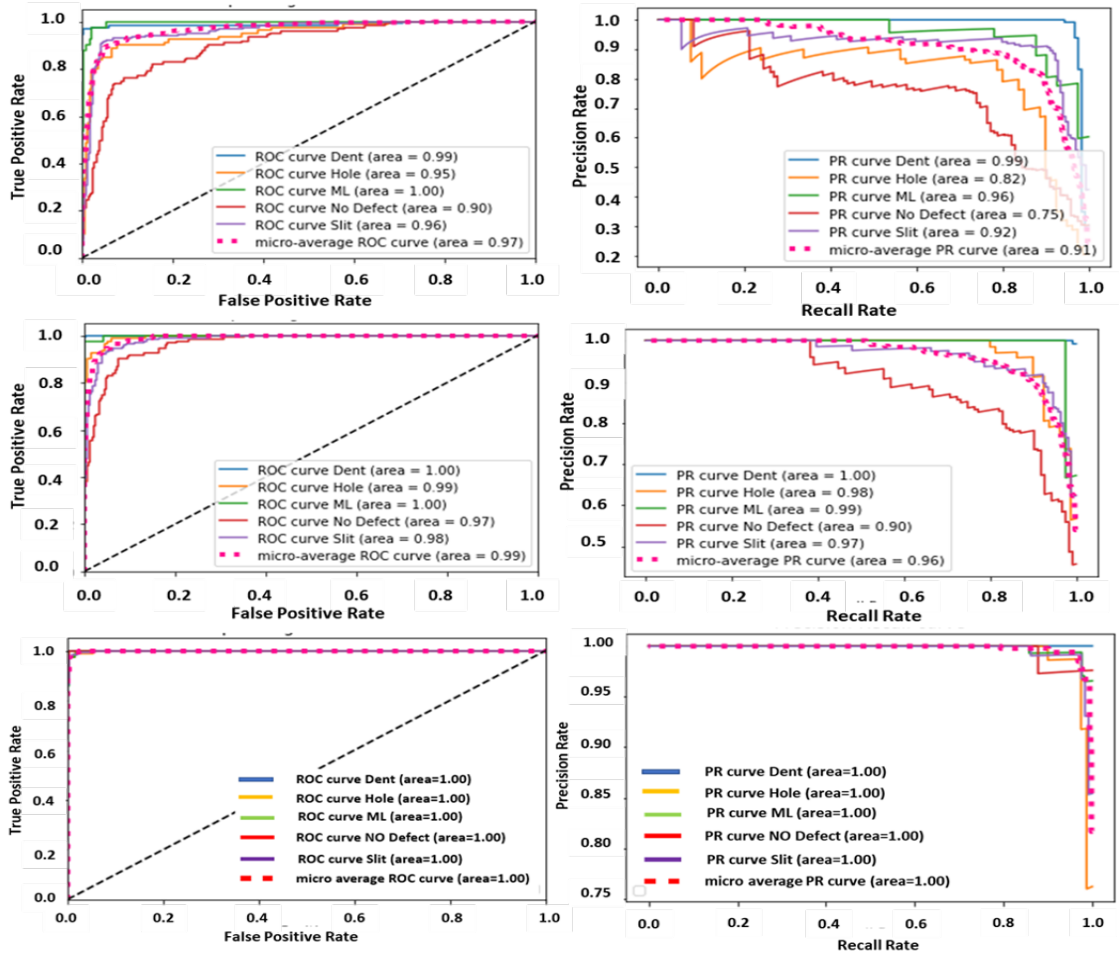

Figure 3.14 ROC (left) and PR (right) curves for different deep learning frameworks on laser-scan data. From top to bottom the plots corresponds to CNN, CNN + TL, HBP.

CHAPTER 4

KRIGING BASED SUB-SAMPLING SCHEME FOR FAST AND EFFICIENT NDE BASED DEFECT DETECTION

4.1 Introduction

In this chapter an under-sampling based data processing scheme via kriging (100) is proposed. To implement the kriging algorithm both experimental and simulated data from Magnetic flux leakage (MFL) based non-destructive evaluation (NDE) has been generated. Monitoring the condition and performing effective diagnosis of the defects within the pipeline is a necessity as the structural integrity of the pipeline decreases with time. Many oil and gas pipeline failures have led to fatalities and significant loss of properties in recent past, hence developing inspection techniques to access their conditions is of considerable importance (61; 62; 101). Since the 1960s magnetic flux leakage (MFL) serves as one of the most widely used non-destructive evaluation (NDE) technique for in-service inspection of pipelines as the pipeline materials are mostly ferromagnetic (102; 103; 104). The working principle of MFL is as follows: ferromagnetic pipe wall is magnetized close to saturation by the aid of a permanent magnet or a coil wound on a ferromagnetic yoke; presence of defects decreases the wall thickness and so, magnetic flux density and reluctance is increased in vicinity of defects (105). A higher fraction of the magnetic flux will thus leak from the pipe-walls near defects into the air. Advanced MFL signal processing techniques that can detect defects with high accuracy across different scenarios have been developed in the recent literatures (106; 29). While these MFL based methods have been very successful to detect defects in noisy situations as well as for dynamic tracking of defect sizes using transfer learning on sequential scans (28), they cannot be used in monitoring large pipelines as they involve near continuous line-scans which is extremely time-consuming. For monitoring massive pipelines, here we develop a fast-approximate algorithm that uses a minuscule fraction of data but provides similar operational performance as the aforementioned data intensive MFL methods.

The developed kriging based inspection procedure works without scanning the entire massive pipelines and only used MFL readings from a very few random scan points in the pipelines, thereby

saving humungous time and cost. The key contributions of this chapter are as follows:

1. Using kriging, we conduct spatial interpolation by using the limited set of sampled data points to estimate the magnetic flux over the concerned continuous pipe-sector.
2. Scan points in the defective areas have very different MFL features than those from non-defect regions. Here the method considers the spatial autocorrelation among the sampled data points and in this process, it allows tracking and characterization of defects from much fewer readings than those permitted in PIG based on continuous scanning.
3. The present work involves developing a 3D finite element (FEM) model in COMSOL to investigate the performance of the proposed technique in detecting the surface and sub-surface defects of various shapes and sizes. See section 4.2 for details.
4. The efficacy of our proposed kriging based fast and approximate method is tested against a wide range of defect categories such as single, multiple and interacting defects. Thresholding is used to segment defective and non-defective areas in pipelines based on MFL readings from (a) exhaustive line scans, (b) kriging interpolated predictions based on significantly smaller sample of scan points. Encouraging results (see table 4.2 in section 4.4) are obtained which shows the predicted defective areas by our proposed method to be in close resemblance with ground truth.
5. The results produced on simulated data were also validated experimentally (see table 4.3 in section 4.4) by developing a set-up in the laboratory (see subsection 4.2.2) that enabled collecting MFL readings akin to real defects.

The rest of the chapter is organized as follows. In the following section, we describe simulation procedures and lab experimental set-ups for acquiring MFL data akin to real-life scenarios. We consider various types of interacting defects in pipe-sectors which possess more threat than the singular defects. In section 4.3, we describe our kriging based fast NDE procedure and its advantages. In section 4.4 , we present the applicability of our proposed methodology. Our results in section 4.4, show that by using only a very small fraction of MFL readings from the datasets of

section 4.2, our proposed method can attain accuracy as good as extensive scan based traditional PIG methods. In section 4.5, we end with a discussion on the prescribed method.

4.2 Magnetic Flux Data Generation

For both synthetic as well as experimental data in this section we generate MFL readings for all the scan points in a metallic pipe-sector. To generate synthetic data, we design a 3D finite element MFL model in the COMSOL Multiphysics modeling software. The synthetic data generation procedure is described next. Thereafter, we present the experimental procedures.

4.2.1 Simulation Designs

Our simulation data generation process is constructed as a magneto-static problem governed by the conventional Maxwell's equation. Figure 4.1 shows the different components of a 3D MFL model in COMSOL along with a defective sample with five interacting defects of varied volumes. The different dimensions and the material properties needed to construct the model is shown in Table 4.1. The benchmark as discussed in (29) is used here to construct the model with the lift-off parameter being set at 2 mm.

Table 4.1 Dimensions (in mm) of the different components used for MFL simulations.

Term	**length**	**width**	**height**	**Material**
Yoke	400	50	40	Ferro-nickel alloy
Magnet	30	80	40	NdFeB
Brush	30	50	40	Ferro-nickel alloy
Specimen	600	400	10	X52 like iron

As shown in Figure 4.1(b) a rectangular surface consisting of multiple rectangular notch defects have been considered. Different scenarios like varying the size of notches, interaction among the notches, evolution of new defects is considered. Parametric sweep in x,y direction is conducted as shown in Figure 4.2 and the magnetic flux in axial direction B_x are collected in the form of matrices. Near the defects extremely fine triangular mesh is used whereas on the rest of the surface tetrahedral fine mesh is employed.

We use a rectangular grid of length $8cm$ and breadth $5cm$. We consider heterogeneous spacing

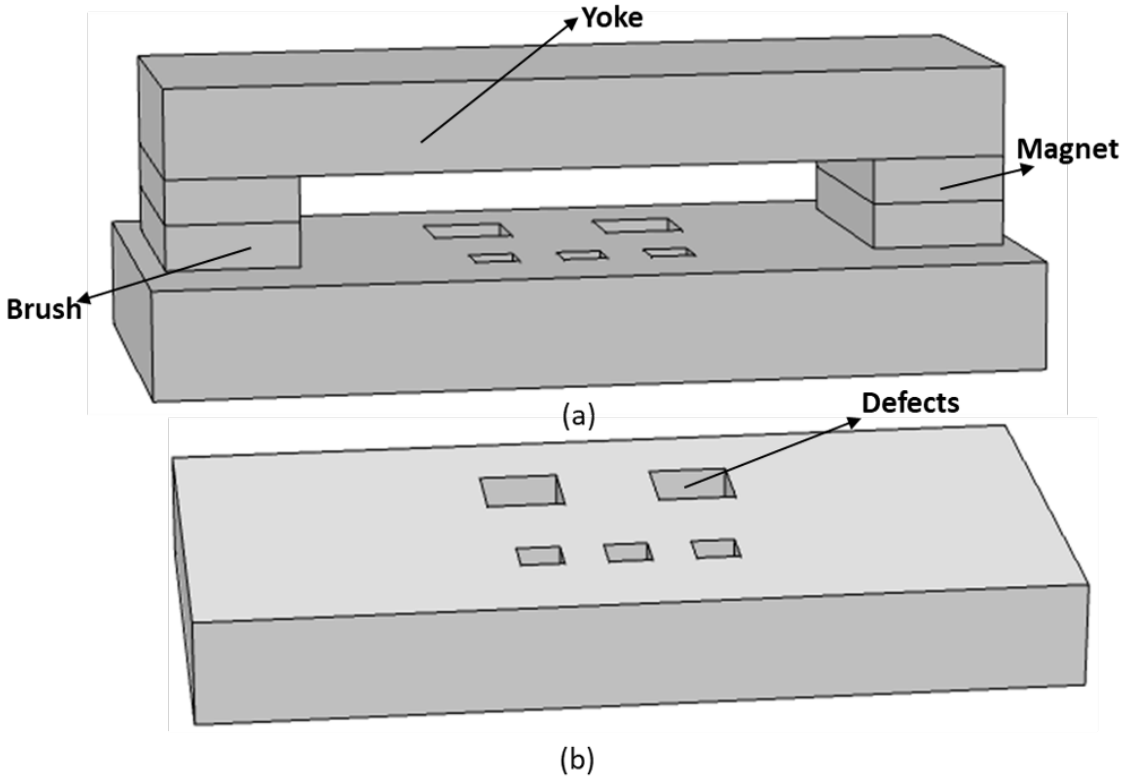

Figure 4.1 (a) Schematic of a conventional 3D MFL model in COMSOL. (b) Schematic of the sample containing arbitrary defects.

along the x and the y directions in the grid. The spacing between grid points is 0.5 mm along the x axis and 0.2 mm along the y axis. Consequently, we have $160 \times 250 = 40000$ grid points. We initially consider four interacting defects of equal sizes. Figure 4.2(a) shows the XY plot of a 2D fine rectangular grid containing the location of these four defects. The design is symmetric and balanced in defect sizes. Later we consider situations where another defect is added to a new location on the rectangular lattice and some of the existing defects evolve to grow larger in sizes. Figure 4.2(b) shows the schematic for this case with the upper ones increasing in size and a new defect cropping up on later inspection.

We generate MFL readings pertaining to five different designs. With the design in Figure 4.2(a) being the baseline, we allow the defects in the subsequent four cases to either increase in size or remain constant. The simulations reflect the perturbation in MFL readings due to not only increase in defect sizes but also due to interactions in flux leakages from the neighboring defects. The 2D fluctuations of the magnetic fields along the axial direction B_x due to spatial movement along x

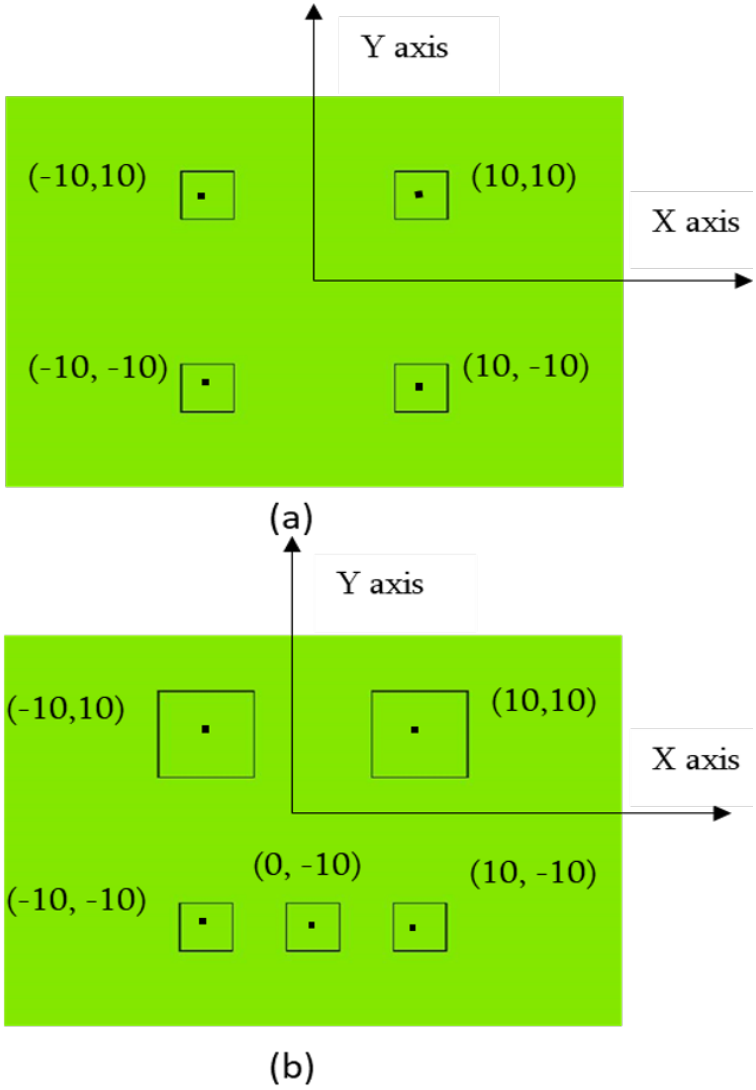

Figure 4.2 Schematic of the location of notch defects on the sample surface with the marked co-ordinates as the origin.

and y axes are observed in COMSOL from these designs. Figure 4.3 plots these MFL readings. From the figure, it is evident that in the vicinity of the defects the perturbation in magnetic field is significantly higher than those from the non-defective points.

The five different designs considered here are:

1. All the defects are squares with length of each side being 5 mm. Their locations in the rectangular lattice of length 2.5 cm and width 1.6 cm is given by the layout in Figure 4.2(a).
2. The upper two defects in case I have increased to squares of length 9 mm whereas the bottom three defects are squares of length 5 mm each. A new defect has evolved in case II compared

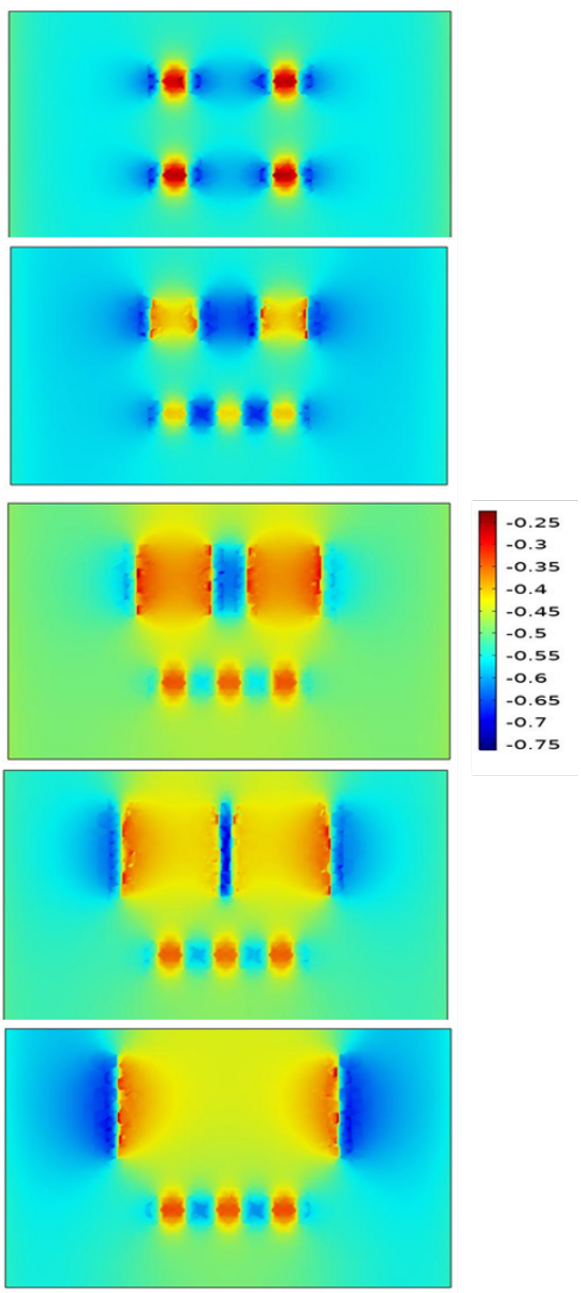

Figure 4.3 2D plots of magnetic flux densities for various defect types (Case I to V from top to bottom) as generated in COMSOL based on FEM. In all the above sub-plots, the x axis ranges from -4cm to 4cm and the Y-axis from -2.5 cm to 2.5 cm.

to case I. The layout is given in Figure 4.2(b).

3. The upper defects have increased in size to squares whose length of each side is now 14 mm.
4. The upper defects have increased to squares with 18 mm sides. The MFL reading plot in Figure 4.3 shows increasing interaction among the upper defects.

5. The upper defects have merged into a single defect on further increasing their dimension to 20 mm sides each. The lower defects remain constant in size.

Figure 4.3 contains the plots of the MFL data generated in COMSOL from the aforementioned five different designs.

4.2.2 Experimental Setup

The working principle of MFL based NDE methods is based on the fact that there will be increased magnetic flux density in the vicinity of the defects (104) which can be measured using a hall effect sensor (static and dynamic fields) or coil (dynamic fields) (105). Following this principle, we design our experimental setup using the following ingredients:

- a permanent magnet based MFL probe,
- a scanning robot arm to move the probe along the sample,
- a direct current (DC) power supply,
- a data acquisition system (DAS),
- other associated units.

The image of the setup and its associated MFL probe is shown in Figure 4.4(a) and 4.4(b) respectively. As shown in the figure 4.4(b), the permanent magnet assembly includes a magnetic circuit consisting of two permanent magnets (NdFeB) and magnetic field sensor to sense the leaked magnetic field in presence of defects. The robotic scanning arm which holds the MFL probe is used to control the movement of the permanent magnet assembly. We use an analog giant magnetoresistance (GMR) magnetometer sensor (AAH002-02E) manufactured by NVE Corporation in our sensor setup. Figure 5 shows the schematic of the entire experimental setup, the pin configuration of the GMR probe and the data acquisition system. The GMR sensor contains four resistors in the form of Wheatstone bridge configuration (107; 108). We chose AAH series of sensor as it has high sensitivity for low field sensing and excellent temperature stability. Also, the small size of the sensor makes it very convenient for mounting in the constructed MFL probe.

The axis of sensitivity of the sensor is parallel to the surface of test material. We record the changes in the axial component of the magnetic flux (B_x) based on the output voltage of the GMR sensor. When the MFL probe is far from the ferromagnetic sample then the output voltage of the GMR sensor is constant. In the presence of defects, the magnetization changes which subsequently changes (B_x) . This produces a resistance change in the GMR sensor thereby altering the output voltage. This voltage is subsequently analyzed by our proposed algorithm. A constant DC voltage of 1.98 V is given as input to our experiment. This low field range and high resolution of the GMR sensor make it ideally suited to measure the residual fields. The pivotal ingredient in our DAS was the National Instrument Data Acquisition card (PCIe-6341) which samples and digitizes the data using an imaging routine and the output is plotted on a computer. The sensor was interfaced with the data acquisition card via Digital Multimeter (DMM) to allow simultaneous data acquisition.

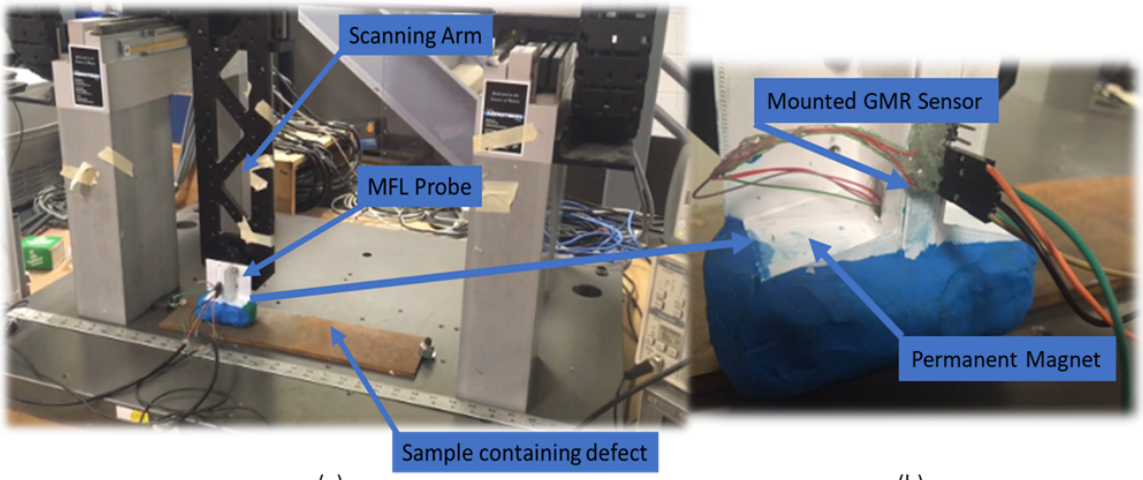

Figure 4.4 The complete experimental set up as shown in (a), the constructed MFL probe as shown in (b) consisting of two permanent magnets and GMR sensor.

We use square steel samples with 50 cm sides. We consider two different experiments one with a circular and the other with a rectangular defect. Both the defects were placed in the center of the steel samples. The rectangular defect is of length 2.5 cm and breadth 1.2 cm whereas the spherical defect has diameter 0.7 cm. Our designed MFL probe recorded flux readings at scan points 1 mm apart producing a 50×50 square grid of MFL records. Their plots are shown in Figures 4.6 (rectangular defect) and 4.7 (spherical defect).

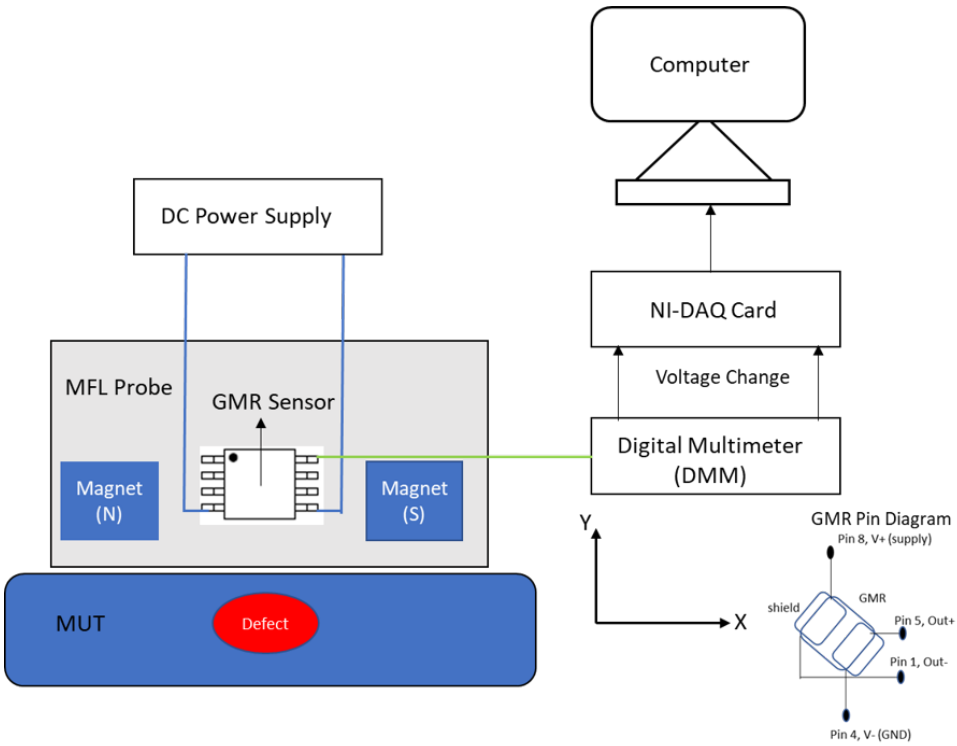

Figure 4.5 Schematic of the experimental and data acquisition setup.

4.3 kriging Based Proposed Methodology

4.3.1 Magnetic flux leakage prediction

Kriging is a method of spatial interpolation that originated in the field of mining (100). Precision provided by the state-of-the-art PIG techniques comes at a price of computational cost. In large-scale NDE applications this cost become crucial and can be prohibitive for timely damage control. Thus, it is extremely important to consider cheap surrogate methods that reduce computational cost while maintaining the required precision.

We use Kriging which is a spatial interpolation method to develop scalable NDE techniques for defect detection in large pipelines. Kriging uses a limited set of under sampled data points to estimate the variable over a continuous spatial domain. The interpolation is based on the spatial arrangement of the empirical observations, rather than on a presumed model (109). Thus, kriging generates estimates of uncertainty surrounding each interpolation. Kriging predictor is a combination of linear predictor and exact interpolator, thereby the value obtained by kriging for any actually sampled location will be equal to observed value at that point and the interpolated values

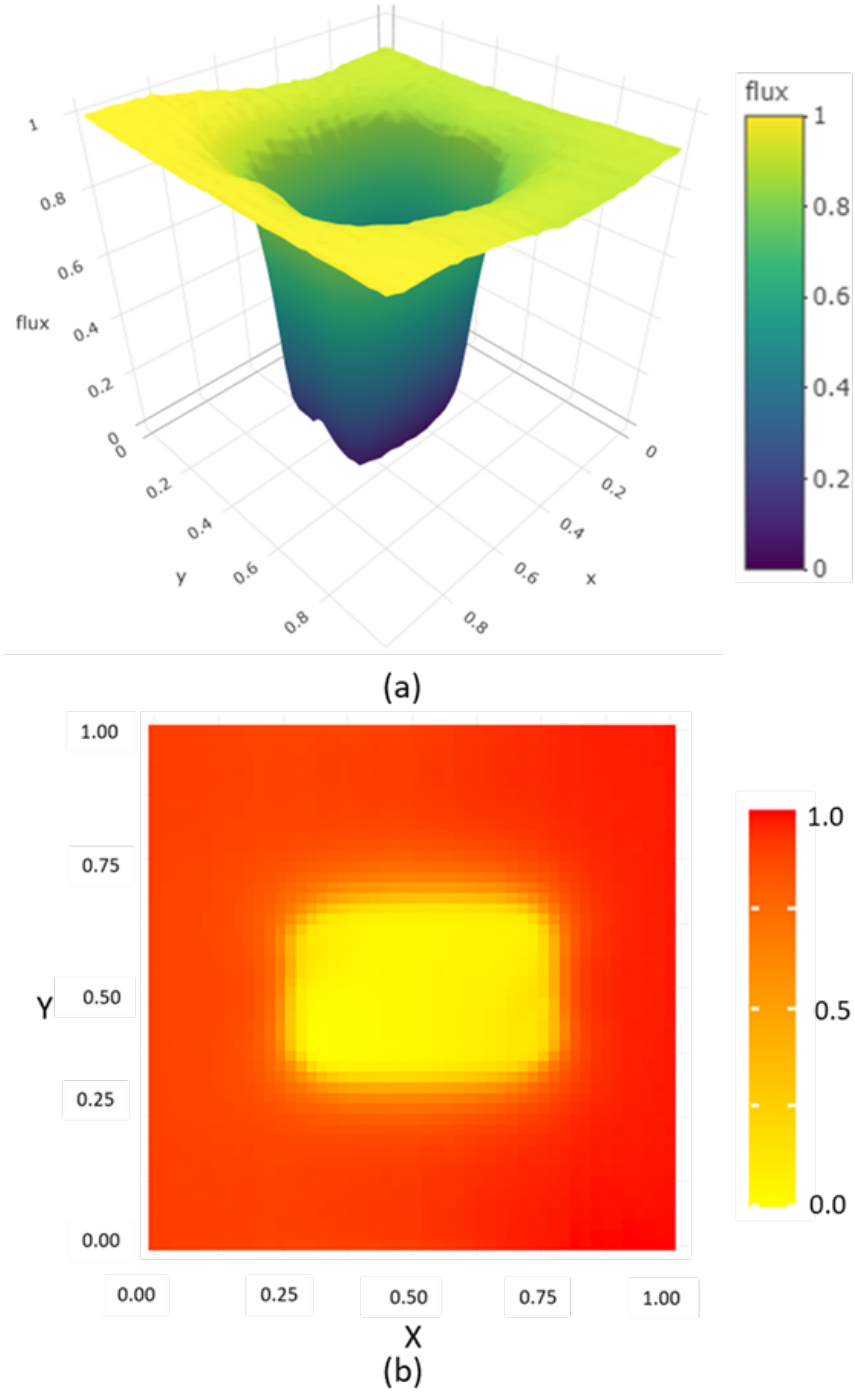

Figure 4.6 3D (in subplot a) and 2D (in subplot b) heat maps of standardized MFL readings from a steel sample with a rectangular defect in its center. The readings were obtained from laboratory designed probes using GMR sensors and two permanent magnets. Both the x and y axes have 50 points with a uniform 1mm gap between them. These axes are rescaled between 0 and 1 on the plots.

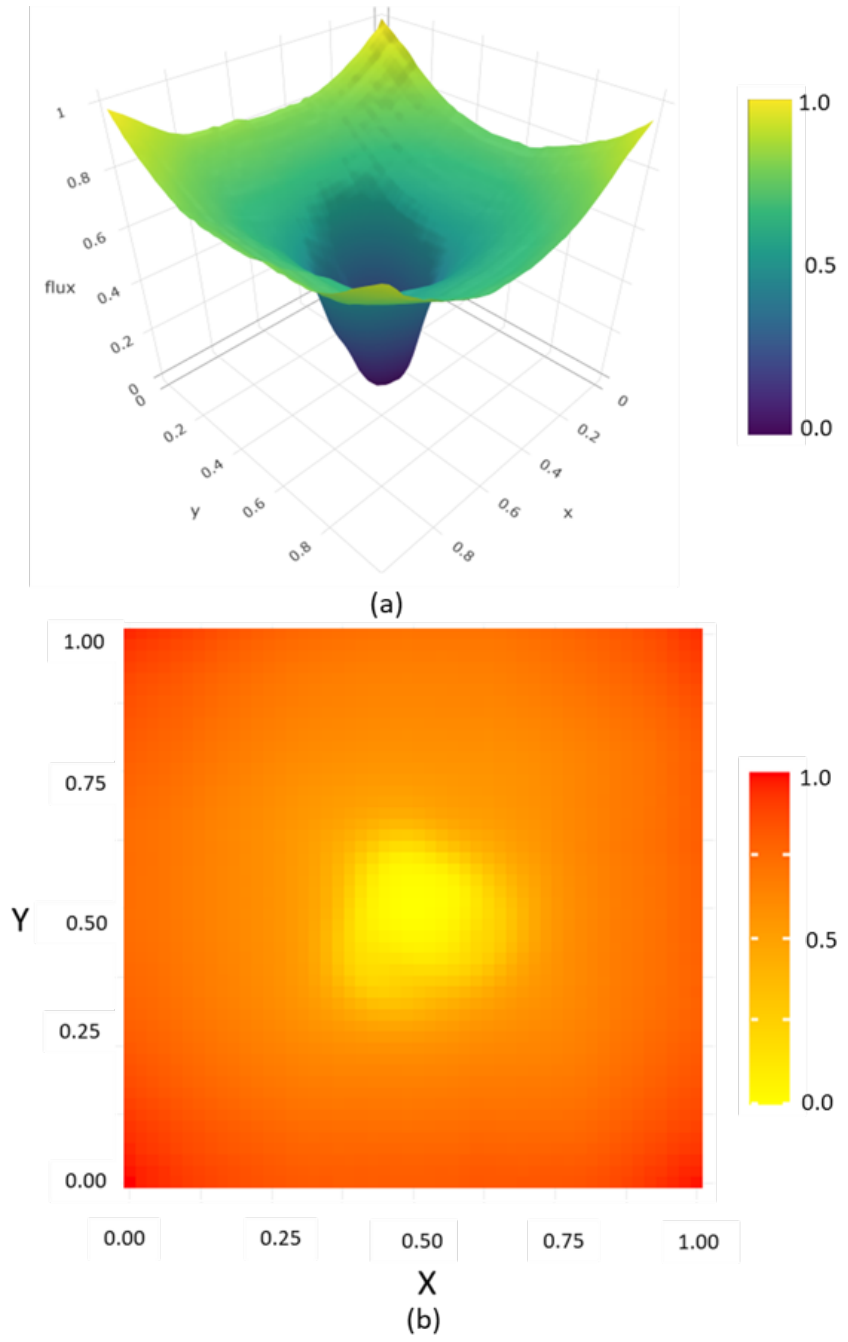

Figure 4.7 3D (in subplot a) and 2D (in subplot b) heat maps of standardized MFL readings from a steel sample with a circular defect in its center. The readings were obtained from laboratory designed probes using GMR sensors and two permanent magnets. Both the x and y axes have 50 points with a uniform 1mm gap between them. These axes are rescaled between 0 and 1.

will be the best linear unbiased predictor (BLUP).

Magnetic flux distribution depends on the location of the co-ordinate points. As there exists a strong spatial correlation among the MFL from neighbor points on the lattice, kriging can provide

a good prediction for the unobserved points by leveraging the correlations with their neighboring observed scan-points. Kriging is basically a two-step process where in the first step the spatial covariance structure of sampled points is determined by fitting a variogram which is a visual depiction of the covariance of each pair of points in sampled data. Thereafter, the weights from this variogram are used to interpolate values in the unobserved points (110).

Let G be the grid of all points used in a traditional PIG line-scan or 2D scan. Let G_S be the sub-set of points in G that are used for kriging. Let M be matrices containing standardized magnetic flux leakages readings for points in G and M_S be the set of all sampled M values at G_S. Let $\hat{M}$ be the predicted MFL values based on our procedure that takes M_S and the location of the scan-points G_S as inputs. The entries $\hat{m}_{(i,j)}$ and $m_{(i,j)}$ in $\hat{M}$ and M are the predicted and recorded MFL values based on our proposed method and PIG line-scans respectively at the $(i, j)^{th}$ point in the grid G. Kriging models the MFL values by a gaussian process. The flux f_{ij} at the $(i, j)^{th}$ location on the grid is modeled as:

$$f_{ij} = \mu_{ij} + \sigma^2 Z_{ij} \tag{4.1}$$

where, the trend function σ_{ij} can be any function with domain in $[0, 1]^2$. Let x_{ij} represents the corordinates of the $(i, j)^{th}$ point in the grid. Consider the axes of the grid G to be standardized so that $G \subseteq [0, 1]^2$. We expand the trend function with respect to the canonical basis of square-integrable functions in $[0, 1]^2$ as

$$\mu_{ij} = \alpha^{\mathrm{T}}(x_{ij})\beta \tag{4.2}$$

where, $\alpha^{\mathrm{T}}(\cdot)$ denotes basis of functions in $[0, 1]^2$ and β is the corresponding basis coefficient. These are the regression coefficients of the model. We assume that the constant variance of the gaussian process is σ^2 and Z_{ij} the local deviation from the trend functions follows independent and identically distributed standard normal distribution. Then, β corresponds to the regression coefficients. However, the trend functions at any two random points x_{ij} and x_{kl} are correlated. The correlation decreases as the distance between the points increases. We model this correlation $R(x_{ij}, x_{kl})$ between the trends at x_{ij} and x_{kl} by the following exponential spatial correlation function:

$$R(x_{ij}, x_{kl}) = \frac{1}{\sqrt{2\pi h}} \exp\left(-\frac{1}{2h}(x_{ij} - x_{kl})^2 \right) \tag{4.3}$$

where, h is the unknown hyperparameter that is tuned. Let A_S and R_S be the matrix of basis functions and correlations for the sampled G_S points, then the parameters are estimated as:

$$\hat{\beta} = \left(A_s^{\mathrm{T}} R_s^{-1} A_s\right)^{-1} A_s^{\mathrm{T}} R_s^{-1} M_s \tag{4.4}$$

$$\sigma^2 = \frac{1}{|G_s|}\left(M_s - A_s\hat{\beta}\right)^{\mathrm{T}} R_s^{-1}\left(M_s - A_s\hat{\beta}\right) \tag{4.5}$$

where, $|G_s|$ is the number of sample points. The kriging prediction (111) for the set of unobserved points $U = G \setminus G_s$ is given by:

$$\hat{M}_U = A_U^{\mathrm{T}}\hat{\beta} + \hat{\beta}^{\mathrm{T}} R_{US}^{-1}\left(M_s - A_s\hat{\beta}\right) \tag{4.6}$$

where, R_{US} is the correlation matrix between the observed and unobserved points and A_U is the matrix of basis functions evaluated at locations in U. The code is implemented in the *R* programming language using the library packages gstat and sp (112; 113; 114).

4.3.2 Sub sampling Rate & Defect Detection

The sub-sampling ratio adjusted for dimension is defined as: Sampling rate per dimension $= \rho_S = (|G| \setminus G_s)^{(1/d)}$, where $|G|$ and $|G_S|$ are respectively the number of scan points in G and G_S and d represents the dimension of the grid. For a fixed threshold h we segment points based on the standardized MFL values. The readings in Figure 4.6 and 4.7 (experimental data) from the defective points are troughs in lattice whereas that in COMSOL simulation (see Figure 4.3) constitute a crest. The orientation of the crest and troughs depend on the different data collection procedures. In our simulation experiments, the variation in magnetic flux signals is calculated directly and plotted. However, from the lab experiments the changes in the output voltage from the GMR sensor are plotted. As the defect detection algorithm used in this chapter only requires the readings from the defective region to be different from that of the background, we are not concerned with whether the readings of defective regions were uplifted or downregulated. Using threshold h and the complete PIG scan data, we segment the points in G into defective and non-defective or trouble-free sets D and T respectively. For most cases h chosen as 0.5 works well with standardized MFL values. Keeping the same value of h and the same criterion we segment our Kriging predictions $\hat{M}$ into defective and trouble-free sets $\hat{D}$ and $\hat{T}$ respectively. Ideally, we would like $\hat{D}$ and $\hat{T}$ to be very similar to sets D and T. We analyze their relations in the following section.

4.4 Results

The larger the sub-sampling rate rho_S the fewer is the number of points used in our proposed method. The greater rho_S is from 1, the faster the MFL data collection step in our procedure will be. However, if $|G|_S$ is too small then we will do a shoddy job in interpolating the MFL values at the unsampled points $G \setminus G_S$ and any subsequent inference on defect location recognition will be highly erroneous. So, we consider popular metrics that capture the operating performance of defect detection algorithms. We report the mean square error (mse) of the MFL prediction based on kriging. We also report the coverage of defect points based on our proposed algorithm as well as its false positive percentages. Next, we define these measures.

The MSE of kriging is calculated as:

$$mse = \frac{1}{|G|} \sum_{(i,j) \in G} \left(\hat{m}_{i.j} - m_{i.j} \right)^2 \tag{4.7}$$

where, $\hat{m}_{i,j}$ and $m_{i,j}$ are the predicted and recorded MFL values based on our proposed method and PIG line-scans respectively at the $(i, j)^{th}$ point in the grid G . As the raw mse values of the predicted MFLs are difficult to interpret, we present the percentage improvement in mse by using Kriging instead of the naïve average by reporting the R-squared statistic:

$$R^2 = 1 - \frac{mse}{variance\ of\ M} \tag{4.8}$$

Lower mse signifies that the predicted MFL values by the kriging model is closer to the exhaustive scan MFL values, in which case the R-squared values will be closer to 1 signifying accurate reconstruction of MFL values by kriging at the unsampled points.

Next, using the below mentioned coverage statistics we evaluate how much of the defective area found in exhaustive scan can be identified by our proposed methodology:

$$Coverage = \frac{size(\hat{D} \cap D)}{size(D)} \tag{4.9}$$

The false positive rate of our procedure is reported via the exceedance metric as:

$$Exceedance = \frac{size(\hat{D} \cap D)}{size(G)} \tag{4.10}$$

High R-squared, high coverage and low exceedance values are desired. Next, we report these

metrics as subsampling rate ρ_S is varied. We consider uniform subsampling designs throughout the chapter.

4.4.1 Results on Synthetic Data

Consider the synthetic datasets described in the five designs of section 4.2 . In Table 4.2, we report the coverage, exceedance and R-squared performance measures of our proposed methodology as the subsampling rate ρ_S varies from 5 to 30. Note that in these COMSOL simulated datasets $|G| = 40000$ and thus a sub sampling rate of $5:1$ means $40000/5^2 = 1600$ random locations in G were considered in G_S. For ρ_S equal to $10, 20 and 30$ we respectively have $400, 100 and 44$ MFL samples.

From the table, we see that when $\rho_S = 5$, we get considerably close defect identification with very low false positives across all the five cases even if we consider only 4% of the MFL readings. The R-squares of the predicted MFL values are quite good signifying that our proposed method can be used with high confidence at $\rho_S = 5$. Figure 4.8 shows the MFL predictions by our method for Case II of Table 4.2. It shows the gradual deterioration in the kriging based MFL reconstruction as sampling rate ρ_S increases. The coverage rate decreases considerably when $\rho_S = 20$ and as expected there is a complete breakdown of the method as with $\rho_S = 30$ for there are only 40 scan points.

4.4.2 Results on Experimental Data

Next, we apply our proposed method on the experimental data described in subsection 4.2.2. The different performance metrics are reported in Table 4.3, Figure 4.9 and 4.10 show that predicted readings for the cases with rectangular and spherical defects respectively as subsampling rates are increases from 2.5 to 10.

From Figures 4.9, 4.10 and Table 4.3, we see that in both the cases our method produces coverages sufficiently close to exhaustive scans till the subsampling rates reach 5. Thereafter, the coverage decreases below the 90% tolerance limit for the rectangular defect. In all these cases, the exceedance is very low. In particular, from the 3D MFL plots in Figures 4.9, 4.10 we observe that the troughs corresponding to the rectangular and spherical defects become greatly thin and

Table 4.2 Compared to exhaustive scans the performance of our proposed method in the FEM simulations designs is reported for different subsampling rates.

Defect Type	Performance Measure	Subsampling Rate			
		5/1	10/1	20/1	30/1
Case I	R-squared	92.41%	65.67%	27.07%	13/75%
	Coverage	83.93%	51.79%	25.00%	7.14%
	Exceedance	0.00%	0.25%	0/19%	0.06%
Case II	R-squared	88.67%	65.48%	47.40%	9.84%
	Coverage	73.02%	69.84%	19.05%	1.59%
	Exceedance	0.38%	0.50%	0.13%	0.00%
Case III	R-squared	93.55%	76.41%	36.82%	9.85%
	Coverage	96.67%	82.33%	73.67%	6.67%
	Exceedance	0.38%	1.94%	5.38%	0.00%
Case IV	R-squared	90.17%	77.80%	52.56%	20.00%
	Coverage	90.91%	71/16%	68.97%	6.90%
	Exceedance	0.50%	1.38%	5.38%	0.44%
Case V	R-squared	96.11%	88.14%	70.16%	27.56%
	Coverage	89.83%	84.88%	71.51%	8.43%
	Exceedance	0.56%	1.00%	3.94%	0.50%

lose much of their shape as subsampling rates is increased from 5 to 7.5. In Figure 4.11, the standard deviation of the estimates associated at each scan point is plotted. In Table 4.4 we report the mean of these standard deviation (MSD) across all the scan points. Note, that as we have used standardized reading values the range of the readings is 1. To understand the relative impact of these standard deviations we report the relative standard deviation (RSD) by dividing MSD by the average difference between the readings from defective and non-defective scan points. Higher RSD means greater confidence in correctly predicting defective scan points. In Figure 4.12 the 95% prediction surface for the two defects is also reported for different sub-sampling ratios. From Table 4.4 we observe that for subsampling rates up to 5, the MSD and RSD are well-controlled. Figure 4.12 shows that the 95% prediction surface are thin enough to provide accurate differentiation between the predicted readings from defective and non-defective scan points.

Table 4.3 Compared to exhaustive scans the performance of our proposed method in laboratory experiments is reported for different subsampling rates.

Defect Type	Performance Measure	Subsampling Rate			
		5/2	5/1	15/2	10/1
Rectangle	R-squared	99.4%	94.6%	85.0%	70.95%
	Coverage	99.3%	95.8%	81.2%	70.42%
	Exceedance	0.3%	1.5%	0.5%	3.68%
Sphere	R-squared	99.5%	98.8%	96.7%	80.47%
	Coverage	100.0%	100.0%	99.7%	80.78%
	Exceedance	0.4%	1.8%	3.2%	0.72%

Table 4.4 Uncertainty bounds for kriging-based prediction on experimental data. Mean Standard Deviation (MSD) and Relative Standard Deviation (RSD) for different sub-sampling rates are reported.

Defect Type	Measure	Subsampling Rate			
		5/2	5/1	15/2	10/1
Rectangle	MSD	0.0166	0.0622%	0.0641%	0.1367
	RSD	18.5940	4.9744	4.8274	2.2614
Sphere	MSD	0.0041	0.0185	0.0307	0.0443
	RSD	26.7102	5.9282	3.5651	2.4689

4.5 Discussion

Based on the results in section 4.4, we observe that with much fewer number of MFL readings, using our kriging-based method we can identify defective areas as accurately as exhaustive line scans. As such, in section 4.4 across a wide spectrum of synthetic and experimental datasets, we found that our kriging-based method can provide more than 90% defect identification coverages and lower than 1% false positive rates when it uses only 4% of the MFL readings used in exhaustive traditional PIG approaches. Thus, our proposed method is fast, cost-effective and highly scalable for inspecting defects in large pipelines. In future, it will be useful to study the improvements due to non-uniform sampling designs and introspect the applicability of our prescribed method to other complex NDE tasks in pipelines such as dynamic tracking of defects.

Here, we show that our proposed kriging-based detection methodology worked well in detecting millimeter sized defects which is popularly done by MFL based NDE inspections (115; 116; 117; 118). Further work is needed to see if aberrations in magnetic flux due to smaller defects can be recovered from under-sampled signals. In this chapter inspection cost is reduced by under-sampling and thereafter reconstructing the MFL signals at all scan points by using Kriging. Compared to MFL readings from extensive scans there is always some information loss in the Kriging based reconstructed signal. The success of our procedure rests on the fact that this information loss does not hamper detection of the presence and location of the defects in the metallic surface. It is to be noted that the information loss will be exacerbated if the MFL values are very noisy; this can happen if there is lift off or gap between the probe and the inspection surface. In such situations, it will be difficult to recover defect locations from under-sampled MFL readings as the signal-to-noise ratio can be very low.

Here, in lab experiments we considered $50cm \times 50cm$ metallic plates with defects. Continuous scans over large pipe surfaces will have severe imbalance between defective and non-defective scan points. In these cases, we can apply our proposed method frame-by-frame by first identifying rectangular frames containing defects and thereafter locating the defects in frames. Most frames will not contain any defect. As our proposed method is witnessed to provide good coverage and low exceedance in defective scan points detection, we expect low false discovery rate and high power in such frame-by-frame detection analysis.

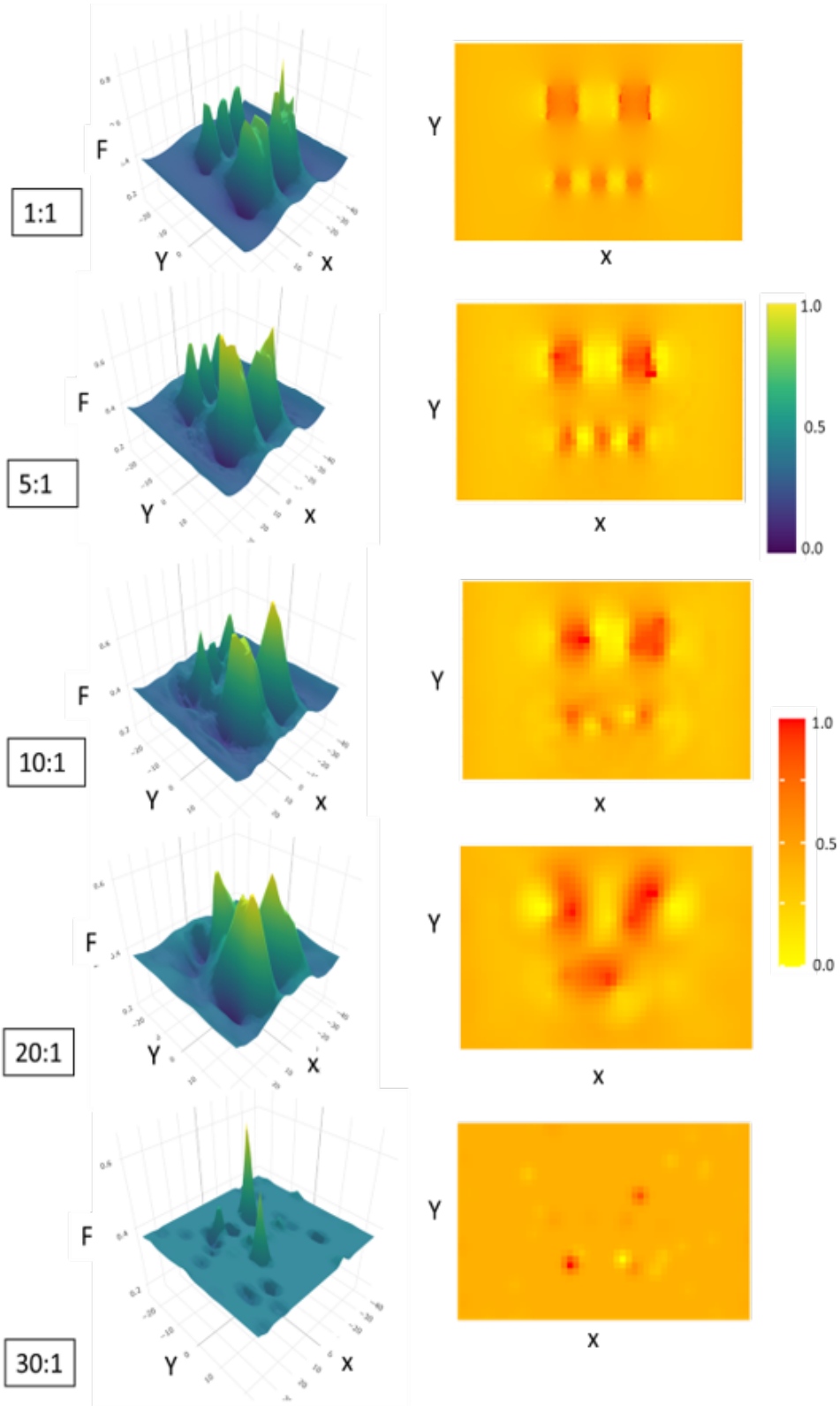

Figure 4.8 From top to bottom, as the subsampling rate (provided in boxes on left for each case) increases, we have 3D (left) and 2D (right) plots of standardized MFL readings predicted by our procedure based on the synthetic dataset described in subsection 4.2.1.

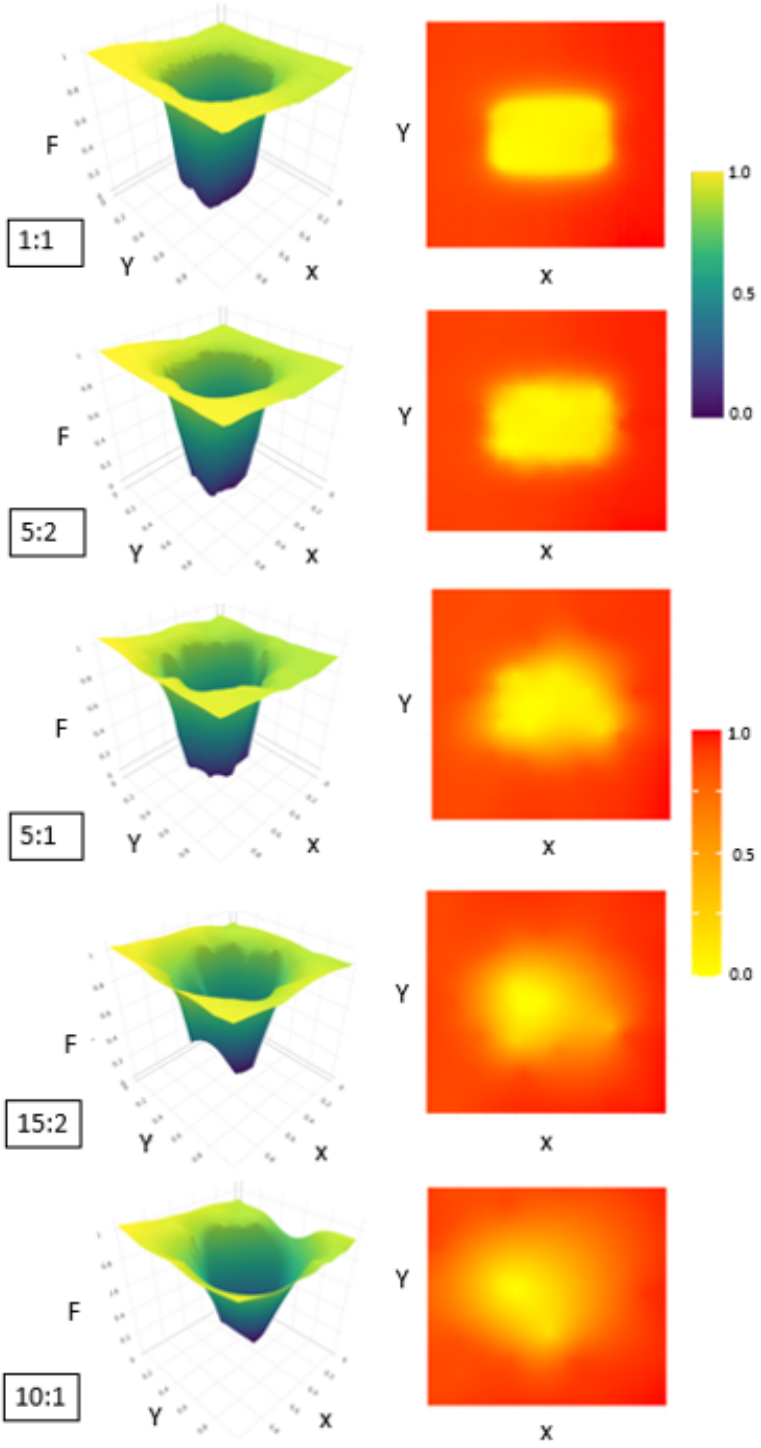

Figure 4.9 From top to bottom, as the subsampling rate (provided in boxes on left for each case) increases, we have 3D (left) and 2D (right) plots of standardized readings predicted by our procedure based on the experimental dataset described in subsection 4.2.2. The topmost figure corresponds to no subsampling and corresponds to the complete data for this example which had readings on square grid with a rectangular defect in its center.

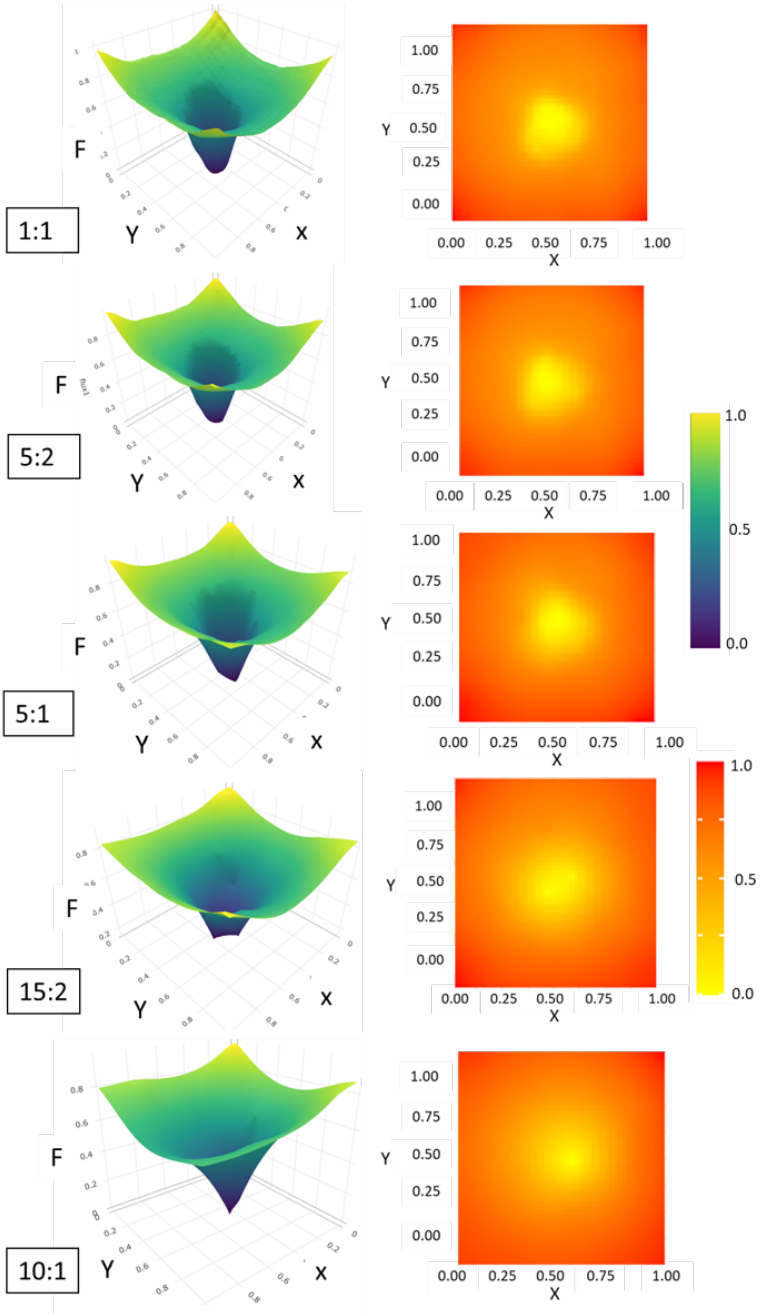

Figure 4.10 From top to bottom, as the subsampling rate (provided in boxes on left for each case) increases, we have 3D (left) and 2D (right) plots of standardized readings predicted by our procedure based on the experimental dataset described in subsection 4.2.2. The topmost figure corresponds to no subsampling and corresponds to the complete data for this example which had readings on square grid with a spherical defect in its center.

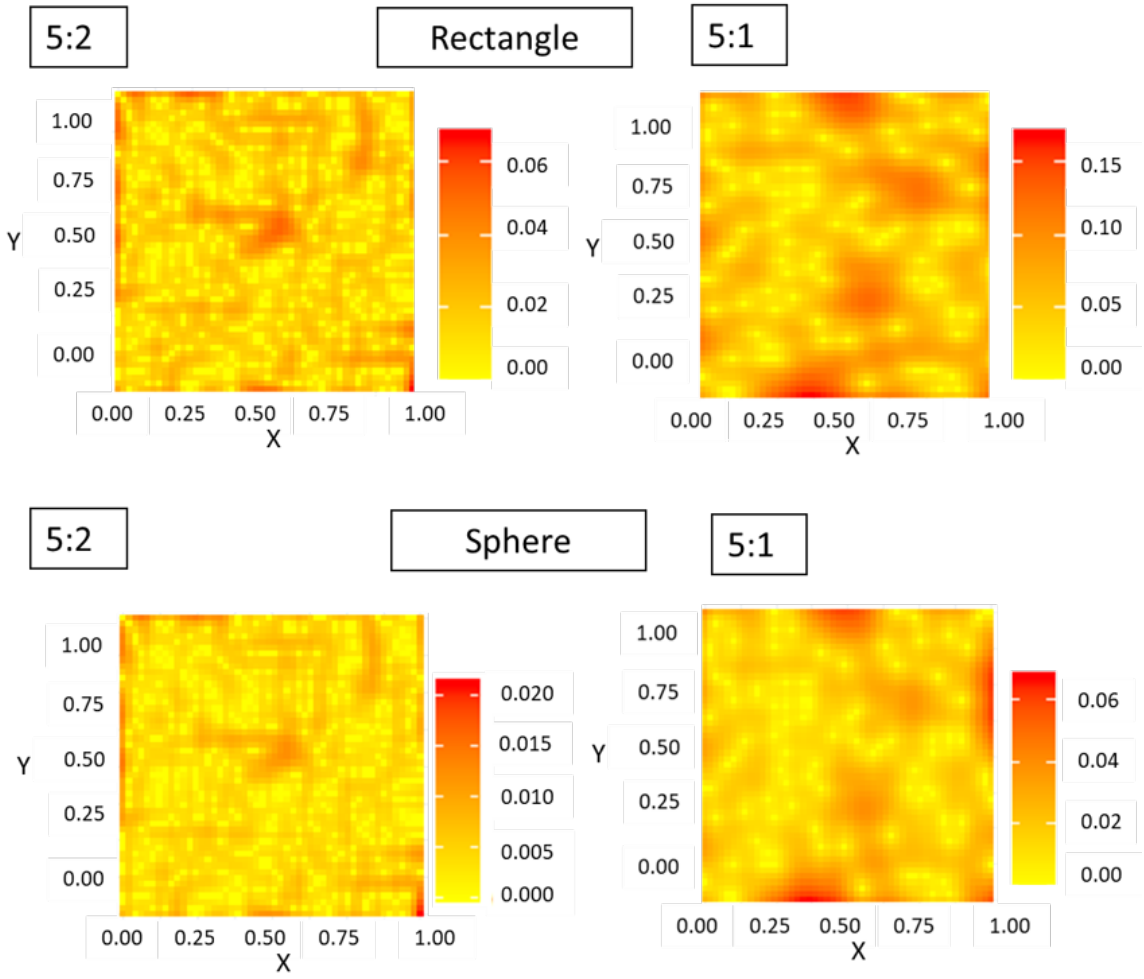

Figure 4.11 Plot of standard deviation of the kriging estimates at every scan point for experimental data (rectangle up, sphere bottom) at sub-sampling rates 2.5 (left) and 5 (right) respectively.

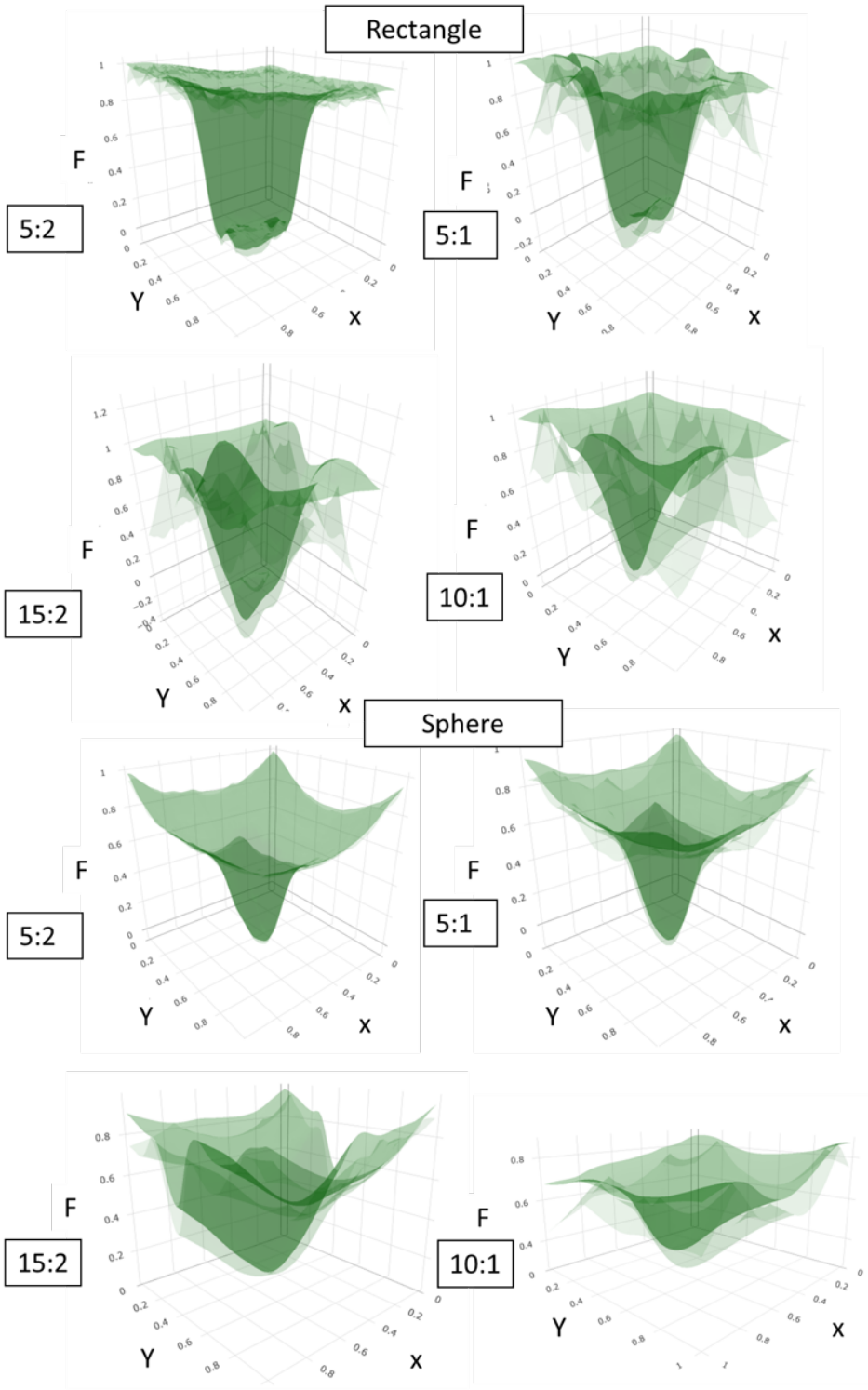

Figure 4.12 95% prediction surfaces for scan points in experimental data. The upper and lower surfaces (prediction surfaces) are in light green and their average is in dark green.

CHAPTER 5

NDE BASED DEFECT TRACKING WITH THE AID OF TRANSFER LEARNING BASED POSTERIOR PROPAGATION

5.1 Introduction

This chapter focuses on the development of a dynamically updated transfer learning method for defect detection under extreme noisy conditions. Setting up non-destructive inspection procedures abiding ideal laboratory conditions is difficult. Uncertainties due to lift-off, fluctuations in sample rate, probe tilt, difficulties in sensor registration due to fast scans result in the generation of noisy data (36; 37). Hence, better defect tracking and detection algorithms need to be developed. In this chapter we have developed a mixture regression based transfer learning framework to track defect growth on later noisy inspections. Currently, we have used our algorithm to test on the simulated MFL data inundated with very high noise . In future we will test our algorithm on the experimentally generated MFL data by our novel in house sensors as well as on the experimentally generated Eddy Current (EC) data which are noisy due to high lift off and speed.

Most of the pipelines in usage are ferromagnetic, hence Magnetic flux leakage (MFL) serves as widely accepted inspection technique for the in-service inspection of oil and gas pipelines (102; 103; 104; 105). In MFL inspection, the pipe wall which is ferromagnetic, is magnetized close to saturation by means of a permanent rare-earth magnet or a coil wound on a ferromagnetic yoke. Presence of defects increases the magnetic flux density in the vicinity of the defect, thereby increasing the reluctance and causing the flux lines to leak in to the surrounding (105). This results in a leakage of magnetic field that can be measured using a Hall effect sensor (static and dynamic fields) or coil (dynamic fields) (106; 119). Advanced adaptive MFL signal processing and characterization algorithms have been developed and demonstrated that can effectively reduce the noise, enhance the inspection data and achieve superior reconstruction performance (120; 121; 122; 123).

In the industry, inspection of pipelines is carried out during periodic maintenance cycles under the preventive maintenance framework. However, the defect that begin at the earlier stages of inspection may grow with the passage of time due to wear at a very non-linear rate. Since each

inspection is time consuming and expensive, there is need for methodologies that reduce the number of inspections. In this chapter we propose a new approach based on transfer learning techniques for processing inspection data combined with defect growth model, and along with the better understanding of the current state of piping materials to enable a condition-based maintenance (CBM). Transfer learning constitute a set of extremely powerful information assimilation techniques for retaining knowledge gained from one problem and applying it to a different but related problem (83; 84). Simply fitting function estimation cannot recognize the defects properly in noisy condition. Therefore, a dynamically updated transfer learning method is developed for defect detection under noisy condition. We consider finite element model to demonstrate such capabilities. A schematic of the MFL tool comprised of a ferromagnetic yoke and two ferromagnetic couplings [20,21] are shown in Figure 1(A). The lift-off is set as 2 mm in our model. Accurate slow scan is performed during a time when there is less or no usage of the pipeline, i.e. during regular preventive maintenance period. The location of the defect can be detected with adequate accuracy by this baseline scan which is referred as Inspection 1. The location and the size of the defect obtained from Inspection 1 (NDE signal intensity obtained by accurate scanning) are incorporated into transfer learning to recognize the defective area from later inspections in the noisy conditions.

The data for validation is obtained using a COMSOL numerical 3D model which allows systematic analysis of experimental parameters. The following sections describe details of model-based signal generation, transfer learning method and implementation results.

5.2 Methods and Materials: Experiment and Dynamic Data generation

The MFL model is treated as a magneto-static problem, which is expressed by Maxwell's equations as:

$$\nabla \times \overrightarrow{H} = \overrightarrow{J_s} \tag{5.1}$$

$$\nabla B = 0 \tag{5.2}$$

where $\overrightarrow{H}$ is the magnetic field intensity, $\overrightarrow{J_s}$ the applied source current density and $\overrightarrow{B}$ the magnetic flux density. In permanent magnetic domain, $\overrightarrow{B}$ is given by:

$$\overrightarrow{B} = \mu \overrightarrow{H} + \mu_0 \overrightarrow{M_0} \tag{5.3}$$

$$\overrightarrow{B} = \mu \overrightarrow{H} \tag{5.4}$$

In other domains, equation 5.4 holds; where μ is the magnetic permeability and $\overrightarrow{M_0}$ is the remaining intrinsic magnetization (28).

The benchmark setting as discussed in (28) is used to construct and validate our 3D model. Figure 5.1 shows the MFL 3D geometric model developed in COMSOL. The detailed description of the parameters and the material properties used to construct the model is discussed in Table 5.1. Representative 3D defects such as sphere, cuboid of various dimensions is modeled and the resulting MFL signals at the test sample surface are recorded and processed (124). Surface plot of MFL amplitude (in the form of matrices) is obtained from simulation, which indicates flux leakage due to presence of the defects. These data are used as representative MFL data pertaining to Inspection 1. As defects at later inspection times cannot decrease in size, the efficacy of the proposed algorithm could be tested against the scenario where on later inspections the defects grow in size. Here, for spherical defects, the size is increased on later inspection times. For the baseline inspection i.e., during inspection 1, a spherical defect of radius 6 mm is considered. Figure 5.2 (b)&(d) show the 3-D plots of MFL data from noiseless readings produced based on our simulation model. The figure shows that as defect grows, the perturbation in magnetic field also become larger. These data are represented in the form of 2-D images for subsequent processing and analysis (see Figure 5.2(a)&(c)).

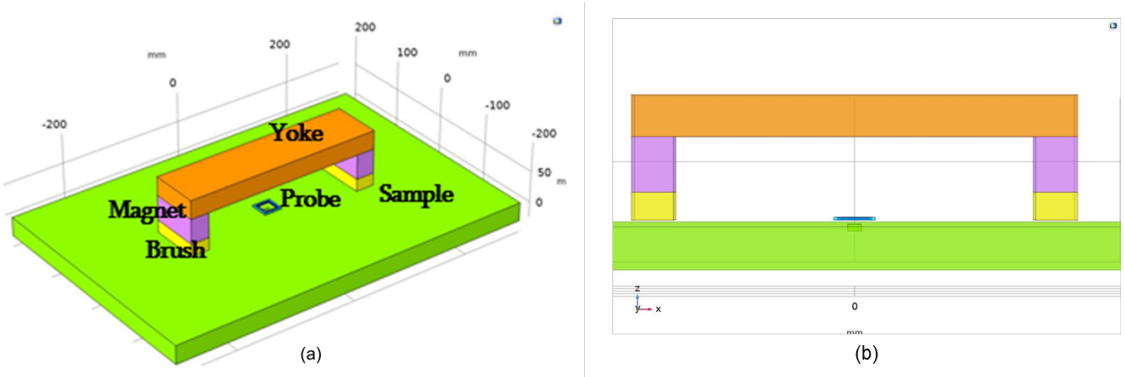

Figure 5.1 (a) shows geometry of the implemented 3D Maxwell model in COMSOL and in (b) shows the side view of the MFL model.

Table 5.1 Various dimensions of the MFL model(l=length, w=width, h=height, all the dimensions reported in mm).

Term	Length(l)	Width(w)	Height(h)	Material
Yoke	320	80	30	Ferro-nickel alloy
Magnet	30	80	40	NdFeB
Brush	30	80	20	Ferro-nickel alloy
Specimen	600	400	10	X52 similar to iron

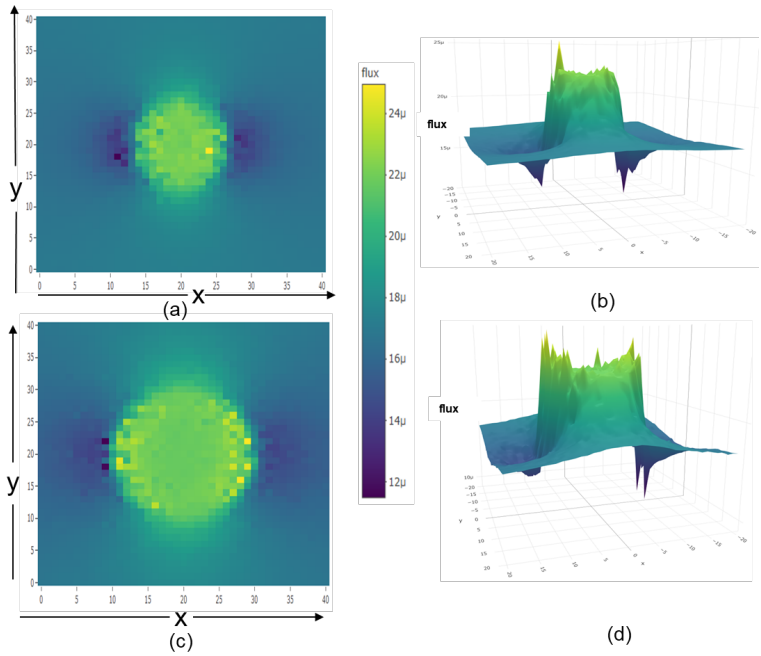

Figure 5.2 (a) and (b) respectively show the 2D-plot and 3D MFL data from noiseless inspection 1 readings for a spherical defect of radius 6 mm as locations vary in the x and y axes. (c) and (d) show these plots for a spherical defect of radius 9 mm without noise. These data were generated by our design using COMSOL.

In order to generate noisy MFL data for the successive inspections, a two-fold approach is taken. Flux readings corresponding to increased defect sizes for each inspection using the previous designed approach are obtained. To produce representative noisy data for the later inspection times, these MFL readings are contaminated with Gaussian noise having a fraction of the signal standard deviations at $t_1 (t_1 \in [0, \text{inf}])$ and at $t_2 (t_2 \in [0, 1])$ of randomly chosen scan points. If the MFL readings were sequenced at a uniform spacing of S_p mm then the signal-to-noise ratio (SNR) is defined as:

$$SNR = \frac{1}{t_1^2 . t_2 . S_p} \tag{5.5}$$

Figure 5.3(a,b,c) shows representative data for three inspection times where the defect is spherical in shape and has grown from 6 mm to 9 mm in size over the span of inspection times with an increment of 1 mm on each passing inspection. The data were collected on a square grid of length equals to 2 cm, and the MFL data readings were made on a lattice with spacing equals to 1 mm, so we had 41^2 scan points at which the readings were made. This data has SNR equals to 1 and are analyzed in Row 3 of table 5.2. Datasets gathered by cheap fast scans have low SNR. Increasing SNR means that the wavelet-based algorithms can detect the defects accurately with ease. The SNR was considered as 1 here, which is an interesting regime to study where wavelet fails and transfer learning algorithm can provide immense benefits. Be noted that inspection 1 data are much less noisy than data from the subsequent inspection times.

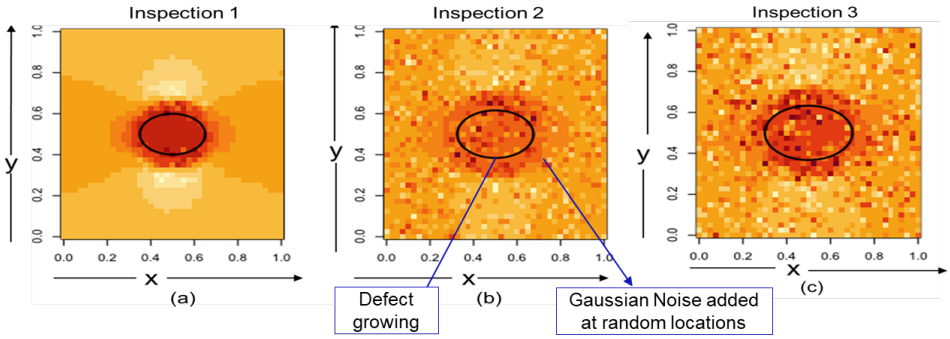

Figure 5.3 a, b, c shows MFL data for inspection times 1 to 3 for a spherical defect which had radius 6 mm, 7 mm and 8 mm on the successive inspections. The readings are collected on a square lattice of length 2 cm and at uniform spacing of 1 mm. The SNR for Day 2 and 3 is fixed at 1 with t_1 and t_2 values given in table 5.2, row 3.

5.3 Algorithms: Tracking Defective Area

5.3.1 Algorithm 1: Defect detection on baseline inspection by mixture regression analysis of low noise data

In this work, mixture regression method is used (125) for fitting a bivariate function to the less noisy MFL data obtained from the baseline inspection. Mixture regression is a potent methodology for identifying the presence of sub-populations within an overall population (126; 127). Location of the scan points is used as explanatory variables based on which the model is trained to provide a probabilistic classification of scan points into two segments: defects and non-defects. For setting up the regression model, vectorization of the matrix-variate MFL data is done at first. Consider the

response vector Y of length N^2 as:

$$Y_l = f(x_i, y_j), \text{ where, } i, j = 1, \cdots, N$$
$$\& l = N(i-1) + j \tag{5.6}$$

where, x_i, y_j denotes the coordinates of the point for the daily probes and N is the number of points at which the MFL data are sampled for both axes. The unknown intensity generating function f is modelled as:

$$f(x_i, y_j) = \sum_{k \in D,S}^{2} P_k \cdot N(\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j, \phi_k^2) \tag{5.7}$$

The number of mixtures is 2; D denotes the class of defective points on the grid and S is the set of non-defects or safe scan points. For each mixture model, the intensity is modeled by a linear surface $\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ with aberrations having variance ϕ_k^2. Here $\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ are the regression coefficients and ϕ_k^2 the regression error for $k = D, S$ and P_k represents the mixing weights for each of the k regression models. If ϕ_k^2 is low, then there is not much variability, which will be the case when there are no defects in the vicinity. For areas near defects, the average intensity $\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ will be different leading to the points in those areas having other mixing density [23]. Thus, spatial relation among the points are leveraged here. Also, for point adjoining defects the variance ϕ_k^2 will be large. While computing, for each point (x_i, y_j), the posterior probabilities of $\hat{\pi_D}(x_i, y_j)$ are evaluated and based on their value, we can classify each point to be either defective (D) or non-defective (S) area. Note that, if the defective area is not very large, then the sector does not contain a harmful defect. Different kinds of defects such as scratches, small, moderate and large defect sizes are considered. In all these cases, it is observed that function estimation matches well with the intensity plots; the defective areas on Inspection 1 can be appropriately recognized by applying the aforementioned function estimation procedure on data, which has zero or very low noise levels. The R package mix tools is adopted (128) for fitting these mixture regression functions. Figures 5.4 show the application of the procedure on Inspection 1's data of Figure 5.3 for the spherical defects.

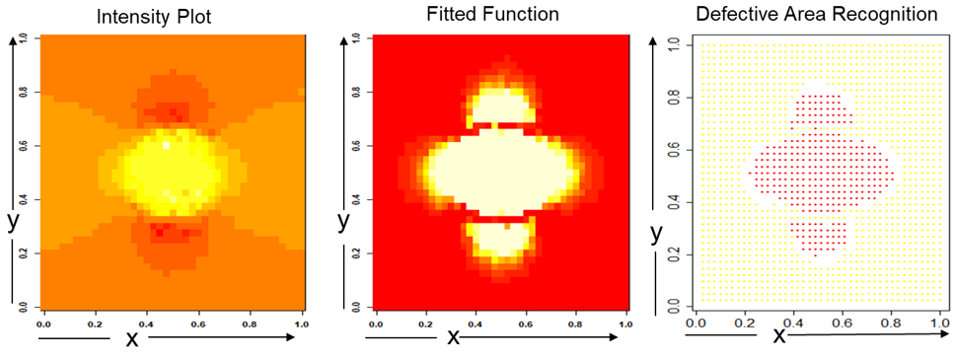

Figure 5.4 (a) shows the 2D MFL data from a spherical defect of radius 6 mm on Inspection 1; (b) shows the fitted function by the 2-mixture regression model and (c) shows the estimated defective scan points plotted in red as based on thresholding the posterior probabilities at 0.8.

5.3.2 Algorithm 2: Defect growth tracking after later cheap scans using transfer learning

The two-mixture regression algorithm fails when the data are noisy. Figure 5.5 shows how the mixture regression approach cannot be used even on wavelet denoised MFL readings from Inspection 2 of Figure 5.3. Its performance on raw MFL Inspection 2 data is even worse. Debauches wavelet of level 3 for denoising the noisy MFL data has been applied. Panel 5D of figure 5.5 shows how the two-mixture model has failed to detect the true defect area on the following days. As a result, the above algorithm cannot be used to analyze data from the noisy scans after Inspection 1. However, if the algorithm is made aware of the location of defects as well as the mixing proportions, then the algorithm can perform better [28] and the proposed method is described in detail below.

Let I_s be the intensity matrices from the daily probe scans for $s = 1, 2, \cdots, I_1$ is the intensity matrix from the Inspection 1 and is quite accurate with very low noise compared to $I_s : s > 1$. For simplicity, consider that new defects have not cropped up in between, and the concern is to update the sizes of the existing defects based on I_1 as the beginning of the inspection. Based on Inspection 1's data, the location and proportion of the defective points sampled in the probe can be estimated very well. By updating the probabilities of the defects $\hat{\pi}_D^{(s)}(x_i, y_j)$ based on I_s for every s, let D_1 denote the location of defect scan points in Inspection 1 for any prefixed threshold τ:

$$D_1 = (x_i, y_j) : \hat{\pi}_D(x_i, y_j) \geq \tau \tag{5.8}$$

Consider $Z_1 = I_1(x_i, y_j) : (x_i, y_j) \in D_1$ as the set of Inspection 1 MFL values on the defect scan

points and $Z_2 = I_2(x_i, y_j) : (x_i, y_j) \in D_1$ be the set of Inspection 2 MFL values on the same scan points, which were judged in the defective area D_1 of Inspection 1. Be noted that Z_2 is very noisy, hence it is important to denoise Z_2. The denoising can be done by fitting a linear model on Z_2 based on Z_1 values as follows:

$$Z_{2k} = \alpha_2 + \beta_2 Z_{1k} + \sigma_2 \epsilon_k, \text{where} k = 1, \cdots, |D_1| \tag{5.9}$$

Here $\alpha_2, \beta_2, \sigma_2$ are intercept, slope and standard deviation parameters and $\epsilon_k : k \geq 1$ are independent and identically distributed (i.i.d.) Gaussian noise. After obtaining the ordinary least square coefficients $\hat{\alpha}_2, \hat{\beta}_2, \hat{\sigma}_2$ the set of predicted value can be evaluated as:

$$Z_2^P = \hat{\alpha}_2 + \hat{\beta}_2 Z_1(x_i, y_j) : (x_i, y_j) \in D_1 \tag{5.10}$$

Z_2^P will be much less noisy than Z_2. In Inspection 1, Z_1 was well modeled by a linear model and so Z_2^P which is the denoised version of Z_2, got by projecting it on Z_1 values, to be also well-modelled by linear models. The neighborhood of the defect points of Inspection 1 are then considered. Let N_1 denote the neighborhood of the scan points in D_1, i.e.,

$$N_1 = (x_i, y_j) : \min_{(x,y)\in D_1} |x_i - x| + |y_i - y| < \rho \setminus D_1 \tag{5.11}$$

where, ρ is a tuning parameter. In these neighborhood points N_1, the defect can spread on Inspection 2, which is the uncertain area. All points outside N_1 and D_1 can be considered to be safe on Day 2. Let S_2 denote that set, i.e., $S_2 = (N_1 \cup D_1)^c$. Let $W_1 = I_1(x_i, y_j) : (x_i, y_j) \in S_2$, $W_1 = I_2(x_i, y_j) : (x_i, y_j) \in S_2$ be the MFL data from Inspection 1 and 2 respectively on these safe points. Again, as W_2 is noisy hence we fit a linear model W_2k as,

$$W_{2k} = \delta_2 + \gamma_2 W_{1k} + \tilde{\sigma}_2 \tilde{\epsilon}_k \text{ where, } k = 1, \cdots, |S_2| \tag{5.12}$$

and similarly, the predicted values based on the least square coefficients W_2^P can be evaluated as:

$$W_2^P = \hat{\delta_2} + \hat{\gamma_2} W_1(x_i, y_j) : (x_i, y_j) \in S_2 \tag{5.13}$$

Similarly, as before $\delta_2, \gamma_2, \tilde{\sigma}_2$ are intercept, slope and standard deviation parameters and $\tilde{\epsilon}_k : k1$ are independent and identically distributed (i.i.d.) Gaussian noise. The least square coefficients

$\hat{\delta}_2, \hat{\gamma}_2$ are then obtained to consider the set of predicted value W_2^P. By projecting on Inspection 1 data, we have de-noised and linearized Z_2^P and W_2^P. Thus, the MFL values in Inspection 2 for all scan points other than the neighborhood points are well regularized by this transfer learning approach. Let U_2 denote the set of Inspection 2 MFL values in the neighborhood points N_1 i.e. $U_2 = I_2(x_i, y_j) : (x_i, y_j) \in N_1$. In the next step, a mixture regression model with two groups on Z_2^P, W_2^P, U_2 response values is developed. As the proportion of points in $(S_2 \cup D_1)$ is much higher than N_1 points, there are many more denoised points in Z_2^P, W_2^P, U_2 than noisy points. So, this mixture regression model can be fitted accurately unlike before. Using the model which is basically trained by the points in Z_2^P and W_2^P, the posterior of the points in N_2 are detected and thus classified as safe or defect. Thus, the D_2 defect area is:

$$D_2 = D_1 \cup (x_i, y_j) \in N_1 : \hat{\phi}_D^{(2)}(x_i, y_j) \geq \tau \tag{5.14}$$

Similarly, for Inspection 3, the neighborhood N_2 of the defective area D_2 of Inspection 2 is considered. The safe points of Inspection 3 are denoted as$S_3 = (N_2 \cup D_2)^c$. Corresponding to these S_3 and D_2 points, the predicted intensities W_2^P and Z_2^P are evaluated based on the above mixture regression fit on Z_2^P, W_2^P, U_2 values of Inspection 2. These intensities are thus used to transfer information from Inspection 2 defect area to Inspection 3.The Inspection 3 intensity Z_3, W_3 corresponding to D_2 and S_3 scan points are denoised by regression on Z_2P and W_2^P respectively. The resultant predictors are Z_3^P and W_3^P. Again, mixture model is fitted on Z_3^P, W_3^P, U_3 values and based on the posterior probability of defect in N_3 the neighborhood points are updated. Thereafter the scan points are classified as defect or non-defect by thresholding these posterior probabilities. The process is repeated for the following days.

5.4 Results & Discussions

In this paper, the efficacy of our algorithms is tested on spherical shaped defects across a wide range of SNRs and sampling designs. Firstly, a sampling design where MFL data are collected on a square lattice of length 2 cm with scan points at 1 mm apart along the either axis is considered. The defect is spherical with radius 6 mm on Inspection 1. Its radius increases by 1 mm on each passing inspection. Three different sets of values of t_1 and t_2 for this defect type are considered

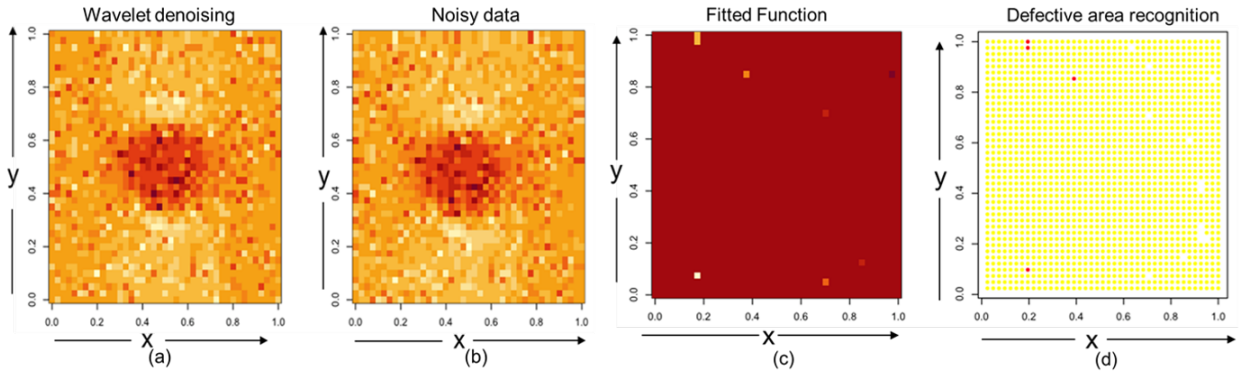

Figure 5.5 (a) Inspection 2 MFL data from Figure 5.3. It is based on a lattice of spacing 1 mm on a pipe sector of length 2 cm. It contains a spherical defect of radius 6 mm. The SNR=1 with t_1 and t_2 values given in the Table 5.2, row 3. (b) Wavelet based denoised version of the data. (c) shows the 2-mixture regression based predicted function and (d) shows the estimated defective scan points plotted in red as based on thresholding the posterior probabilities at 0.8.

based on which noise-contaminated Inspection 2 and Inspection 3's data was generated on the basis of the design as described in Section 5.2. Row 1-3 of Table 5.2 contains the result pertaining to this spherical defect type. The top row of Figure 5.3 demonstrates the corresponding MFL data for a particular set of t_1 and t_2 values (row 3 of Table 5.2) that makes SNR equals to 1. In all the cases, the mixture regression approach works well on Inspection 1 data as illustrated in Figure 5.4 and recovers Inspection 1's defect location and volume precisely. However, naïve application of the mixture regression approach on Inspection 2's data performs terribly in recovering the defective area. The performance of our proposed Transfer Learning (TL) based mixture regression algorithm and that of the naïve method (Non-TL) which does not transfer information from Inspection 1 is thoroughly compared, and a mixture regression model on later Inspections 2 and 3 in isolation is fitted and developed. For evaluating the performance of any NDE based detection methodology (M) on these kind of data types, it is suggested to consider the following two sets:

D_s^M =set of defective scan points detected by method M on inspection s.

D_s^T =set of true defective scan points on inspection s.

We can define the coverage of defect by method M on inspection s as:

$$\text{Coverage}_s = \frac{size(D_s^M \cap D_s^T)}{\text{size}(D_s^T)} \tag{5.15}$$

To check if method M has grossly overestimated the size of the defect, we further define:

$$\text{Exceedance}_s = \frac{\text{size}(D_s^M \cap D_s^T)}{\text{size(G)}} \tag{5.16}$$

where, G is the grid of sampling points for MFL data. Methods with high coverage and low exceedance values are desired. From Table 5.2, it is concluded that our proposed method has near perfect coverage and very low exceedance for all the different regimes considered here. Its performance drops as the SNR is decreased from 2 to 1, however, the drop is not by much and its performance is still commendable. In contrast, the naïve Non-TL method that do not use the dynamic nature of the problem has poor performance. It is found that the TL based approach has full coverage of the defect area on all 3 inspection times across all regimes along with acceptable exceedance (the maximum exceedance is 7%). In contrast, the naïve algorithm that models each inspection separately without transferring previous inspections' information cannot attain good coverage on inspections 2 and 3. Figure 5.6 shows the defect growth tracked by our algorithm on Inspection 3 where the SNR was used by as Design 3. In Figure 5.7, we report the root mean square error (RMSE) for the predicted signal intensity by our proposed approach at the detected defective scan points. We report the RMSE relative to the inherent variation of the signal intensity at the defective points and it was witnessed to be well controlled.

Table 5.2 Performance based on mixture regression and transfer learning.

Shape	Design	SNR	Inspection Time	Evaluation Metrics	Proposed TL Method	Non-TL Method
Spherical defect increasing 1mm/inspection	$L = 2cm, t_1 = 1, t_2 = 0.5, S_p = 1mm$	2	2	Coverage of defect	100.00%	3.13%
				Exceedance	0.91%	0.23%
			3	Coverage of defect	100.00%	31.25%
				Exceedance	4.76%	1.13%
Spherical defect increasing 1mm/inspection	$L = 2cm, t_1 = 1, t_2 = 1, S_p = 1mm$	1	2	Coverage of defect	97.55%	11.35%
				Exceedance	2.20%	0.00%
			3	Coverage of defect	96.23%	16.23%
				Exceedance	2.62%	0.30%
Spherical defect increasing 1mm/inspection	$L = 2cm, t_1 = 1.41, t_2 = 0.5, S_p = 1mm$	1	2	Coverage of defect	94.79%	0.00%
				Exceedance	0.59%	0.00%
			3	Coverage of defect	94.67%	34.33%
				Exceedance	4.16%	1.07%

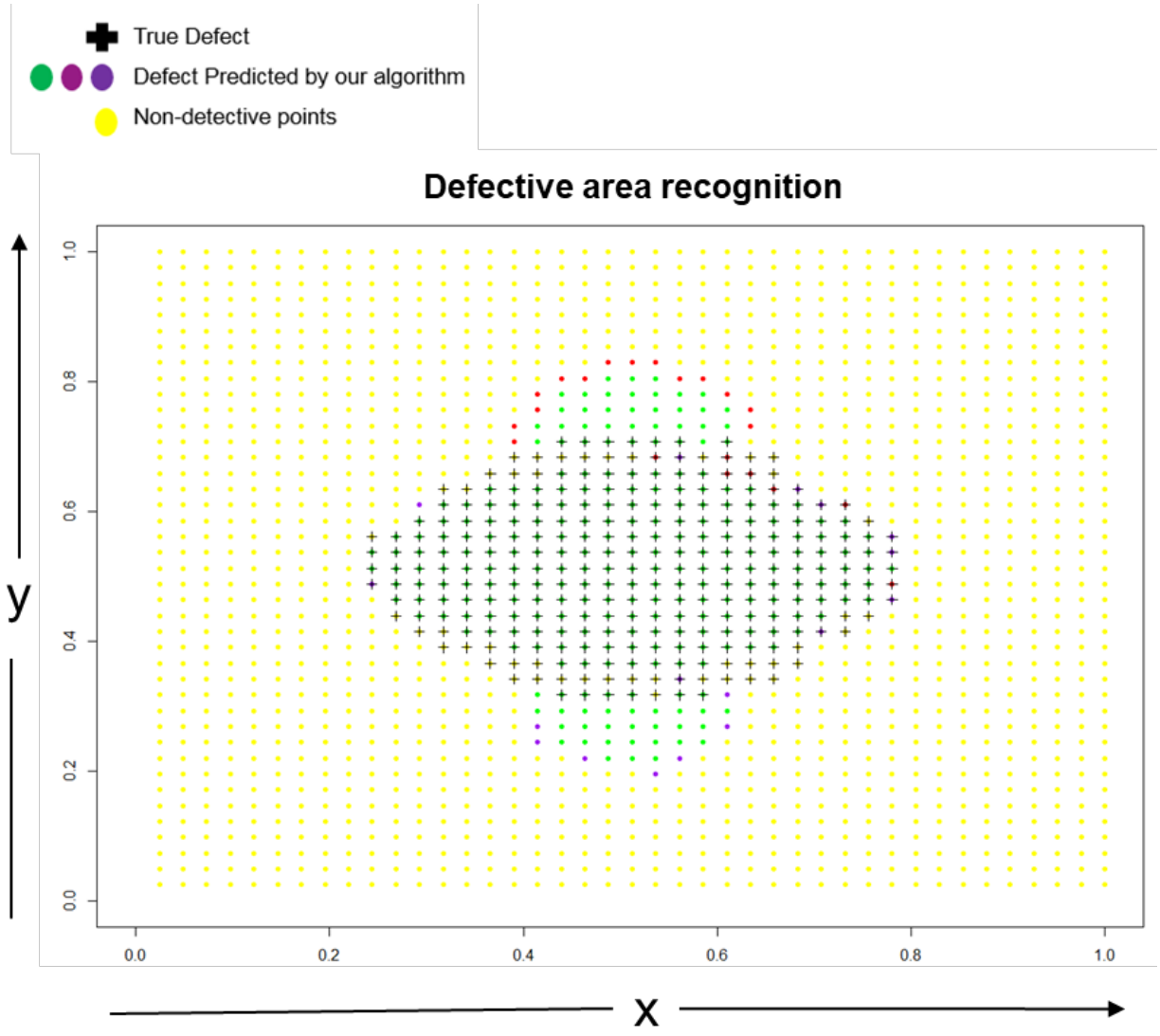

Figure 5.6 Defect tracked by the algorithm for Inspection 3 and by design 3.

5.5 Conclusion

The defect size cannot be properly estimated with limited data size and noisy measurement. A novel framework by applying bivariate function based estimation to the noisy data obtained from the later inspections using fast and cheap scanning has been investigated and demonstrated. In order to properly recognize the defect under noisy condition, the algorithms can dynamically update the grid by the auxiliary information obtained from baseline inspection i.e., Inspection 1, where the location and the size of the defect is used as a transfer learning to update the size of the defect. As the defect on the later inspection times can't decrease in size, the efficacy of our algorithms are

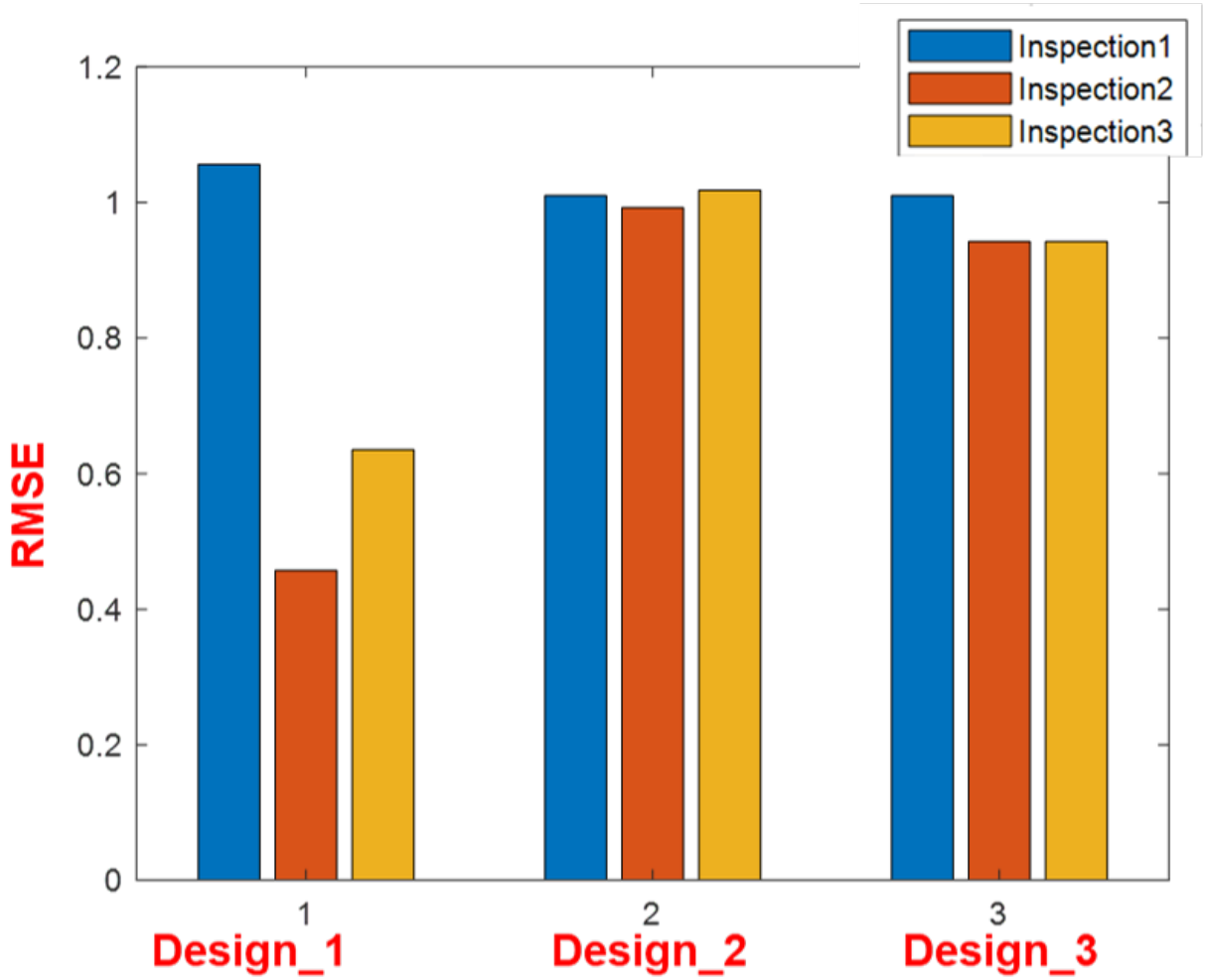

Figure 5.7 RMSE of predicted signal intensity at the defective scans points for the different designs of Table 5.2.

rigorously tested and satisfactory results have been achieved.

Portions of this chapter appeared in the publication (129; 29)

CHAPTER 6

DYNAMIC DEFECT DETECTION IN FAST, ROBUST NDE METHODS BY TRANSFER LEARNING BASED OPTIMALLY BINNED HYPOTHESIS TESTS

6.1 Introduction

Non-destructive evaluation (NDE) plays an important role in condition-based monitoring (CBM) and performing effective diagnostics of the safety-critical infrastructures such as oil and gas transmission pipelines, aircrafts, railways, bridges, nuclear power plants, battery inspection in electrical vehicles, load bearing metal structures in highways etc. (61; 62; 130). To diagnose structures effectively, NDE techniques utilize various inline inspection (ILI) techniques to access them without causing additional harm, in a non-invasive manner (26). With the advancement in sensing technology, different NDE modalities (27; 28; 29; 30; 31; 12; 52; 11; 32; 33; 10; 34; 9; 35; 8) are being rapidly used. However, in real situations, it is often very challenging and expensive to establish non-destructive inspection procedures that follow the optimal less noisy laboratory conditions in which these methods were created and tested. For this reason, NDE data are collected under uncertainties such as high scanning speed, lift-off, probe drift, fluctuations in material surface and different data acquisition step sizes (12; 36; 131; 116; 132). Conducting fast inspections makes the signal distorted and noisy. Lift-off refers to the distance between the sensor and the material under test (MUT) while probe drift denotes the deviation of probe from perpendicularity. Thus, NDE data are most often distorted and contaminated by noise due to various scanning speeds and the presence of unknow lift offs due to coarseness of MUT surface, non-conducting coatings, or irregular paintings on MUT (38; 7). Variation in data acquisition step sizes, misalignment of the probe and uneven sample surface also contaminate NDE data. These uncertainties in data collection process result in different types of signal corruption. Defect detection algorithms developed based on sensor data under controlled laboratory conditions do not provide accurate defect analysis under these uncertainties. In present NDE literatures, very few adaptive algorithms are present that perform accurate defect diagnostics under noisy regimes (6; 12; 29; 28). However, to the

best of authors' knowledge there is no such literature that can perform dynamic defect tracking on inspections corrupted due to the stated uncertainties.

Here, a miniaturized in-house MFL probe which is cheap, highly flexible, and easy to implement is designed. It does not require cumbersome hardware setup needed in conventional MFL Pipeline inspection gauge (PIG) (51). The commonly used non-destructive evaluation approaches includes ultrasonic testing (UT), Eddy Current (EC), pulsed eddy current (PEC), microwave testing, Electromagnetic acoustic transducer (EMAT) testing, capacitive sensing, thermography and magnetic flux leakage-based inspection. Among all these techniques, MFL is widely used in inspection of the conducting materials, mostly ferromagnetic oil and gas pipe wall inspections (102; 103; 104). MFL follows a simple working principle as follows: ferromagnetic pipe wall is magnetized close to saturation by means of permanent rare earth magnet or a coil wound on a ferromagnetic yoke. The presence of defects decreases the pipe wall thickness, thereby increasing the reluctance and magnetic flux density in the vicinity of defects. This phenomenon results in a higher fraction of magnetic flux to leak into the surrounding, which can be measured using a Hall effect sensor (static and dynamic fields) or coil (dynamic fields) (105; 106; 119). Unlike other popular NDE methods, it can be employed over a wide range of conducting specimens as it enjoys the following advantageous properties over the competing NDE methods: (a) Thermographic measurements have low sensitivities while using cheap thermal cameras for imaging and are non-robust to small temperature variations (4; 1). X-ray imaging is expensive as well as proper screening is needed to take care of detrimental ionizing radiations (33); (b) laser profilometry and structure light based optical inspections do not work on ferromagnetic conducting pipe walls as it results in unwanted reflections (99); (c) Ultrasonic testing requires a coupling medium and physical contact with MUT (32). Air coupled UT does not need coupling but requires long sound pulse for excitation where accurate timing measurement is difficult. Moreover, because of high impedance of metal structures it is not applicable for ferromagnetic metallic pipe wall inspection (5); (d) Microwave imaging operates at high frequencies in GHz ranges which require extensive data acquisition circuitry and fail to penetrate deep in conducting materials (2); (e) In eddy current inspection, the optimized

frequency needs to be chosen to take care of skin depth (31) and capacitive sensing exhibits non linearity as the output voltage is inversely proportional to the distance between the sensing and driving electrodes (12).

6.1.1 Dynamic defect tracking in streaming NDE data

A set-up for periodic monitoring of MUT by NDE methods at regular intervals is considered. Thus, the NDE data arrives in as a streaming data sequence and the goal is to carefully estimate the changes in the state-of-health of the MUT between two successive scans. In the real-world, in most industrial sectors, inspections are carried out during periodic maintenance cycles under preventive maintenance framework. Corrosions and other defects in the metallic surfaces on the MUT are dynamically tracked. The defective areas in the MUT such as corrosions in the inner wall of pipelines often remain controlled over time but can also further aggravate by increasing in size in which case alarms need to be raised for human intervention, particularly when the defective areas are large enough to be immensely detrimental to the safe functioning of the MUT. Also, new corrosions and defects can also crop up over time which need to be detected and tracked. Since the less noisy inspections are time consuming and expensive, there is need for methodologies that reduce the scanning times. Here, a new approach that can significantly reduce the NDE data collection cost and inspection time for this dynamic defect monitoring problem is addressed. A dynamic setup is considered where there is an occasional detailed, highly accurate, expensive MFL scan in between regular low cost noisy MFL scans. The overall cost of the proposed NDE set-up is only a small fraction of defect tracking procedures that are based on NDE data from frequent less noisy but quite costly MFL scans. In all dynamic NDE systems huge volume of data needs to be regularly analyzed over time which need development of automated data analysis algorithms that can accurately detect formation of defects in metallic pipes and precisely track their growth over time. While our proposed set-up is cheap, it is challenging to develop automated data analysis algorithms for defect tracking based on noisy MFL scans. For this purpose, we develop an efficient transfer learning based (133) defect tracking algorithm.

6.1.2 Transfer learning based binned hypothesis tests (TLBH) for defect detection

In expensive, less noisy MFL scans, the signal-to-noise ratio (SNR) for defect detection is very large and usually pointwise hypothesis testing (134; 135) method that classify each scan point as defect or non-defect typically works well. However, with increase in noise in NDE data, pointwise hypothesis testing fails to work (12) due to low SNR in the data. Binned or aggregated hypothesis testing methods (136; 137; 138; 139; 140) are popularly used in low SNR cases. In a host of modern applications (141; 142; 143) binned tests that aggregate neighbourhood information is shown to be very successful for signal detection in very low SNR regimes. Aggregation helps in increasing SNR of the aggregated signals and consequently defects can be defected. However, aggregation leads to loss of resolution of the data and too much aggregation can lead to massive overestimation of the defect size resulting in very high FDR. It is extremely important to optimally select the aggregation level in binned test so that the test is not only powerful for defect location detection but also does the least over estimation of defect sizes. The optimal aggregation-level for binned test can be easily constructed if we know the true SNR of the noisy MFL signals. Since, these are unknown we develop a transfer learning based technique for estimating the SNF in nosy MFL signals. Transfer learning is a set of powerful information assimilation techniques for retaining knowledge gained from one problem and applying it to a related problem (84; 133). Leveraging the streaming MFL inspection data set-up, we transfer accurate knowledge of locations and sizes of existing corrosions from the detailed MFL scan to the subsequent noisy MFL inspections. Using this support information of defective scan-points we estimate the deterioration of SNR in the noisy MFL scans. It is shown that the proposed binned hypothesis test has optimal defect coverage and false discovery rate (FDR) control (see Section V). A fascinating property of the proposed optimal binning hypothesis testing based defect detection method is that it is agnostic of the noise type in the low-cost MFL scans. So, the method can be applied even if the underlying noise producing uncertainties massively changes over time.

6.1.3 Organization of the chapter

In Section 6.2, description of experimental MFL data generation under different types of uncertainties for the proposed cost-effective dynamic defect tracking set-up is given. Section 6.3 illustrates the highly heterogenous nature of noise in these MFL datasets. Under different noisy regimes, the deterioration of SNR and the subsequent problems faced by point-wise hypothesis testing methods are shown. In Section 6.4, detailed description of the transfer learning based binned hypothesis testing (TLBH) method is provided and its properties are discussed. In Section 6.5, it is demonstrated that the proposed TLBH method can track different types of defects with high precision. The performance of TLBH with the following competing methods (a) point-wise hypothesis testing methods (b) naïve binned hypothesis testing methods (c) Wavelet based spatial denoising and thresholding (d) dynamic mixture regression models for defect detection is compared in this section. In Section 6.6, a brief discussion about the eddy current based inhouse experimental setup and the applicability of our TLBH algorithm on the EC data are provided. It is shown that TLBH far out-perform the competing methods in dynamic growth tracking for our streamed MFL tracking set-up. The chapter concludes with a discussion in Section 6.7.

6.2 MFL Experiment and Data Generation

A MFL probe based on the magneto-static problem governed by conventional Maxwell's equation (29) is developed. The experimental setup is designed using the components as discussed in chapter 3 of this thesis. Figure 6.1 shows the schematic of the entire experimental setup. The holder of the MFL probe fits two cubical permanent magnets composed of NdFeB material. Two different holders are constructed to hold magnets of two different dimensions. Figure 6.1(b) shows one such holder of dimension $80 \times 72 \times 30$ mm while another larger holder is shown in Figure 6.1(e). Two cubical permanent magnets of dimension 25 mm are placed inside the holder with opposite polarities facing the MUT. Large cuboidal permanent magnets of dimension $40mm \times 4mm \times 20mm$ are also used. Hall sensor of dimension $25mm \times 14mm$ is placed in the middle of the holder between magnets of opposite polarities. We use Aerotech AGS1000 programmable XYZ scanner for scanning and TLE493D-A2B6 three axis-Hall effect sensor produced by Infineon to measure

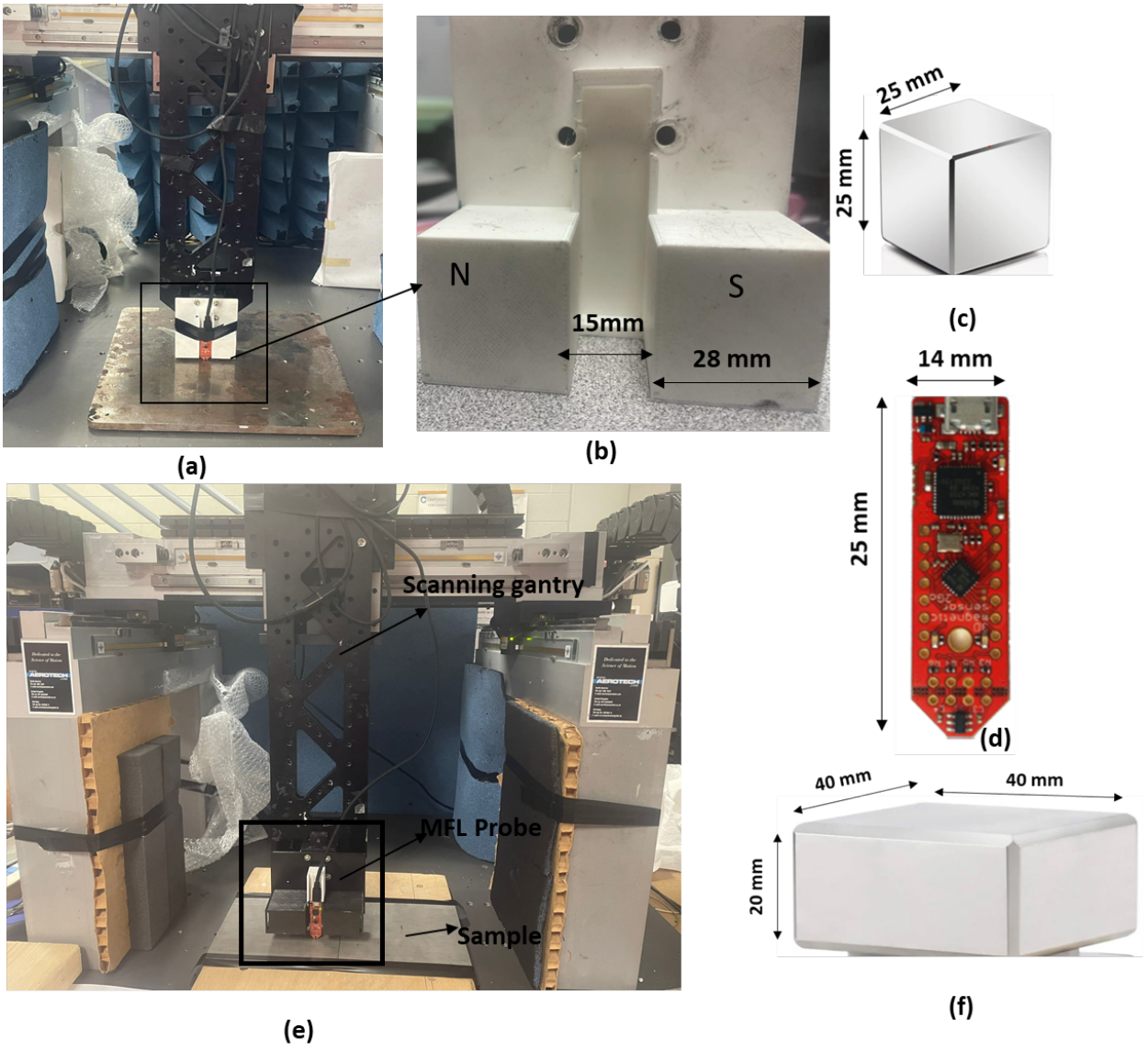

Figure 6.1 Panel A shows the schematic of the developed MFL probe scanning MUT, (b) Dimensions of the probe holder, (c) Permanent magnet used in sensing, (d) Infenion TLE493D-A2B6 sensor used in sensing, (e) Entire setup with gantry, (f) Dimension of larger permanent magnets.

the axial, radial and tangential components (in X, Y, Z directions) of the magnetic field leakages. TLE sensor consists of three main blocks (a) Power mode control system comprising of oscillators, (b) Main sensing unit consisting of Hall probes, multiplexer and ADC controller, (c) I2C interface, register files. The sensing unit of the sensor is connected sequentially to a multiplexer and then to an Analog to Digital Converter (ADC) which is connected to Data acquisition system (DAS). The DAS consists of a National Instrument Data Acquisition Card PCIe-6341 where the magnetic flux signal is sampled and digitized by a routine and the output is recorded.

To understand how the data collected by MFL sensors is altered as uncertainties increases in system, we collect data across different regimes where the defect has increased in size during later scanning. Our designed MFL probe recorded flux readings at scan points of varying step sizes of 0.5mm and 2mm respectively. A square grid of 100×100 mm is scanned for this purpose. With 0.5mm step size, it results in 40,000 grid points whereas for 2mm step size, only 2500 scan points are recorded. The scan time for 0.5mm step size is about 1 hour 30 minutes whereas the scan time for 2mm step size is only about 18 minutes. The efficacy of our algorithms in dynamic defect tracking on later noisy inspections are tested across following three defect cases:

- On later scans defect size can remain constant
- On later scans defect size get increased
- On later scans new defect has cropped in

In Figure 6.2 , the axial B_x, tangential B_y, radial B_z and magnitude $|B| = \sqrt{(B_x^2 + B_y^2 + B_z^2)}$ components of magnetic flux signal sensed by our MFL sensor on one such experimental sample containing notch like fabricated defect is shown. The scanning is done at lift off of 1 mm and step size of 0.5 mm. The MFL sensor is capable of detecting both the large and small defects under less uncertainty due to small lift off. Our algorithms are tested on the axial component B_x of the flux signals in the subsequent sections. Figure 6.3 shows the axial component of the magnetic flux readings across four Inspection cases. Here, Sample 1, 2, 3 and 4 (sample dimensions stated in later section) are scanned at increasing lift offs of 1mm, 7 mm and 9mm respectively. With the increase in lift off, it is observed that difference in magnetic flux readings between the defective areas and background decreases. Hence it is impossible to track the defect growth on later scans based on this noisy data. The scans are conducted at step size of 0.5mm. Figure 6.4 shows the increase in uncertainty on later inspections due to the increase in step size. Scans on samples 2,3,4, are done at step size of 2mm and at lift of 2mm. This shows that detecting defective scan points based on flux readings become difficult as the step size increases.

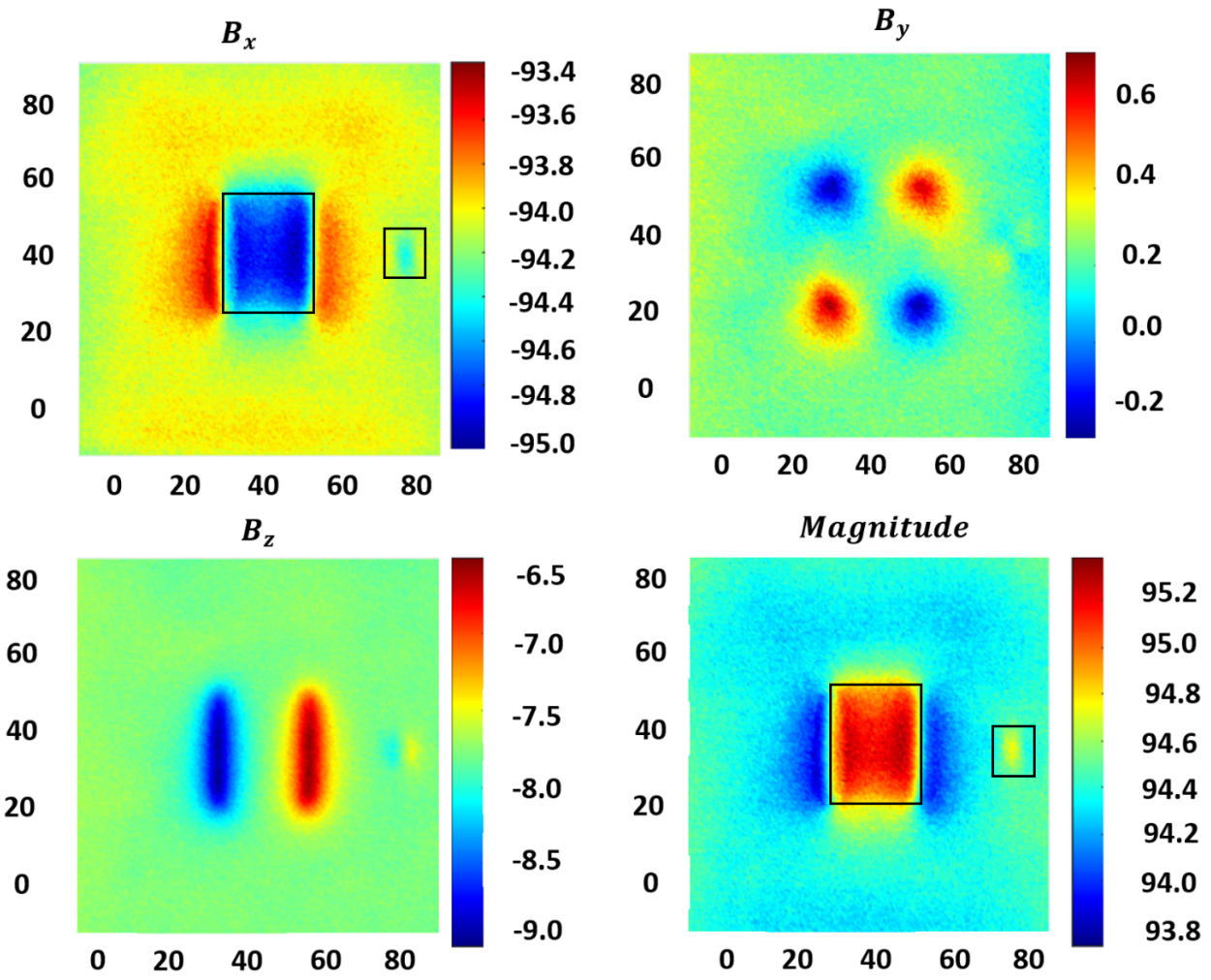

Figure 6.2 Image plot of the three axial magnetic flux readings detected by the MFL probe when scanned at lower lift off 1mm. At low uncertainty both the large defect and small defects get detected (as marked in black square).

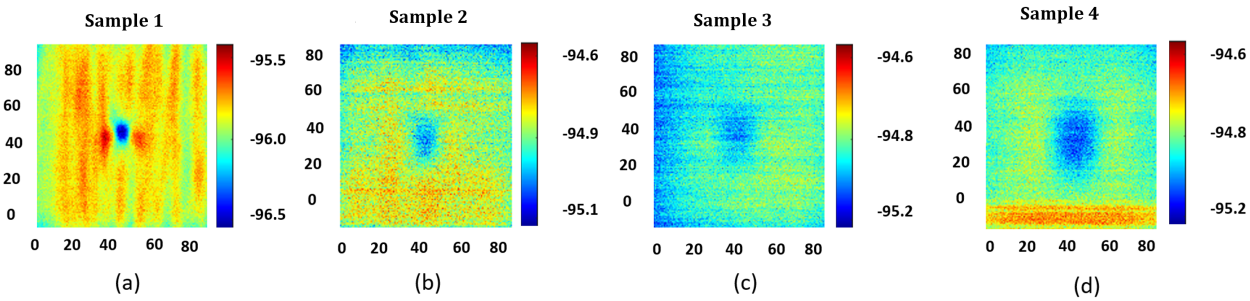

Figure 6.3 Plot the axial component of magnetic flux of different samples showing the difficulty of defect tracking at higher uncertainties due to lift off.

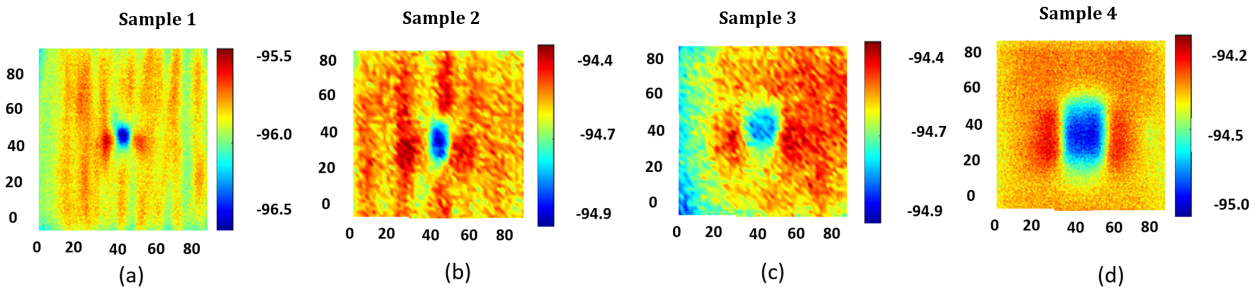

Figure 6.4 Plot the axial component of magnetic flux for different defect dimensions and lift offs showing the difficulty of defect tracking at higher uncertainties due to increase in step size.

6.3 Deterioration of SNR under different types of uncertainties

In section 6.3, the SNR of the experimental data sets that were collected based on the procedure described in the earlier section is studied. The objective is to evaluate how noise increases as NDE data is collected under uncertainties with increasing magnitude. For this objective MFL data collected from surfaces with a single defect is considered. Machined defects of different sizes and characteristics are fabricated. Collection of sequence of MFL data for each data is done as step sizes and lift off are increased. To systematically summarize these matrix-variate MFL data a simple bivariate normal mixture model for the MFL reading $X_{ij}^{(l)}$ from uncertainly level l at the (i, j) grid point is examined. For any fixed l, assume $X_{ij}^{(l)}$ is independent and identically distributed (i.i.d.) from $N(\Delta_1[l], \sigma_1^2[l])$ if (i,j) is actually a defective scan point and $X_{ij}^{(l)}$ is i.i.d. from $N(\Delta_0[l], \sigma_0^2[l])$ if (i,j) is actually a non-defective scan point. Tables 6.1 and 6.2 (to be placed in this section) show these parameters for the following four different defective samples as lift-off uncertainty is increased.

- Sample 1: Defect dimension $(l \times w \times h)$ $8mm \times 8mm \times 5mm$
- Sample 2: Defect dimension $(l \times w \times h)$ $2cm \times 1cm \times 5mm$
- Sample 3: Defect dimension $(l \times w \times h)$ $2cm \times 2cm \times 5mm$
- Sample 4: Defect dimension $(l \times w \times h)$ $3cm \times 2.5cm \times 5mm$

Where l, w, h represents the length, width and depth of the fabricated machined rectangular notch defects respectively.

The step size used in the MFL data summarized in table 1 is 0.5 mm. Table 6.2 reports the summary of the MFL data when the step size is 2mm. As the MFL data was collected in lab, we knew the position of the defects and the parameters reported in the table were calculated using this

Table 6.1 The mean and the standard deviations of the flux readings of the non-defective and defective regions are reported for the samples in Table 1. Regions are detected based on readings with no lift off. Here the flux readings are recorded at scan points 0.5mm apart. From the obtained parameters, SNR values are reported.

Sample No	**Lift off based Uncer-tainty**	Δ_0	σ_0	Δ_1	σ_1	**SNR**
1	No Lift off	-95.86	0.09	-96.41	0.07	4.78
	1mm	-96.11	0.09	-96.49	0.08	3.23
	2mm	-96.4	0.07	-96.57	0.05	1.99
	5mm	-96.66	0.07	-96.75	0.05	1.03
2	2mm	-94.46	0.07	-94.85	0.05	4.51
	5mm	-94.71	0.06	-94.89	0.05	2.26
	7mm	-94.83	0.06	-94.94	0.05	1.43
3	No Lift off	-94.44	0.11	-94.78	0.06	2.82
	2mm	-94.63	0.09	-94.83	0.06	1.79
	5mm	-94.89	0.07	-95	0.06	1.19
	7mm	-94.97	0.06	-95.06	0.05	1.07
4	No Lift off	-94.06	0.15	-94.84	0.08	4.71
	2mm	-94.47	0.09	-94.91	0.08	3.7
	5mm	-94.69	0.07	-94.95	0.07	2.7
	7mm	-94.8	0.06	-95.01	0.07	2.21
	9mm	-94.87	0.08	-95.02	0.05	1.65
	11mm	-94.96	0.05	-95.07	0.06	1.38

Table 6.2 The mean and the standard deviations of the flux readings of the non-defective and defective regions are reported for the samples in Table 6.2 . Regions are detected based on readings with no lift off. Here the flux readings are recorded at scan points 2mm apart.

Sample No	**Lift off based Uncer-tainty**	Δ_0	σ_0	Δ_1	σ_1	**SNR**
1	No Lift off	-89.83	0.1	-90.33	0.11	3.39
	2mm	-90.39	0.07	-90.61	0.08	2.02
	5mm	-90.79	0.07	-90.86	0.1	0.53
2	No Lift off	-94.48	0.06	-94.69	0.09	1.96
	2mm	-94.5	0.06	-94.65	0.13	1.09
	5mm	-94.49	0.24	-94.71	0.11	0.82
	7mm	-94.83	0.05	-94.87	0.05	0.53
3	No Lift off	-94.67	0.09	-94.91	0.05	2.41
	2mm	-94.65	0.09	-94.86	0.07	1.83
	5mm	-94.88	0.07	-95.02	0.06	1.54
	7mm	-94.98	0.06	-95.08	0.07	1.05
4	No Lift off	-94.2	0.17	-94.35	0.26	0.48
	2mm	-94.33	0.14	-94.47	0.21	0.56
	5mm	-94.55	0.09	-94.61	0.14	0.4
	7mm	-94.43	0.23	-94.31	0.34	0.29
	9mm	-94.89	0.07	-94.97	0.06	0.78
	11mm	-95	0.06	-95.05	0.06	0.66

knowledge.

Note that the bivariate normal mixture model considered here can be further improved as the readings at the scan points are spatially correlated. The bivariate normal mixture model with i.i.d. readings at grid points is mainly considered here as it provides the simplest and most intuitive and quite accurate understanding of the noise characteristics with increase in uncertainties in the data collection process. Based on the model, the difference in readings from a random defective and a random non-defective scan point is $N(\Delta_1 - \Delta_0, (\sigma_1^2 + \sigma_0^2)^{\frac{1}{2}}$. Thus, the SNR for defect detection in this model is $R = |\Delta_1 - \Delta_0|(\sigma_1^2 + \sigma_0^2)^{-\frac{1}{2}}$. In tables 6.1 and 6.2, these SNR values are reported. From the data it is observed that $\Delta_1 < \Delta_0$ across all regimes. The standard deviation σ_1 and σ_0 are in the same scale though they can differ by bit. As lift-off uncertainty increases in both tables, the gap $\Delta_0 - \Delta_1$ decreases sharply. Though the standard deviations also decrease but their decrease is not at the same rate, and this results in a massive decrease in R as lift-off-based uncertainty increases. On comparing table 6.2 with table 6.1 it is observed that the reduction in SNR is much pronounced when step size is doubled. Also, note that the reduction in SNR for the same level of lift off uncertainty is different across the four samples and the smaller defects have much higher reduction of SNR. It is to be noted that defect characteristics such as sizes and depth influence the nature of corruption in the MFL readings due to the presence of lift off and different data acquisition rates. In real world applications, the corruption level or noise type is unknown as there is no information of the defect characteristics beforehand. This implies that efficient defect detection data analysis method is a necessity that can work well across different noise regimes. Also, an efficient defect detection algorithm in the set-up needs to work well in noisy regimes where the true noise level cannot be known.

In tables 6.1 and 6.2 the first two moments of the noise distribution are reported. In Figure 6.5 the entire distribution of flux readings at the defective and non-defective scan points at different lift-offs for two different defect sizes are shown. Figure 6.5 shows that the bivariate normal model though very simple is a decent approximation model for explaining variations in MFL readings due to lift-off based uncertainties. Both the violin plots illustrate the decrease in difference of mean

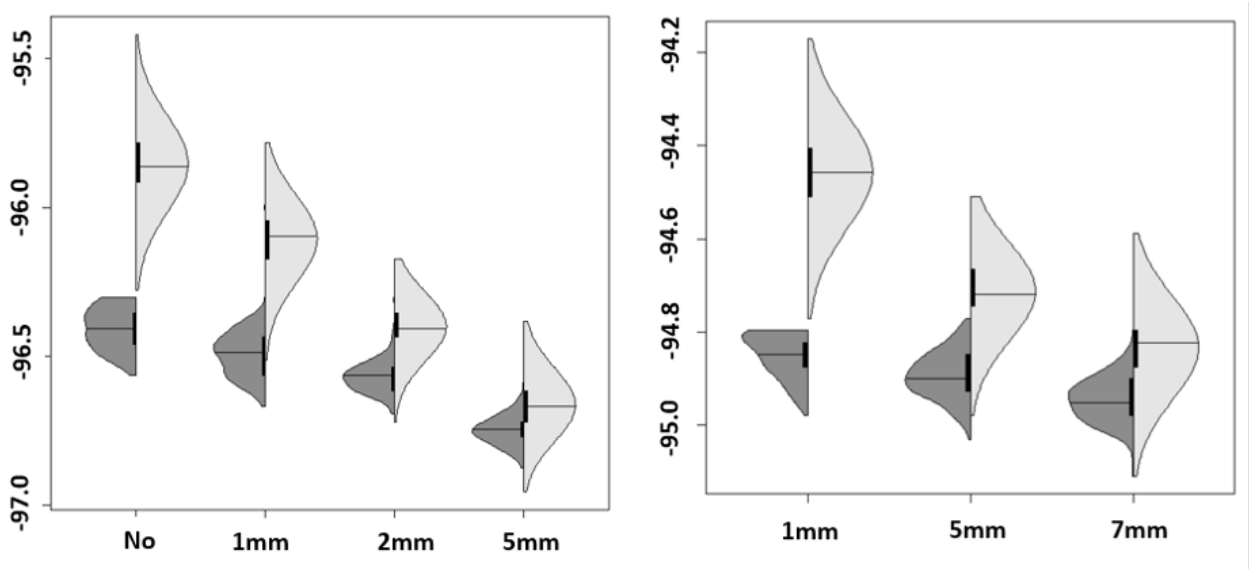

Figure 6.5 Violin plots showing distributions of flux readings for four cases of Sample 1 (a) no lift-off (b) 1 mm lift off (c) 2mm lift off (d) 5 mm lift off (in left) and three cases of Sample 2 a) no lift-off (a) 1 mm lift off (b) 5mm lift off (c) 7 mm lift off (in right). The distributions of the readings from defective scan points are shown in dark grey and distributions of non-defective areas are shown in light grey.

flux readings between the defective and non-defective area with increase in uncertainties in data collection procedure. As shown in Figure 6.5, the numerator of R (difference in mean) decreases whereas denominator (noise variance) stays mostly invariant, with increase in uncertainty. This causes a significant drop in SNR.

To understand the impact of the reduction of SNR with uncertainty on our task of locating and characterizing defect size, consider a pointwise hypothesis test (136; 137) to detect between defects and non-defective scan-points, i.e., to test hypotheses $H_0 : X_{ij}^{(l)} = N\big(\Delta_0, \sigma_0^2\big) \text{vs} H_1 : X_{ij}^{(l)} = N\big(\Delta_1, \sigma_1^2\big)$ for the grid point at $(i, j)^{th}$ coordinate. The most powerful test based on the Neyman-Pearson lemma classifies the grid point as defect if: $(\sigma_0^{(l)})^{(-2)}(X_{ij}^{(l)} - \Delta_0^{(l)})^2 - (\sigma_1^{(l)})^{(-2)}(X_{ij}^{(l)} - \Delta_1^{(l)})^2$ is large. The test will be applied for all grid-points in the lattice and the points classified as defects by the test will constitute towards the predictive defect size. Figure 6.6 shows the schematic for the procedure based on a single square defect at the center of the sample. In panel (a) of Figure 6.6, we have MFL readings for a less noisy scan where the defect is in blue, and the background is in green; in panel (b) we have the same data collected by a much noisier inspection. The dotted points represent the increase noise in panel (b) compared to panel (a).

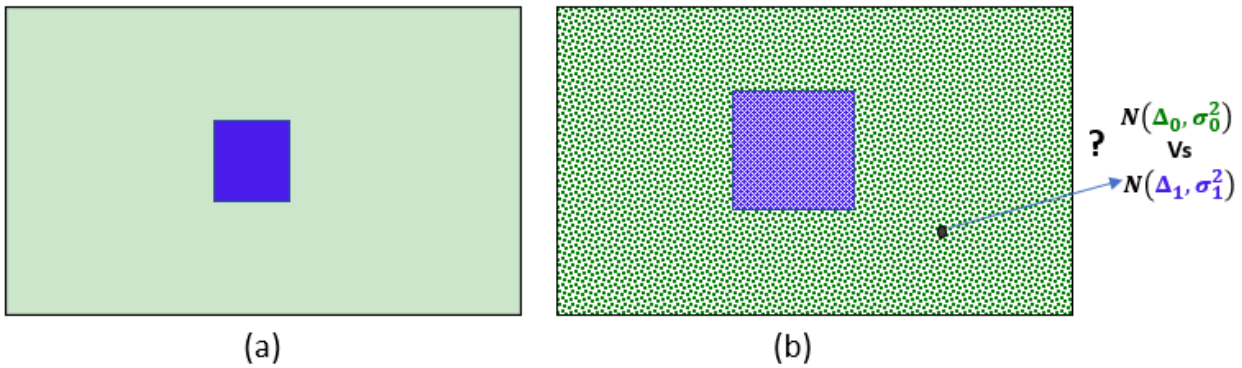

Figure 6.6 (a) Schematic showing the defective areas (marked in blue) and background (in light green) based on low noise data, (b) Noisy defective and non-defective areas of noisy data where pointwise testing fails shows the estimated defective scan points plotted in red as based on thresholding the posterior probabilities at 0.8.

Pointwise hypothesis testing will feature better in panel (a) than panel (b). To see this, we calculate the power and the false discovery rate (FDR) of the testing procedure. For simplicity, assume that $\sigma_0^{(l)} = \sigma_1^{(l)} = \sigma^{(l)}$. Then, the test classifies the (i,j) point as a defect if $X_{ij}^{(l)} - \frac{(\sigma_0^{(l)}+\sigma_1^{(l)})}{2}$ is large. We want this test to have very high power in detecting defective scan-points for unless we have very high probability in detecting defective scan-points we will underestimate defect size. We consider a test with power β. The false discovery rate (FDR) is the probability with which such a test misclassifies a non-defective point as a defective point. We prove in the appendix that the FDR for the most-efficient pointwise hypothesis test of power β is $\Phi(\Phi^{(-1)}(\beta) - 2R^{(l)})$ where, Φ is the standard normal cumulative density function and $R^{(l)}$ is the SNR. These mis-classified non-defective points can lead to severe overestimation of the defect sizes. Typically, defects constitute a small fraction p of the considered lattice in any inspection window. Thus, the overestimation of defect size by the most-efficient power β test is $p^{(-1)}\Phi(\Phi^{(-1)}(\beta) - 2R^{(l)})$ times of the original defect size. Table 6.3 shows the defect size overestimation by the test with 95% power as SNR and true size of the defect varies. Note that, the reported overestimation proportion are the theoretically achievable best possible rate as in practice the parameters $\Delta_1^{(l)}, \Delta_0^{(l)}, \sigma_1^{(l)}, \sigma_0^{(l)}$ are also unknown and need to be estimated from the data. In our application, it is extremely easy to estimate the $\Delta_0^{(l)}, \sigma_0^{(l)}$ with high precision based on the background. Estimating $\Delta_1^{(l)}, \sigma_1^{(l)}$ is the challenging part. Their estimation error will further inflate the defect overestimation metric. Comparing tables 6.1, 6.2

with table 6.3, we see that even with moderate uncertainty in the data collection process, defect sizes are grossly overestimated. This creates a huge problem as it implies that based on such this test, we will be wrongly classifying a lot of panels with benign exceedingly small defects such as scratches as defective panels.

Table 6.3 Oracle overestimation error for detecting defects at 95% power.

SNR	**1.00%**	**2.50%**	**5.00%**	**7.50%**	**10.00%**	**20.00%**
4	0	0	0	0	0	0
3.5	0	0	0	0	0	0
3	0	0	0	0	0	0
2.5	0.04	0.02	0.01	0.01	0	0
2	0.94	0.38	0.19	0.13	0.09	0.05
1.5	8.85	3.54	1.77	1.18	0.89	0.44
1	36.32	14.53	7.26	4.84	3.63	1.82
0.5	74.22	29.69	14.84	9.9	7.42	3.71

In the following section, we found that only pointwise independent hypothesis testing methodologies but spatial testing methods that borrow strength from neighboring scan points are also unable to operate well in noisy regimes. We provide a binned hypothesis testing method to increase SNR. A crucial step in implementing the binned hypothesis testing method is to use transfer learning for estimating $\Delta_1^{(l)}$, $\sigma_1^{(l)}$ in the bivariate normal mixture model set-up.

6.4 Transfer learning based binned hypothesis tests (TLBH) for defect detection

In section 6.4 we develop a binning based hypothesis testing methodology [45,46] to increase SNR and conduct hypothesis testing. Binning of the grid points can be done by several ways (a) random binning, (b) aggregate binning by considering the average of the block of grid points, (c) aggregate binning by observing kernel density estimate and (d) batch-based binning. In binned hypothesis testing, we perform tests after aggregating the data into bins. Here, we consider an equally weighted blocked binning test. The binning is done by considering K points of the grid in the form of rectangles. The choice of the block size is dependent on the difference in mean of the null and alternate distribution. K is a tuning parameter in our model. Increasing the value

of K results in poor resolution whereas decreasing the value of K results in erroneous pointwise classification. The set-up of the binned test is as follows. Considering the flux readings to be independent and identically distributed (i.i.d.), for K points in a block in the rectangular grid, the null distribution becomes: $H_0 : X_1, X_2, \cdots, X_K i.i.d N(\sigma_0, \sigma_0^2)$ whereas the alternate distribution is $H_1 : X_1, X_2, \cdots, X_K i.i.d N(\sigma_1, \sigma_1^2)$. Since the readings are i.i.d, for the mean of the rectangular grid $\overline{X_K}$ the variance will be reduced by K times. For testing the mean of the K^{th} block $\overline{X_K}$ the null distribution will be $H_0 : \overline{X_K} = N(\Delta_0, \frac{\sigma_0^2}{K})$, whereas the alternate distribution will be $H_1 : \overline{X_K} = N(\Delta_1, \frac{\sigma_1^2}{K})$. As the noise variance decreases compared to unbinned pointwise test ($K = 1$), the binned hypothesis testing provides better defect tracking on noisy inspections. In the previous section, we showed that in the presence of uncertainties, the means and standard deviations of the noise and defect MFL distributions vary greatly. As we do not know the noise type beforehand, we need a test that is agnostic of the type of noise but works universally well across all noise regimes. For this purpose, we use transfer learning. We project information regarding the location of defective areas from earlier less noisy scans. Accurate estimation of the locations of the defective scan points can be made from the MFL data with very low noise. We transport these locations to noisy MFL data collected in the following scans. In the noisy MFL data collected later, the defects can stay same or grow in size. Thus, the transferred location of the defective scan points for the earlier less noisy probe helps us in precise estimation of $\Delta_1^{(l)}, \sigma_1^{(l)}$ in the current noisy data collected by the cheap probe. Estimating $\Delta_o^{(l)}, \sigma_0^{(l)}$ from the non-defective scan-points is easy as there is a lot of background signals in the data. Once these parameters are correctly estimated, we bin the data into blocks and implement a power β test. Figure 6.7 illustrates the difference in the framework of binned hypothesis testing and point-wise hypothesis testing. The green and the filled blue boxes respectively denote the supports of non-defect (background) and defect scan points from an earlier scan. The defect has grown from before and the growth area is shown in light blue. The parameters $\Delta_1^{(l)}, \sigma_1^{(l)}, \Delta_o^{(l)}, \sigma_0^{(l)}$ are however calculated based on the smaller blue area and the green outlined area. Though the binning algorithm increase SNR of the signal it does not always outperform unbinned ($K = 1$) tests as the block wise aggregation in binned tests

leads to loss of resolution. Particularly, in large SNR a binned test will be sub-optimal. For this reason, it is extremely important to optimally choose the block size K in binned test. An optimally binned test should be adaptive to the SNR in the data [47, 49-51]. As using transfer learning, we have accurate estimates of all the four parameters of the bivariate normal mixture model, we can accurately estimate the SNR $R^{(l)}$ corresponding to the current data. Binning on a block of size K increases the SNR to $R^{(l)}\sqrt{K}$. If p was the size of the defect in previous scan, the overestimation error of a power β test is controlled at $p^{(-1)}\Phi(\Phi^{-1}(\beta) - 2R^{(l)}\sqrt{K})$. We select the minimum $K \geq 1$ such that the above is controlled below 10%.

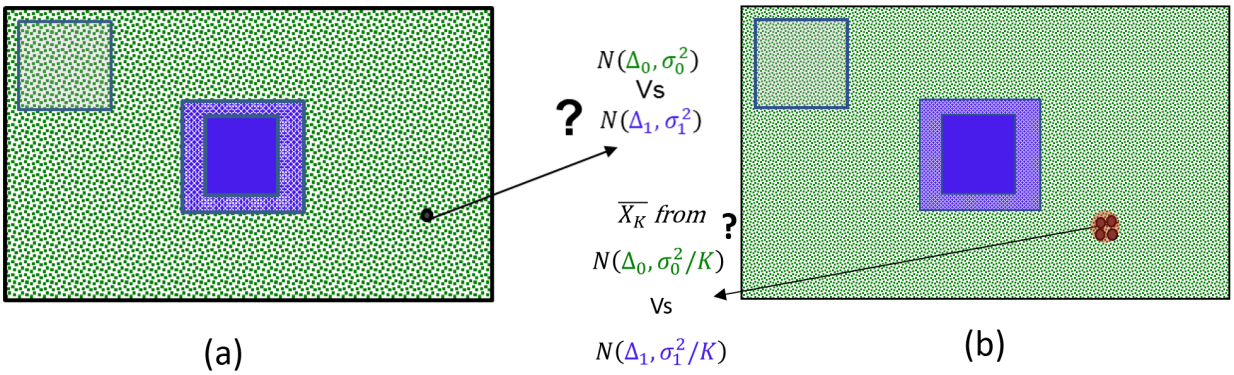

Figure 6.7 Schematic of pointwise hypothetical testing, (b) Schematic of optimal binned hypothesis testing based on multiple adjoining points (rectangular grid is shown in yellow).

Figure 6.8 shows the results of TLBH algorithm in tracking defect size for one such case where the defect has increased in size on later inspections. The baseline inspection is done on Sample 1 at low lift off. On Inspection 2 conducted at higher lift off of 2 mm, the defect size remains constant. On subsequent inspections on Inspection 3, 4 and 5 conducted at higher lift offs of 5 and 7 mm respectively, TLBH algorithm is successful in detecting the defective areas. The coverage and FDR are reported in comparison Table 6.5 along with the defect inspection patterns in Table 6.4.

6.5 Comparison with other competing defect tracking algorithms

6.5.1 Mixture Regression

In this sub section, we have used a mixture regression method (125) for fitting a bivariate function to less noisy MFL data as developed in details in Chapter 5. Mixture regression is a

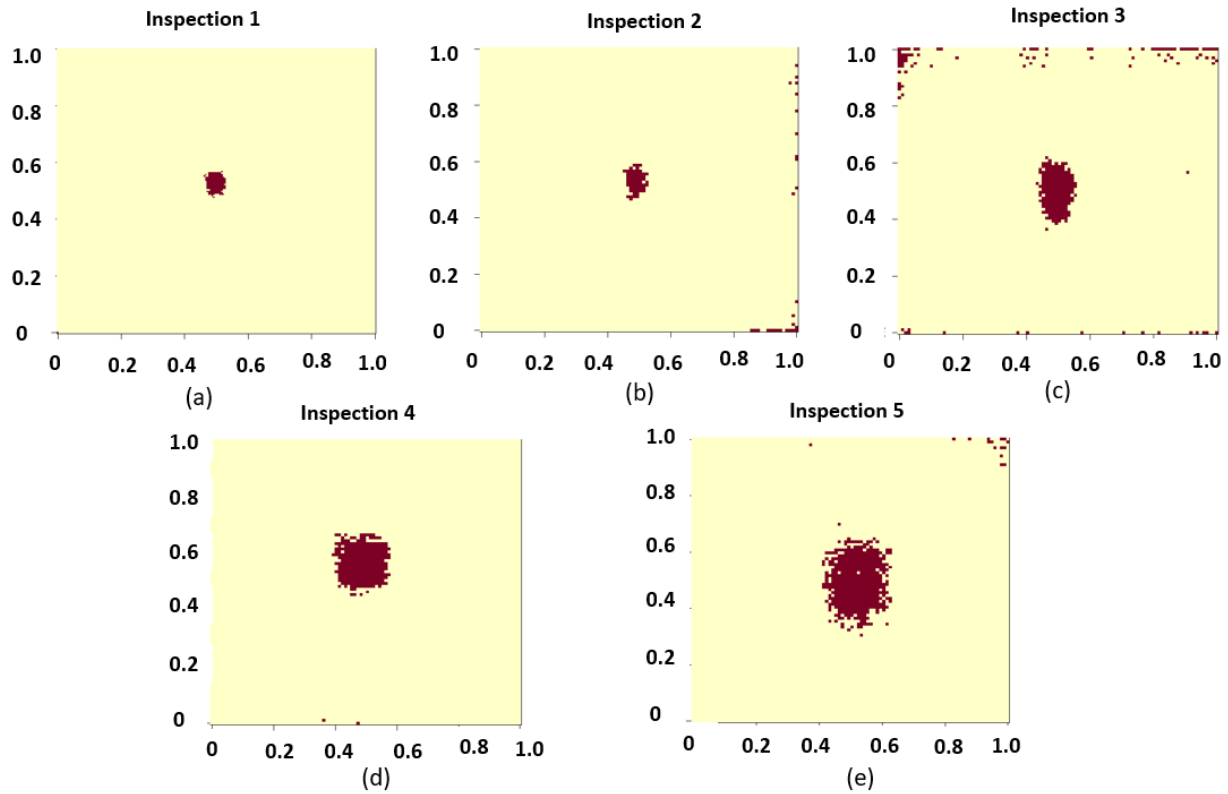

Figure 6.8 Baseline Inspection 1 is done on Sample 1, (b) In Inspection 2 the defect size remains constant and the scan is done at 2 mm lift off. TLBH algorithm is able to track the defect area in the presence of noise, (c) Inspection 3 is done on Sample 2 at 5mm lift off where the defect has increased, (d) Inspection 4 conducted on Sample 3 at 5 mm lift off, (e)Inspection 5 is conducted at 7 mm lift off. In all the subsequent inspections TLBH algorithm tracks the defect growth (defective area in brown, background in light yellow).

powerful statistical technique (127; 126) for detecting sub populations within a larger population. The location of each scan point here is treated as an explanatory variable and used to train the model, which generates a probabilistic classification of scan points to fall in either of two categories: defect or non-defect category. At first the matrix-variate MFL data is vectorized to set up the regression model. The unknown intensity generating function f can be next modelled as:

$$f(x_i, y_j) = \sum_{k \in D,S}^{2} P_k \cdot N(\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j, \phi_k^2) \tag{6.1}$$

where, x_i, y_j denotes the coordinates of the scan points in grid. The number of mixtures is 2 as a scan point can either fall in a defect or a non-defect cluster; D denotes the class of defective points on the grid and S is the set of non-defective scan points. For each mixture model, the intensity is modeled by a linear surface$\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ with aberrations having variance ϕ_k^2. Here

$\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ are the regression coefficients and ϕ_k^2 the regression error for $k = D, S$ and P_k represents the mixing weights for each of the k regression models. If ϕ_k^2 is low, then there is not much variability, which will be the case for defect less background scenario. For areas near defects, the average intensity $\theta_0^{(k)} + \theta_1^{(k)} x_i + \theta_2^{(k)} y_j$ will be different leading to the points in those areas having other mixing density. Thus, spatial relation among the points are leveraged here. Also, for point adjoining defects the variance ϕ_k^2 will be large. The mixture model is trained using an expectation-maximization (EM) algorithm, where the mean and covariance of each component are iteratively updated until the convergence criteria are met [54]. For each point (x_i, y_j), the posterior probabilities of $\hat{\pi_D}(x_i, y_j)$ are evaluated and on the basis of this value, we classify each point to be either defective (D) or non-defective (S) area. The R package mix tools [55] is adopted for fitting these mixture regression functions. Figures 6.9 show the application of the procedure where it succeeds on defect detection on less noisy MFL data.

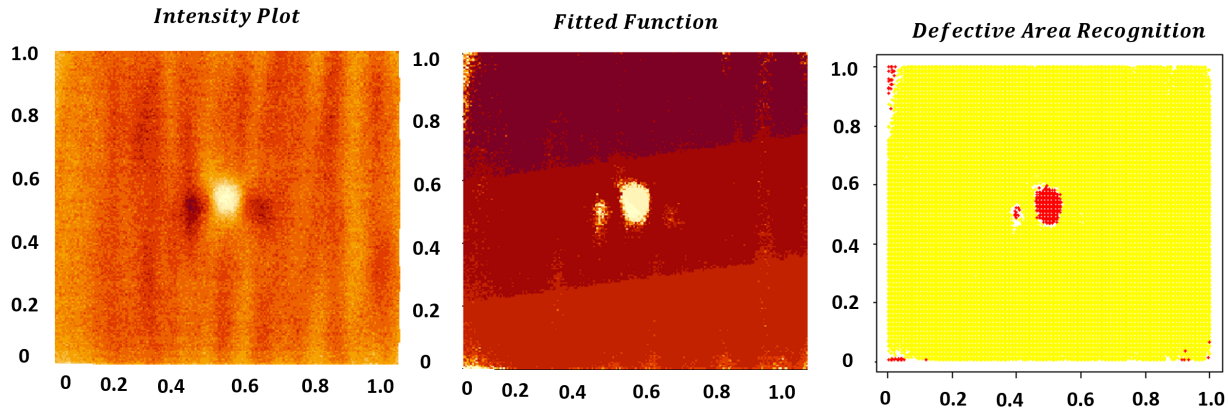

Figure 6.9 Mixture regression succeeds in automatically detecting the defective areas based on less noisy MFL scan.

However, the two-mixture regression fails when the data are noisy due to uncertainties. In Figure 6.10, we show that, on conducting scan at lift off of 2mm on Sample 1, the mixture regression fails to detect the defective areas. As a result, this algorithm is not fit to be used to analyse data with noise. In Table 6.5 the comparison of the novel optimized TLBH algorithm with that of mixture regression and thresholding based on wavelet based denoising on the MFL data collected with various step size and lift off based uncertainties is presented.

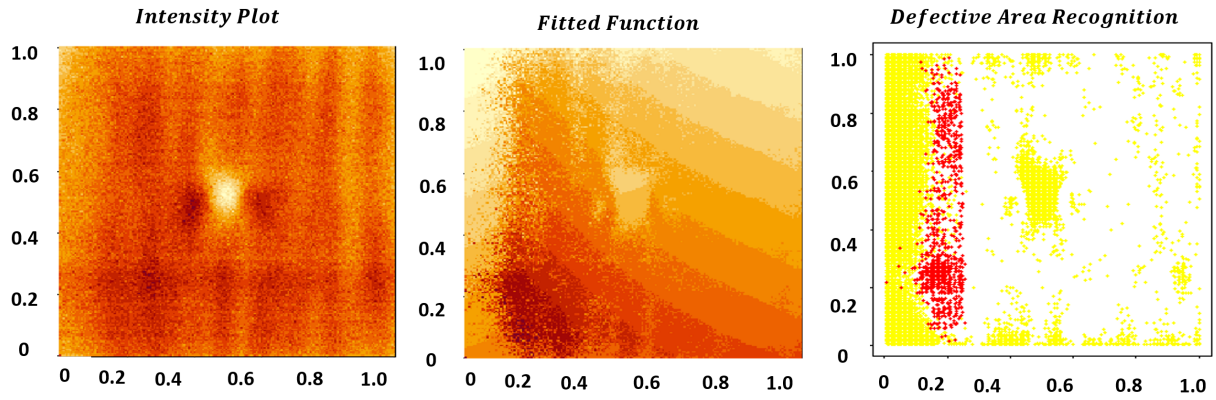

Figure 6.10 Mixture regression fails to detect the defective area due to the increase in lift off of 2mm.

6.5.2 Element wise thresholding based on denoised MFL data

For comparison, denoising based on the waveslim package of R [56] is conducted next. *Waveslim* denoise a raster scan image using discrete wavelet transform (DWT) method. On the denoised version of the MFL image, then elementwise thresholding is applied. The naïve thresholding estimator $\hat{\theta}_T[l] = 1\hat{y}_l < c_\alpha$ is considered on the reconstructed image from the wavelet coefficients, which classifies l_{th} scan point as defect if the MFL reading is below the α^{th} quantile of flux readings in the grid. In Table 6.5 two values of α at 25% and 5% are chosen. There is a trade-off between coverage and overestimation. At lower values of α, both the coverage and overestimation are low. However, at higher values of α the coverage increases with the cost of increase in overestimation.

6.5.3 Mixture regression with aid of Transfer learning for dynamic defect tracking

The naïve (non-TL) mixture regression algorithm is improved by transfer learning [57] where the defect locations and mixing proportions are leveraged as discussed in (29). The intensity matrix from less noisy scan denoted as I_1 is quite accurate with very low noise compared to the later noisy ones. The objective is to update the sizes of the existing defects based on I_1. Based on Inspection I_1's data, the location and proportion of the defective points sampled in the probe can be estimated. In Figure 6.11 the schematic of the working principle of Transfer learning aided Mixture regression (TLMR) is presented. At first the non-TL mixture regression is applied on the less noisy data. Scan points with posterior probability greater than 0.8 which is a tuning parameter are considered as

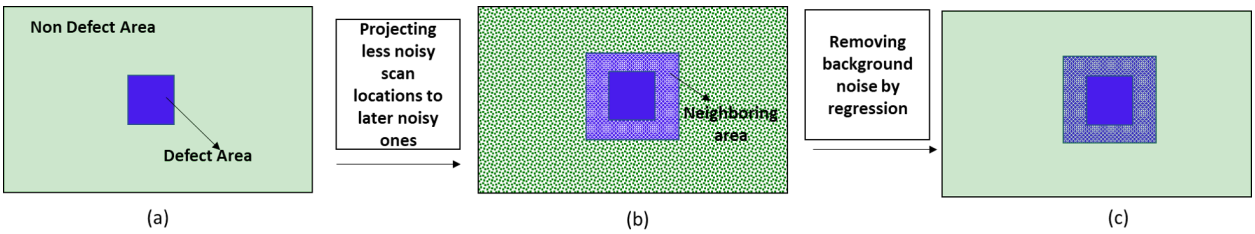

Figure 6.11 (a)Less noisy scan, (b)projection of the background information from less noisy to noisy scans, (c) Regress the defect and non-defect values of noisy scans by less noisy ones.

defective area. The defective area is in deep blue (say) and the background is in light green as shown in Figure 11 (a). The noisy scans are marked by dots to represent noise. In the later scans, the defect can grow in size which is marked by light blue as neighboring area. As later scans are noisy, hence the defective and non-defective areas from I_1 are projected on the same locations on later scans so as to obtain an estimate of the defect. Then as shown in step (c) of figure 6.11 , linear regression is applied to regress the defective and non-defective areas of later scans by I_1 and replace the values of those areas in later scans by the predicted values from regression. This results in the denoise of the defective and non-defective areas as shown in (c). Only the neighboring area is noisy. As the amount of noisy area has diminished to a large extent, we developed mixture regression model with two groups on the response values consisting of denoised areas by regression and the noisy neighboring area. In Figure 6.12, mixture regression is implemented on one such noisy scenario where non-TL mixture regression fails to detect the defective areas on scanning Sample 1 at a lift off of 5mm (shown in Figure 6.12 (b)). On leveraging the defect location and mixing proportions from the less noisy scans, mixture regression succeeds in detecting the defective areas as shown in 12 (c). It is to be noted in this scenario the defect size stays constant.

Another scenario where a defect (Sample 2) of dimension $2cm \times 1cm$ (in blue in figure 6.13 (a)) has increased to Sample 3 of dimension $(2cm \times 2cm)$ is shown in Figure 6.13. Sample 2 scan is less noisy as it is scanned at lift off of 2 mm whereas Sample 3 is noisy being scanned at 7mm. On less noisy scan, mixture regression is fitted and mixture regression is successful in tracking the defective areas. Figure 6.13 (c) shows the sample 3 experimental data that is noisy where the defect cannot be observed. 13(d) shows the defective area recognition on Sample 3 after fitting mixture regression by transferring location from less noisy Sample 2. However, figure 6.13 (e) of the target

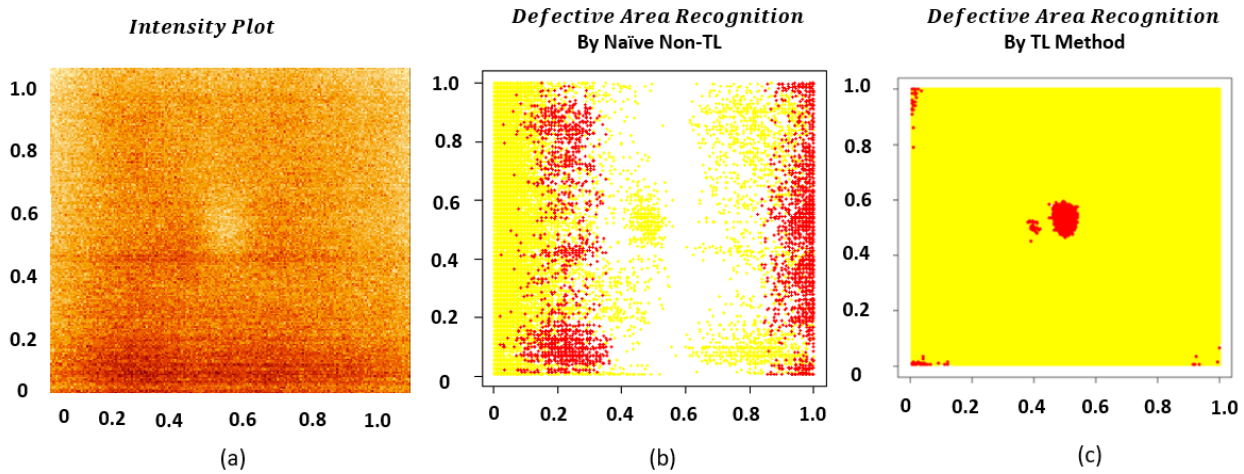

Figure 6.12 (a) Raw flux readings of Sample 1 at lift off 5mm, (b) Estimated defective (in red) and non-defective (in yellow) by mixture regression without TL, (c) Estimated scan points by TL aided mixture regression. Estimation is done by thresholding the posterior probabilities at 0.8. In all the sub figures the Y and X axes are normalized in [0,1].

flux image exhibits a bit erroneous defect estimation of 13 (d).

This happens due to the following reasons: (a) The neighboring area is a tuning parameter in our model where the defect can grow in later scans. If in reality, the defect has grown by a significant amount between two successive inspections then the algorithm may fail to track the defect. (b) The neighboring area is noisy and there, while fitting the mixture model, we are performing pointwise classification based on posterior probabilities. Since the neighboring area is noisy, there is a high chance of misclassification. Moreover, this TL based mixture regression is an iterative algorithm with high time consumption. In order to track defect growth all previous inspections are to be considered. Thus, TLBH outperforms all the stated algorithms in defect tracking. In Table 6.4 we have exhibited various dynamic defect tracking test cases. Here different sample numbers represent different growing defects across Inspections (defect dimensions discussed in Section 6.3). Table 6.5 shows the comparison of the TLBH algorithm with wavelet based denoising and mixture regression in terms of coverage and FDR. Our algorithm worked well uniformly across almost all the Inspections with good coverage and moderately less FDR. Other proposed methods like Mixture regression and wavelet-based thresholding worked well when the defect size gets large. However, when the defect size is small with increase in uncertainties, SNR gets reduced. In these situations, the competing algorithms either fair bad on coverage or have overestimated the defect size by a

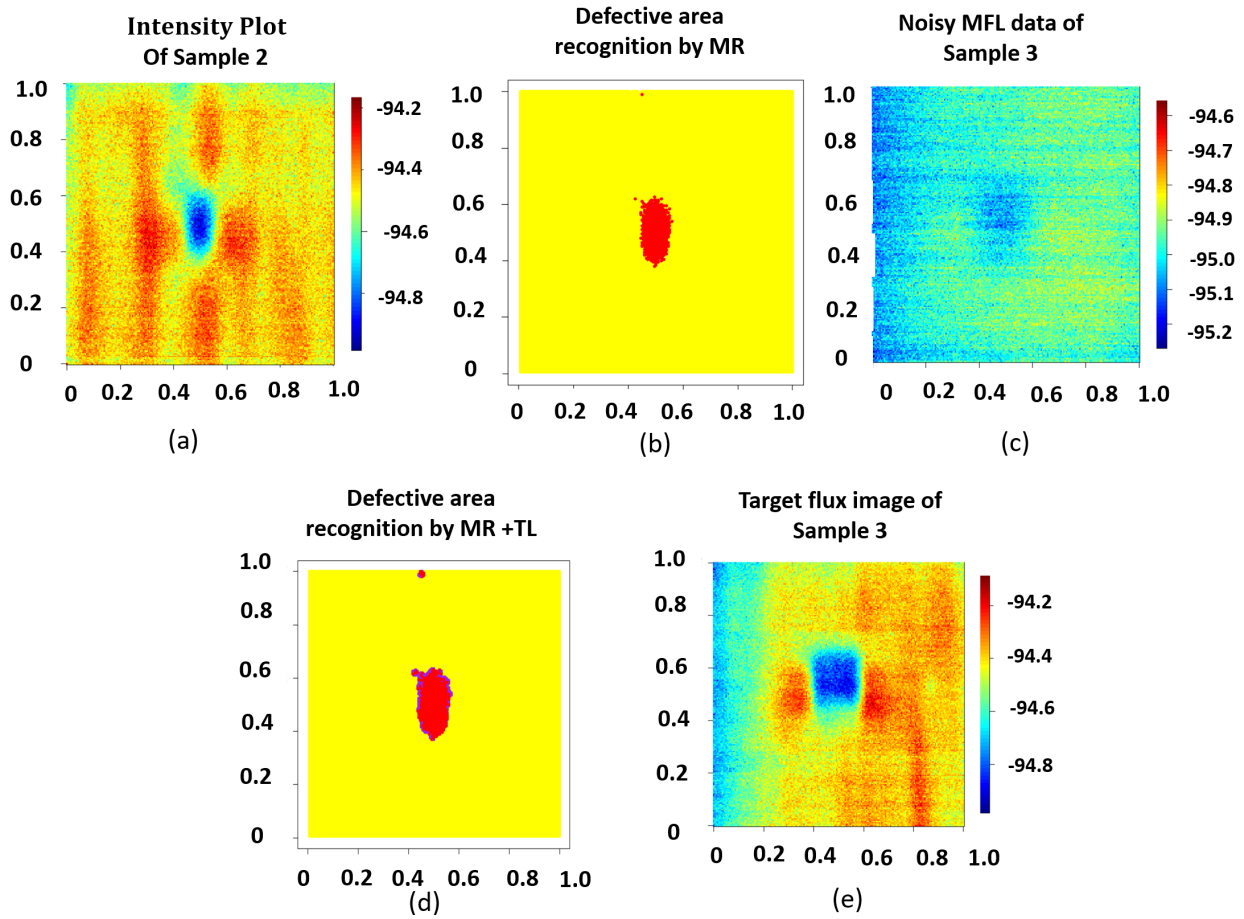

Figure 6.13 (a) MFL flux intensity of sample 2 at low lift off of 2mm, (b) mixture regression succeeds in detecting the defective areas based on less noisy scan, (c) flux intensity of Sample 3 at much higher lift off, (d) defective area tracked by TL aided MR, (e) Target flux image of Sample3.

large extent.

6.6 Application of TLBH in Eddy current based defect tracking

Eddy current (EC) testing is a popular NDE technique like MFL for inspection of metallic conducting pipelines. In this section we have shown the applicability of our TLBH algorithm on the experimentally generated EC data. The EC technique take into consideration electromagnetic induction based on the inductive properties of the alternating current (AC) to detect anomaly [58,59]. On application of time varying magnetic field, eddy currents get generated in the conducting specimen. Figure 6.15 shows the block diagram of the EC experimental setup along with the inhouse EC sensor. The EC sensor consists of two coils (a) transmitter coil, (b) receiver coil to sense the signal from MUT. The diameter of each coil is 6mm and each coil has 2 layers. Number

Table 6.4 Different dynamic defect tracking test cases for comparison.

	Inspection1	**Inspection2**	**Inspection3**	**Inspection4**	**Inspection5**
Case 1	Sample 1 at no lift off	Sample 1 at lift off of 2 mm	Sample 2 at lift off of 5mm	Sample 3 at lift off of 5mm	Sample 4 at lift off of 7mm
Case 2	Sample 1 at no lift off	Sample 2 at 5mm lift off	Sample 2 at 7mm lift off	Sample 2 at 7mm lift off and step size of 2mm	-
Case 3	Sample 1 at no lift off	Sample 1 at 1mm lift off	Sample 1 at 2mm lift off	Sample 1 at 5mm lift off	Sample 3 at lift off of 5mm
case 4	Sample 1 at no lift off	Sample 3 at 5mm lift off	Sample 3 at 7mm lift off	Sample 4 at 2mm lift off	Sample 4 at 5mm lift off
Case 5	Sample 1 at no lift off	Sample 4 at 2mm lift off	Sample 4 at 7mm lift off	Sample 4 at 11mm lift off	-

Table 6.5 Detailed comparison of the three methods: TLBH, Mixture regression and wavelet-based thresholding following the test cases of Table 6.4. α_1, α_2 are chosen as 0.05 and 0.25 quantile respectively.

		TLBH		**Mixture Reg**		**Wavelet at α_1**		**Wavelet at α_2**	
	Inspection	**Coverage**	**Overestimation**	**Coverage**	**Overestimation**	**Coverage**	**Overestimation**	**Coverage**	**Overestimation**
Case 1	1	1	0	1	0.67	1	10.7	1	56.89
	2	0.88	1.12	0.007	11.55	0.84	10.8	1	57.47
	3	0.99	2.5	0.635	0.69	1	5.32	1	30.21
	4	0.92	0.56	0.27	1.08	0.37	0.6	0.97	3.87
	5	0.93	0.39	0.3	0.004	0.85	0.4	0.99	4.71
Case 2	1	1	0	1	0.67	1	10.7	1	56.89
	2	0.96	2.7	0.635	0.69	1	5.32	1	30.21
	3	0.82	1.72	0.11	0.23	0.98	5.37	1	30.19
	4	0.66	1.05	0.1	0.21	0.54	6.8	0.84	42.72
Case 3	1	1	0	1	0.67	1	10.7	1	56.89
	2	0.99	0.02	0.01	9.77	0.87	10.9	1	56.32
	3	0.87	1.12	0.007	11.55	0.844	10.8	1	57.47
	4	0.53	4.25	0.0026	8.51	0.39	11.5	0.89	57.24
	5	0.92	0.94	0.27	1.08	0.367	0.6	0.97	3.92
Case 4	1	1	0	1	0.67	1	10.7	1	56.89
	2	0.92	0.94	0.27	1.08	0.367	0.6	0.97	3.87
	3	0.64	0.35	0.25	1.7	0.325	0.664	0.97	3.92
	4	0.98	0.01	0.56	0.007	0.844	0.312	1	4.71
	5	0.96	0.24	0.43	0.006	0.87	0.28	1	4.73
Case 5	1	1	0	1	0.67	1	10.7	1	56.89
	2	1	0.04	0.56	0.007	0.844	0.312	1	4.73
	3	0.91	0.14	0.3	0.004	0.85	0.78	0.99	4.71
	4	0.62	0.21	0.24	0.88	0.81	1.12	0.84	4.69

of turns in each layer is 9, making a total of 18 turns in each coil. The spacing between each coil is 0.2mm. A waveform generator is used to generate a low frequency (500kHz) sinusoidal signal which gets feed to the transmitter coil. The presence of crack like defects redistributes the induced current, which cause a change in the eddy current thereby altering the impedance of the receiver coil. The reflected signal captured by the receiver coil is then passed to the lock in amplifier which generates a signal proportional to difference between the reflected and a reference signal. The signal is next passed to National Instruments data acquisition card and to computer for visualization of the real and imaginary components of the impedance.

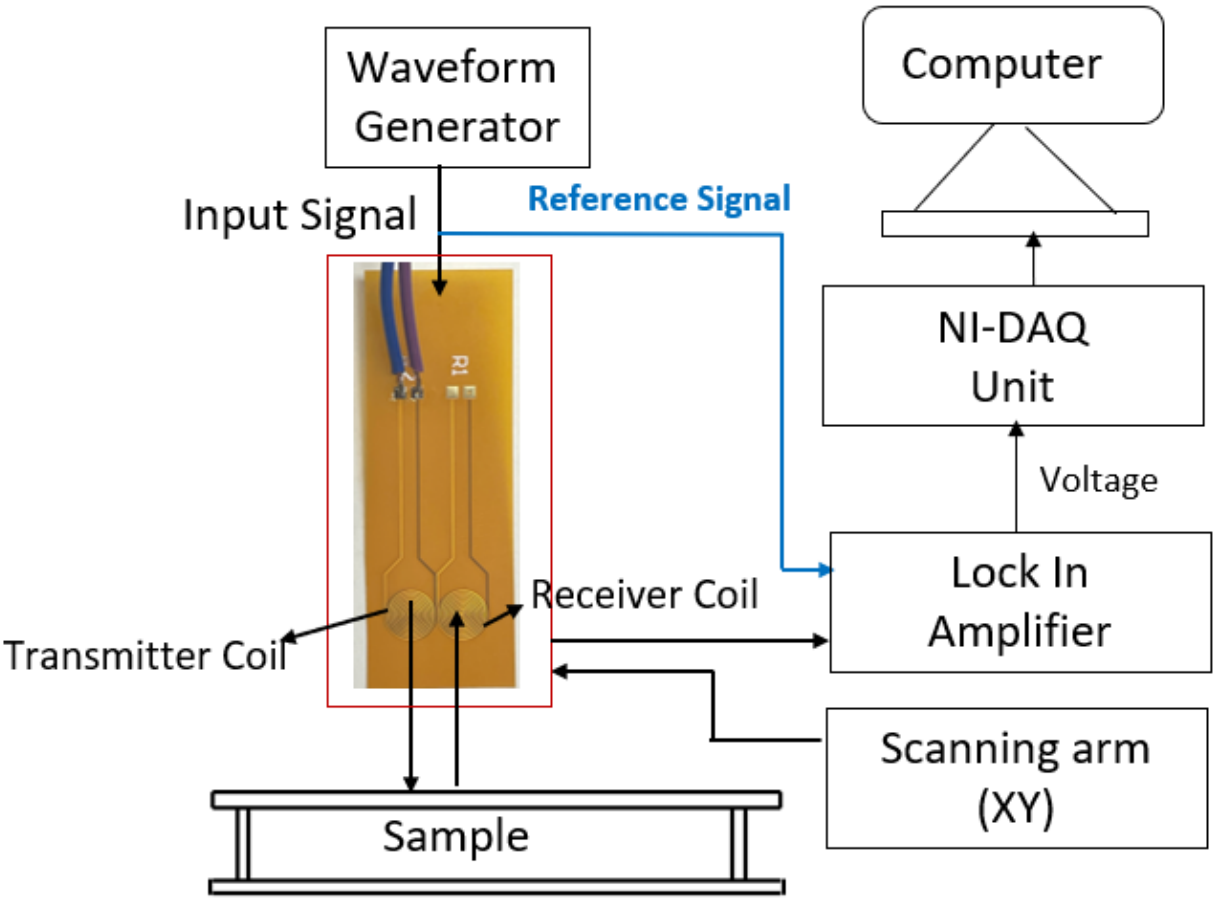

Figure 6.14 Schematic of the EC experimental setup.

Our algorithm is validated on various steel samples containing sub millimeter slits and rectangular notch defects. Uncertainty is introduced by scanning at various lift offs and step sizes as

reported in Table 4. The defect dimensions are as follows: (a) Slit of dimension $1cm \times 2mm \times 1mm (l \times w \times h)$, (b) Slit of dimension $1cm \times 3mm \times 2mm$, (c) dimension of rectangular notch defect $8mm \times 8mm \times 5mm$, (d) which increases later to dimension of $2cm \times 1cm \times 5mm$, (e) notch becomes of dimension $2cm * 2cm * 5mm$, (f) defect dimension $3cm \times 2.5cm \times 5mm$ In our analysis, we have used the amplitude of the impedance values where $amplitude = \sqrt{(real^2 + imaginary^2)}$. In figure 6.15, it is observed that there is signal uncertainty due to the variation in depths of the slit like defects. In figure 6.15, slit of depth 1mm results in weak defect signal compared to a slit of depth 2mm when scanned at a lift off of 3mm. The step size of each scan is 1mm. Due to agnostic nature of the noise due to the various uncertainties, our algorithm needs to be robust for accurate detection.

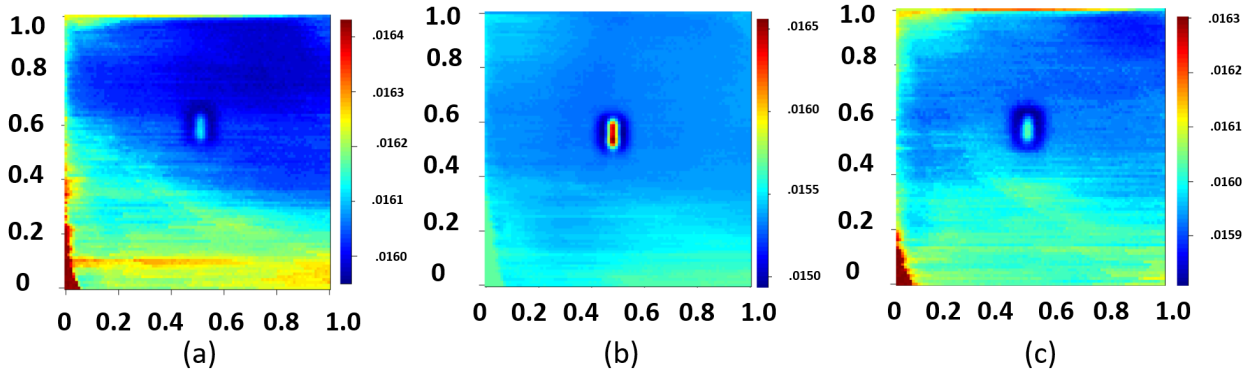

Figure 6.15 (a) Image plot of the amplitude of impedance readings for a slit of depth 1mm, (b) Image plot of a slit of depth 2mm. Both the scans are done at lift of 3mm. (c) shows the image plot of 2mm depth slit at lift off of 5mm.

Figure 6.16 shows the impedance signal images across three different lift off scenarios of sample 2 and sample 4. With the increase in lift off (10 mm) the acquired signal is corrupted with high noise. Similar to Tables 6.1 and 6.2 , for the EC experiments we report the parameters of the distributions in Table 6.6 using oracle information for different lift off scenarios and then have evaluated the SNR. Table 6.6 shows there is a sharp decline in SNR with the increase in liftoffs.

In Table 6.6 we have exhibited various dynamic defect tracking test cases. Here different sample numbers represent different growing defects across Inspections (defect dimensions discussed in Section 6.3). Table 6.7 shows the comparison of the TLBH algorithm with wavelet based denoising and mixture regression in terms of coverage and FDR. Our algorithm worked well uniformly across

Table 6.6 Mean and standard deviations of the impedance readings of the non-defective and defective regions are reported. Regions are detected based on readings with no lift off.

Sample No	**Lift off based Uncertainty**	Δ_0	σ_0	Δ_1	σ_1	**SNR**
Slit 1	No Lift off	0.012	0.00021	0.009	0.00023	8.92
	3mm	0.016	0.00002	0.016	0.00009	0.92
Slit 2	No Lift off	0.0158	0.0002	0.0137	0.0001	10.8
	3mm	0.016	0.0003	0.0154	0.0001	2.5
	5mm	0.016	0	0.016	0.0001	0.51
Sample 1	No Lift off	0.01	2.00E-07	0.007	0.00154	0.17
	3mm	0.02	6.00E-06	0.016	0.00005	0.34
Sample 2	No Lift off	0.021	0.001811	0.0066	0.000889	7.1
	5mm	0.016	0.00004	0.0159	0.000049	1.72
	10mm	0.015	0.000008	0.0152	0.000009	1.47
Sample 3	No Lift off	0.017	0.00018	0.014	-0.00256	1
	2mm	0.016	0.00011	0.01645	0.00038	0.96
	5mm	0.015	0.00004	0.01566	0.00022	0.98
	10mm	0.015	0.00001	0.01511	0.00003	0.96
Sample 4	No Lift off	0.016	0.002233	0.00687	0.000397	4.2
	2mm	0.016	0.000295	0.01642	0.000136	1.7
	5mm	0.015	0.000135	0.0159	0.000073	2.75
	10mm	0.015	0.000014	0.01509	0.000012	2.86

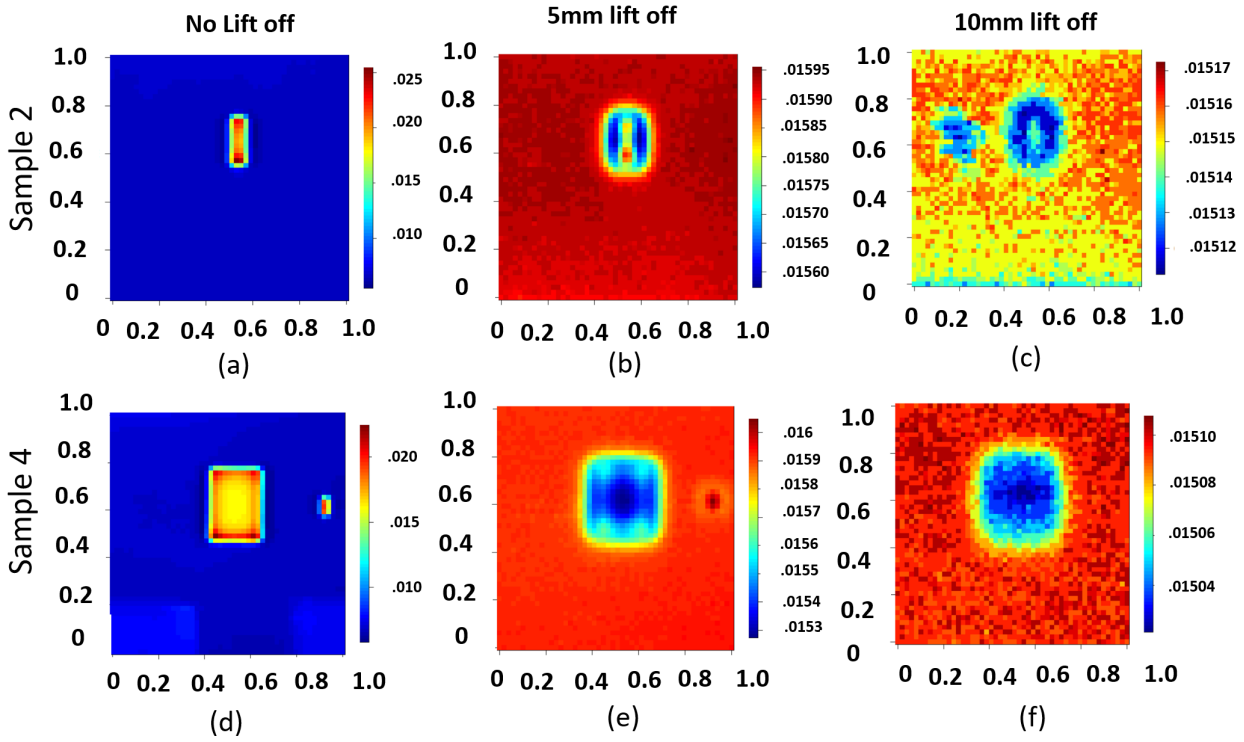

Figure 6.16 Image plot of the impedance readings from three different lift off scenarios of Sample 2 and 4. Here the scanning step size is 2mm.

almost all the Inspections with good coverage and moderately less FDR. Other proposed methods like Mixture regression and wavelet-based thresholding worked well when the defect size gets large. However, when the defect size is small with increase in uncertainties, SNR gets reduced. In these situations, the competing algorithms either fair bad on coverage or have overestimated the defect size by a large extent.

Table 6.7 Dynamic test cases for EC showing the coverage and Overestimation by TLBH method.

	Inspection1		**Inspection2**		**Inspection3**		**Inspection4**	
Case 1	Slit 1 at no lift off		Slit 1 at 3mm lift off		Slit 2 at lift off 3mm		Slit 2 at lift off 5mm	
	Coverage	Over	Coverage	Over	Coverage	Over	Coverage	Over
	1	0	0.86	0.24	0.96	0.04	0.81	0.46
Case 2	Sample 1 at no lift off		Sample 2 at 5mm lift off		Sample 3 at 5mm lift off		Sample 4 at 10 mm lift off	
	Coverage	Over	Coverage	Over	Coverage	Over	Coverage	Over
	1	0	0.92	0.14	0.86	0.42	0.9	0.56

In Table 6.7, dynamic defect tracking by TLBH method is shown on two cases. In case 1, the baseline Inspection 1 is done for Slit 1 at no lift off. As observed in Table 6.6, the SNR of

the baseline inspection is high. This also gets reflected in Table 6 as coverage for the baseline inspection is 1 with no overestimation. On performing Inspection 2 on Slit 1 at 3mm lift off results in significant drop in SNR. Hence there is a need to transfer the information from Inspection 1. Then TLBH is applied on modified Inspection 2 data. This results in better coverage and less overestimation. Next, for inspection 3, Slit 2 is scanned at same lift off of 3mm. The depth of Slit is 2mm and is 1mm more than Slit 1. Performing TLBH after transferring baseline information results in better coverage and estimation than Inspection 2 which is also justified by looking at SNR in Table 4. The coverage and overestimation from Inspection 4 also has high resemblance with the drop in SNR. Overall, we have achieved more than 80% coverage and less than half of exceedance even for very thin submillimeter slit like defects under uncertainties. In case2, we have shown dynamic defect detection in EC from notch defects. There also we have achieved satisfied coverage and FDR rates.

6.7 Conclusion

A robust dynamic defect detection and tracking framework under various uncertainties in NDE data is developed. This framework offers a fast, accurate analysis and can be used in periodic inspection of metallic pipelines and other inspections. Cheap, flexible and miniaturized MFL and EC based NDE sensors are developed for collection of the experimental data. Several dynamic monitoring algorithms are developed that are agnostic of the noise type in the experimental scans and can even perform detection where the underlying noise producing uncertainties vary with time. To understand the noise corruptions and reduction in SNR across various experimental regimes, a mathematical model based on bivariate mixture modelling is developed and hypothesis testing is conducted based on Neyman-Pearson lemma. This mathematical framework shows the inefficiency of pointwise hypothesis testing to operate in noisy regimes and exhibits the tradeoff between FDR and power analysis. In order to properly recognize the defect growth during later noisy periodic scans a novel transfer learning based binned hypothesis testing (TLBH) is developed and validated against both MFL and EC data. The choice of optimal binning parameter is done by transfer learning where the location and size of the defects from detailed baseline inspections are

transferred. The degradations in SNR and the defect sizes of later noisy scans are then estimated and are compared with other developed algorithms such as mixture regression (Non-TL), TL aided mixture regression and wavelet based denoising. The results show that TLBH algorithm is very effective in dynamic detection as it results in high coverage (power) and low over estimation (FDR) across all the experimental scenarios.

CHAPTER 7

MISCELLANEOUS WORKS: APPLICATION OF DIFFERENT MACHINE LEARNING MODELS FOR NDE BASED MATERIAL CHARACTERIZATION AND DEFECT CLASSIFICATION

7.1 NDE based cost-effective detection of obtrusive and coincident defects in pipelines under uncertainties

7.1.1 Introduction

In this subsection of chapter 7, a novel defect detection algorithm based on nearest neighbor based divergence measure in the wavelet transformed domain has been proposed. This algorithm facilitates accurate detection and classification of harmful defects from benign minor scratches under noisy regimes. Defects of different shapes and sizes are considered in this work, particularly we are interested in studying the contrasts between the behaviors of nondestructive evaluation (NDE) metrics from large defects and minor scratches. Large defects may lead to mayhem if not detected in time. On the other hand, the minor scratches and cavities within the pipeline that are harmless at that moment, and just increase maintenance costs when automatically detected and subjected for further extensive scrutiny. Currently, pipeline operators use various inspection methods and analytical programs costing huge sums of public and private funds to prevent, detect, mitigate and inspect for threats. In this context, it will be very helpful if harmless minor defects in the pipelines are actually diagnosed as benign in the very first step and not passed on for expensive higher-level introspections. Another important attribute here would be studying the interaction of multiple defects that might be in close proximity. It might happen that several minor cavities each harmless on its own could affect NDE outcomes much severely than very large defects. Hence, we study scenarios where two or more threats can occur coincidentally and independently of each other. These "coincident threats" might result in a likelihood of failure greater than that due to either threat individually (interacting) or merely the superposition of the threats (co-existing). Interactive threats are the merge of two or more defects in a pipe segment, the result of which is more damaging than either of the individual threat itself.

Here we have constructed defects of seven categories, (1) cuboids of varied lengths, breadth and

height (2) cylinders of varied radii and heights (3) spheres of different radii (4) coincident cube and cylinder as multiple defects (5) coincident cube and sphere (6) coincident sphere and cylinder (7) coincident cube, cylinder and sphere together as multiple defects. A finite element methods (FEM) model, which is implemented on ANSYS developed in our group is adopted to simulate the MFL inspection and generate defects signals. 2D FEM is not sufficient here because the resulting MFL signals are single channel data, whereas the actual signals are multi-channel. For the same MFL geometry and materials, a 3D modelling technique was utilized for more accurate representation of the field distributions in the magnetic circuit due to considerable contribution of the end effect, which is neglected in 2D modelling technique(106). Accurate 3D defects are modeled and resulting MFL signal at the test sample surface is recorded and processed as discussed in previous chapters.

In our MFL FEM simulation, existence of defects will increase the magnetic resistance and distort the magnetic field. The bent magnetic flux, which leaks out of the material surface will form a magnetic leakage field that are thereby measured by the magnetic sensitive sensors. Here, we have considered the magnetic flux in z direction only and stored the magnetic field intensities in the form of images (matrices) so that we can apply wavelet basis as a weighted sum of variations of a much simpler signal. It is shown that the values we obtained from the basis function reveal more features than that of the original MFL signal. Applied adaptive thresholding on the obtained wavelet coefficients is applied in order to denoise them. It is observed that for those defects signals, which exceed our predefined threshold, the variance for them exceeds that of the non-defects and hence a statistical analysis distinguishing the large fissures from benign ones can be done with high accuracy. After applying thresholding on the wavelet coefficients, we apply k-nearest neighbors (K-NN) classifier to distinguish between the non-defects and defects. We observed that defects have larger distance than the maximum of the distance between the non-defects, which serves as a decision criterion for effective classification of any incoming new unknown defect data.

7.1.2 Data Generation

The data used in this section is generated by simulation means using Ansys Maxwell. A 3D MFL model has been developed similar to that used in Chapter 5. Permanent magnets are

the excitation source of the system, as well as the load, and the characteristics of magnets are defined in various material properties. When using ANSYS to treat with permanent magnets, they are translated automatically into equivalent current and apply on every element and node of the model. Corresponding boundary conditions in this model satisfy the following relationship: (a) magnetic scalar potentials (MAG) were used to specify flux-normal (homogeneous Neumann boundary condition), flux-parallel (Dirichlet boundary condition), and far-field zero with MAG set to be 0. (b) The outside air model adopts INFIN47 far-field elements, so they are required to flag the surface of an infinite element, which is pointing towards the open domain.

Here we are observing the perturbation of magnetic flux density vector in z direction, i.e. B_z in the presence of multiple coincident defects at the surface of the pipeline under noisy conditions with different uncertainties. To categorize the defects as non-defects and defects, we investigated multiple defects of various shapes and sizing ranges as shown in Table 7.1.

Table 7.1 Range of defects in training data (units: mm).

Shape	length	width	height	radius
Cuboidal	0.05-20.0	0.05-20.0	0.05 – 8.0	
Cylindrical			0.05 – 8.0	.05 to 7.0
Spherical				.05 to 8.0

Figure 7.1 shows the simulation results illustrating the different nature of the perturbations of magnetic field density in presence of non-obtrusive defects as well as large defects. From figure 7.1, it is clear that in the presence of large defects the perturbation in the magnetic field is larger. However, it is challenging to categorize magnetic field data from unknown pipe sectors as defects or non-defects with higher precision. We present a detection algorithm based on wavelets analysis in the following sub section.

7.1.3 Detection algorithm: Working Principle & Details

A nearest neighbor (NN) based binary classifier (144; 145) to detect between obtrusive and non-obtrusive defects is developed in this paper. The NN classified is constructed on coefficients from multi-resolution wavelet basis (146; 147). Treating the magnetic field intensity from each pipe sector as a bivariate function the corresponding wavelet representation is first computed.

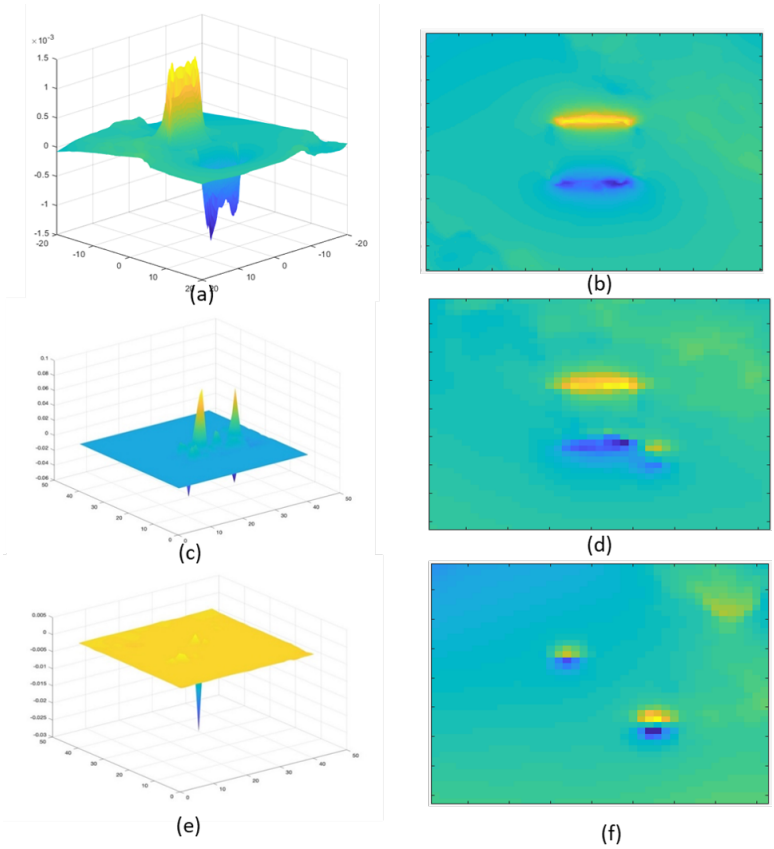

Figure 7.1 From top-left, (a) shows the 3D fluctuation of the magnetic field intensity along the z-axis in the presence of a single obtrusive defect; (b) on its left shows the 2D fluctuation of the same magnetic field as we move spatially along the x- and y-axis. This bivariate representation of the magnetic field is subjected to wavelet based detection analysis; (c) and (d) respectively demonstrate the 3D and 2D perturbation of the magnetic field when there are two large interacting defects in the zone; whereas (e) and (f) in the bottom row illustrate the magnetic field where there are no large defects but a few minor cavities or scratches.

Thereafter, the wavelet coefficients are adaptively estimated with high precision. Next, the distance between the wavelet coefficients for the defects and the non-defects are computed and a cut-off distance for detecting the defects from the non-defects is tabulated. Wavelets being both frequency as well as spatially localized are better equipped to model discontinuities in the magnetic fields due to emergence of spikes in zones with defects, cavities and miscellaneous irregularities [14]. Figure 7.2 shows as the NDE recordings are in advertently affected with system noises and measurement aberrations, wavelet analysis helps in de-noising these readings.

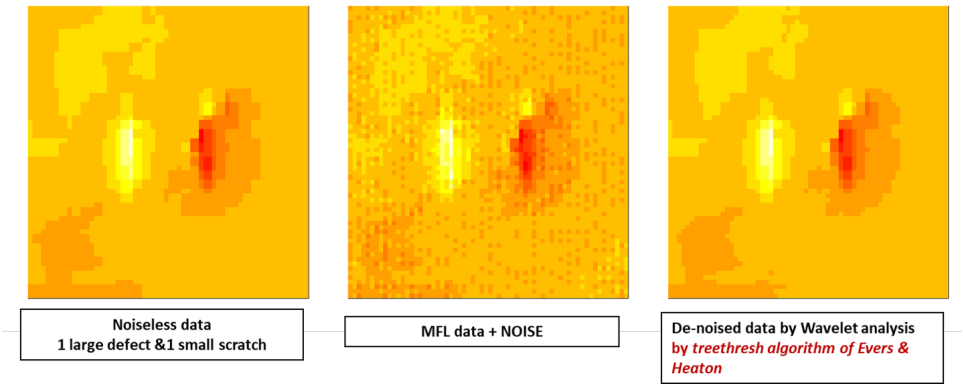

Figure 7.2 Left shows the heat map of the flux distribution of the noiseless MFL data containing a defect, middle figure shows the data infected with noise, right figure shows the denoised data by wavelet analysis by treethresh algorithm of Evres & Heaton.

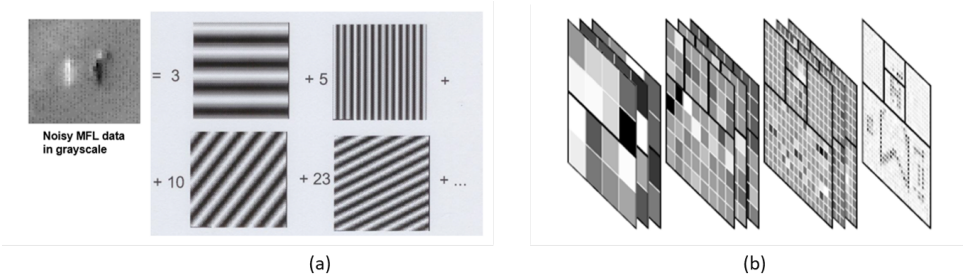

Figure 7.3 (a) Schematic showing wavelet basis of different levels, (b) The basis used here is "DaubLeAsymm", Daubechies' least-asymmetric wavelets.

7.1.3.1 Feature Extraction and Heuristic Working Principle.

We first compute two-dimensional Discrete Wavelet Transform (DWT) based on Mallat's pyramidal algorithm (147). Let I_i be the true signal intensity from the i^{th} experiment involving pipe sectors. It can be well-approximated by wavelet coefficient up-to level L.

$$I_i \approx \sum_{l=1}^{L} \sum_{j=1}^{n_l} c_{ij}^{l} \Psi_j^l \tag{7.1}$$

where, $\Psi_j^l = 1, \cdots, n_l$ is the two dimensional wavelet basis of level l and c_{ij}^l is the corresponding coefficient based on the expansion I_i.Let $c_i^l = c_{ij}^l : j = 1, \cdots, n_l$ represent the vector of the coefficients at level l for the i^{th} experiment. Let DC be the set of all experiments where there are only minor cavities but no harmful defects and D be the set of all experiment where there are harmful large effects.

The heuristic working principle depends on the fact that the higher resolution wavelet coefficients

containing lower energy will be relatively sparse and substantially different between members in D and D^C. Unlike members in D^C, multiple sharp peaks are present in the magnetic intensity field sampled in experiments with large defects and those will be captured in the finer resolution basis coefficients. This separation in the higher resolution coefficients is the heuristic rationale for our nearest neighbor algorithm to work successfully. For each i, define the distance for the nearest neighbor in the non-defective population D^C by:

$$h_i^l = \min_{j \neq i, j \in D^C} distance(c_i^l, c_j^l) \tag{7.2}$$

where, the *distance* operator calculates the Euclidean distance between the basis vectors. For an intensity field sampled from an unknown experiment k, we compute h_i based on its coefficients and the coefficient in D_C. For any fixed level l, we select a cut-off value T. If $h_k < T$, then we classify that the unknown k^{th} experiment is from a pipe sector that does not contain any harmful defects; otherwise, we classify that the pipe sector has detrimental defects and needs to be immediately taken care of. In the following section we will describe our choice of T.

7.1.3.2 Adaptive Estimation of the distance measures

In realistic NDE settings, we do not actually observe the true signal intensity I_i but observe a noise contaminated version of I_i, which we represent by $\hat{I}_i$. Thus, we do not observe the basis coefficient c_i^l directly and need to estimate them by $\hat{c}_i^l$. It is shown that direct usage of DWT usually results in wrong estimate, and as such it has been established in (146) that using a truncated version of the DWT outputs $\hat{c}_{ij}^l$

$$\hat{b}_{ij}^l = sign(\hat{c}_{ij}^l)(|\hat{c}_{ij}^l| - \hat{\sigma}_i^l \times \sqrt{2 \log n_l}); j = 1, \cdots, n_l \tag{7.3}$$

where, $\hat{\sigma}_i^l$ is the standard deviation of wavelet coefficients. Usually, the wavelet coefficients of signals will be sparser at the fine resolution scales, and denser at the coarser scales. Thus, efficient estimation methods need to moderate the truncation magnitude level by level. Also, instead of having a fixed threshold size for truncation as shown in equation 7.1, threshold sizes adaptively chosen based on the signal peculiarities have been witnessed to produce better estimation performance. As such (148; 149; 150) establish that adaptive thresholds far outperform the $\sqrt{2 \log n_l}$

threshold size described in equation 7.1. Here, we use the adaptive estimation methodology of (148) that can efficiently handle the cases where the underlying magnetic intensity is not homogeneous throughout but instead has distinct regions with different strength characteristics. For estimation of the wavelet coefficients, we first identify these different regions and perform separate estimation in each by adaptively adjusting the threshold sizes according to the local signal strength. These creates efficient locally adaptive estimates of c_i^l. We next compute nearest neighbor estimates c_i^l by applying a scalable and fast algorithm on the adaptive coefficient estimates (151; 152; 150) and obtain near neighbor distance estimates within a specified error bound. Based on these distances, we derive precise estimates $\hat{h}_i^l$ of the distances in equation 7.2.Figure 7.4 shows the KNN based clustering schematic where the red dots are the representation of coefficients of Non-defective samples whereas the blue dots represent that of defective ones.

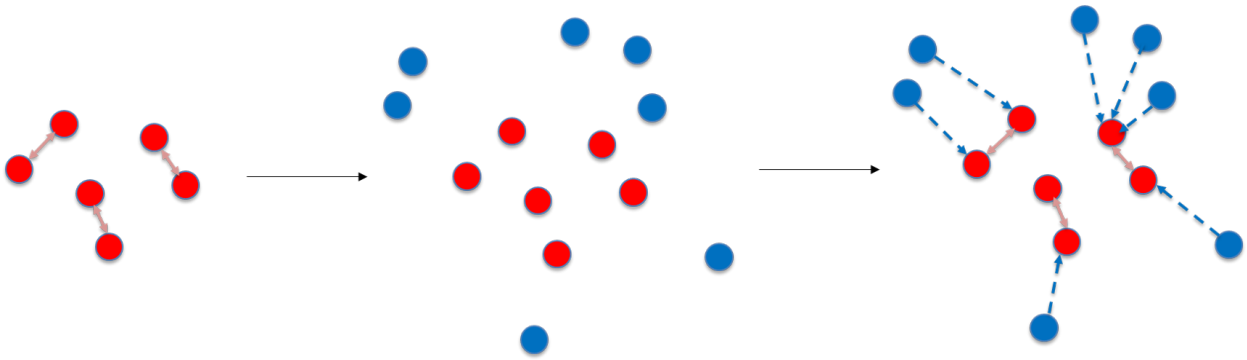

Figure 7.4 Schematic shows the clustering working principle where we calculate the nearest neighbor distance for each point in non-defective set (red) in left & calculate the nearest non-defective neighbor distance for each point in defective set in right.

7.1.3.3 Results & Discussions

Depending on the choice of the threshold, we get different success rates in our two-fold objectives: (a) minimize the classification of harmless defects as obstructive ones, and (b) maximize the correct classification proportion of real obstructive defects. The latter dealing with correct detection of defects is called sensitivity while the former leads to false alarms for detection by wrongly classifying harmless defects and thus increasing maintenance costs. Using higher values of the cut-off T defined above will give us more sensitivity but also increase the proportion of false alarms.

In this case study, we test our proposed detection methodology on the dataset of section II. We classify any experiment with at least a defect of length greater than 5 mm as obtrusive. There are 36 obtrusive and 69 non-obstructive defects associated magnetic intensity fields. We consider distances based on wavelets of level 6 only, i.e., $l = 6$. Based on the grid size used in the data collection step of this case study, this was the finest resolution that was obtainable. Coarser resolutions were substantially denser.

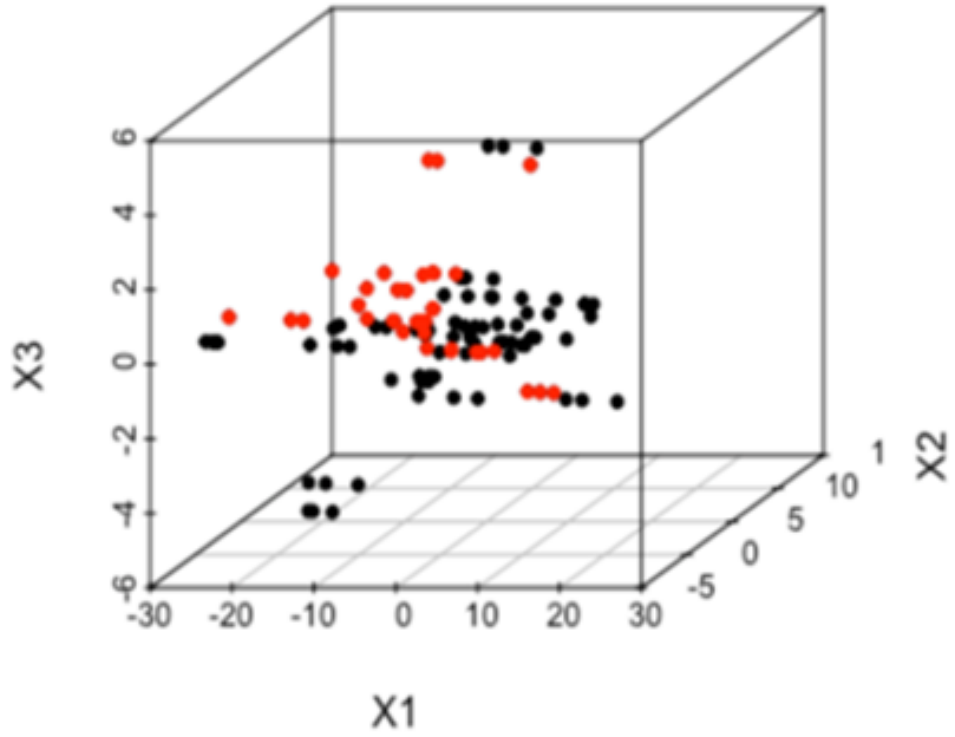

Figure 7.5 Wavelet coefficients at level 6 in the reduced three-dimensional space.

In Figure 7.5, in red we have obtrusive defects and in black non-obtrusive defects. Figure 7.5 shows that there is visible separation in the wavelet coefficients between the obtrusive and non-obtrusive defects even when the high-dimensional wavelets are reduced to the perceptible three dimensions. Figure 7.6 shows that there is also palpable variation in nearest neighbor distances between the two sets with obtrusive having greater distances from the non-obtrusives in comparison to the distances between non-obtrusive.

Here we simultaneously report the percentage of false alarms in incorrect classification of non-obtrusive defects as well as the percentage of correct classification of obtrusive defects (sensitivity).

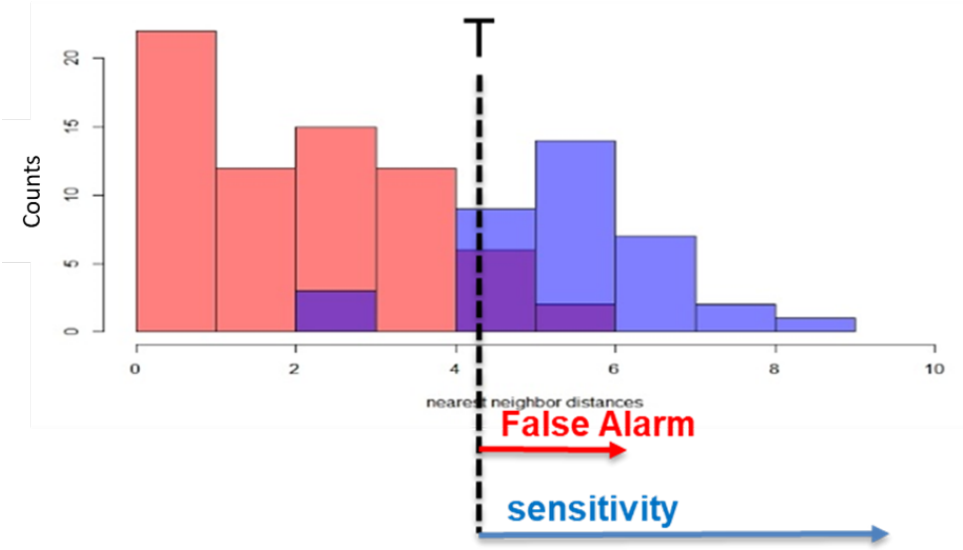

Figure 7.6 Histogram of the Euclidean distance of the wavelet coefficients of a bivariate magnetic flux from the nearest non-obtrusive flux field. In blue, we have distance of obtrusive fluxes and in red we have distances among non-obtrusive fluxes.

Table 7.2 Report of the working characteristics of the detection method.

False Alarm: Train	5.8	10.1	18.8	29	34.8	39.1
False Alarm: Test	40	44.8	46.7	47.6	48.6	48.6
Sensitivity	88.9	91.7	91.7	94.4	97.2	100

Table 7.2 provides us with the percentage of correct detection of large defects when we apply our methodology to the aforementioned dataset. We consider 174 samples with 50% of the data used for training and the rest for testing. If we allow the false alarm rate to be around 40% then we get near perfect detection of obtrusive defects in the training sample. However, the corresponding testing false alarm rate needs to be slightly increased to 48% for getting perfect detection in this study. Figure 7.6 shows the receiver operating characteristic (ROC) curve of Table 7.2. A sharp increase in power is witnessed as the false alarm test rate is allowed to increase above 35%. As we need to prevent the misclassification of obstructive defects, which may lead to mayhem, we suggest having a false alarm rate of 50%, which still benefits us in saving half the costs over competing vanilla methods that does not differentiate between minor and harmful defects.

In Figure 7.7, the ROC curve represents the increase in the proportion of detection of obtrusive

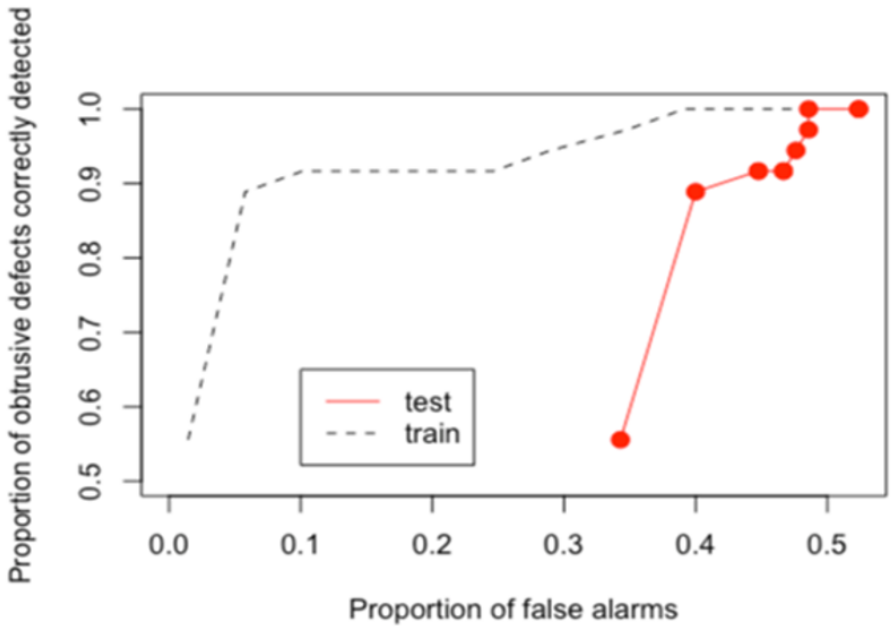

Figure 7.7 Plot of Sensitivity vs false alarm.

defects (sensitivity) as the proportion of false alarms for incorrect classification of non-obtrusive magnetic intensity fields are allowed to increase. These percentages are reported in Table **??**. In red, we have the test errors and the dotted black line represents the corresponding training sensitivities.

7.1.4 Conclusion

We present a wavelet based nearest neighbor algorithm that can detect large harmful defects in pipelines with very high accuracy. Our proposed methodology can also simultaneously distinguish minor harmless defects and cavities. Our proposed method will save expenditure over competing NDE methods that can-not distinguish between defect types and would recommend these minor defects for further scrutiny along with the large defects. *Portions of this section appeared in the publication (28)*

7.2 Accurate Material Characterization of Wideband RF Signals via Registration based Curve Fitting Model using Microstrip Transmission Line

7.2.1 Introduction

Solving inverse problems is commonplace in the field of Nondestructive Evaluation (NDE).Recent advancements in machine learning (ML) have provided us with a large bag of analytical tools for

solving inverse problems (29; 12; 51). However, most potent ML models require a lot of data for training and estimating large number of model parameters. The problem with application of such methods in NDE is the dearth in the volume of real experimental data as collection of experimental data is both costly and time consuming (99). This is particularly true for the NDE problem of radio frequency (RF) based material characterization as acquiring the data from fabricated devices with controlled substrate properties and shape is very difficult and laborious. To mitigate these challenges faced by ML methods in RF material characterization we consider a registration based data augmentation framework by which data driven predictive models can leverage the information stored in easy to generate large simulated data-sets and increase their predictive efficacy in experimental data.

We consider the NDE problem of Radio frequency (RF) based material characterization where the goal is to estimate the electrical parameters such as permittivity, permeability and respective loss tangents, with an input of physical parameters of material and RF structure, and its frequency response. Conventional applications of RF based material characterization make use of idealized frameworks where there exist electromagnetic (EM) physics based theoretical solutions. Microstrip line and coplanar strip lines are two examples, which have the EM physics-based mathematics defined for standard geometries and we can back-calculate the electrical properties based on those theoretical relations (153; 154). This work is an attempt to increase the scope of RF methods beyond such idealized set-ups that need exact mathematical solutions. The goal is to use data-driven machine learning models to characterize a material with any complex shape (not limited to single layer planar sheet) and complex RF structure (not limited to transmission lines). For viability, we test the efficacy of the proposed approach against widely studied microstrip transmission line methods. We list our contributions below:

1. We generate vast amount of simulated data through finite element methods. In this simulation data, the electrical parameters are precisely defined with no noise. We use various data-driven ML models to train on this simulated data and report their performance. Several recent works (155; 156; 157; 158) have considered material characterization using machine

learning in micro strip transmission lines and microwave antennas. But, in the existing literature, the sources of the generated data set are similar to the source of the simulation means. We, on the other hand focus on applying these models trained on large simulated datasets to noisy experimental datasets. We consider observing the frequency response across different physical parameters of material and RF structure. The objective is to estimate the electrical parameters of the unknown material substrates by measuring the frequency responses using Vector Network Analyzer (VNA). For microstrip lines embedded on unknown substrate we want to estimate the electrical parameters of the substrate from the frequency response.

2. We demonstrate that several machine learning model produces good fit on the simulated data as these data sets are large and noiseless. However, the coefficients learned by the model from the trained simulated data can not be directly used to explain the variation in the electrical parameter in experimental data. As the data generation methods are different, for materials with similar electrical characteristics there can be severe misalignment between the simulated frequency responses and the observed experimental responses (see Figure 7.20). This highly limits the usability of simulated data to predict responses in the experimental data. To mitigate this challenge, we propose to use a small fraction of the experimental data for calibrating ML models trained on simulated data. Using very few experimental samples as calibration set, we employ a very simple registration technique based on intercept correction. Compared to ML models trained on simulated data only, the proposed registration method greatly improves the predictive efficacy of the ML algorithms on the experimental data. A great advantage of using this simple registration scheme is that we need very few experimental samples and so the proposed approach can be employed in data scare experimental regimes.

3. The observed data are measured reflection and transmission coefficients. They are smooth functions of frequencies. We first arranged our collected data in a structured format such that we have scalar responses along with both scalar and functional covariates. The response which is the permittivity is scalar, the covariates such as height of the substrate, length and width of the microstrip line are scalars whereas the transmission and reflection coefficients

are smooth functions of the frequencies. The simulated and experimental data can be based on different set of frequencies. To address this issue, we first fit a curve based on local linear regression to each of the observed functional S parameters. Based on estimates from the fitted curves we construct a data set with S parameters from a common set of frequencies. See Figure 7.13 and its associated text in Subsection 7.2.4.

4. As we have very few experimental samples, the experimental data has more features in columns (which are the observed *S* parameters at different frequency values) than the number of records in rows. This happens as the number of frequencies in the parametric sweeps is much larger than the different number of substrates that we can experiment in the lab. For this purpose, we consider dimensionality reduction techniques on the observed feature space of the *S* parameters. We consider here dominant principal component (PC) direction based reduction technique. Different regression models on these PCs are then applied on the simulated and experimental data sets. Both parametric and non-parametric regression methods such as random forest regression are implemented. For the simulated data-sets, complex higher dimensional model are also implemented and their performance is compared to the PCs based regression methods. For experimental data sets, due to data scarcity, only PC based models are implemented. Two registration framework called intercept calibration and registration by data augmentation is developed in Sec. IV to calibrate the PCs based regression models on the experimental data. We show how knowledge of a tiny portion of experimental data helps in transferring a significant amount of information present in the simulation data set for predictive analysis of the experimental data. Figure 7.8 shows the data analysis framework used in this study. We found that the RF regression based model on the PCs produces the least test error after calibration and yields considerable benefits for wideband RF material characterization on the experimental data.

This work is organized as follows: In section 7.2.2 we discussed the simulation and experimental data generation framework. In Section 7.2.4, we describe the data structure after curve smoothing, different feature selection models conducted in this study. It also contains the study of different

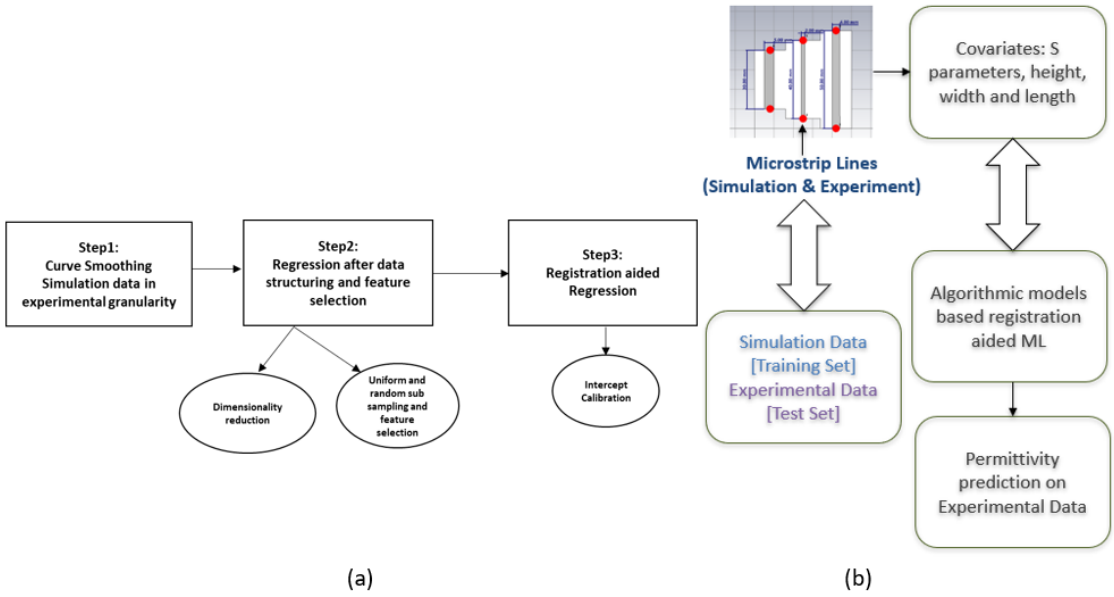

Figure 7.8 Schematic of Data Analysis framework conducted in this study.

regression methods on the simulated data set. Section 7.2.5 shows the impact of registration on better prediction of the permeability in experimental data.

7.2.2 Micro strip Transmission line Simulation & Experimental Setup

Standard microstrip transmission lines with length L and width W are simulated on RF substrates having dielectric permittivity ϵ_r, magnetic permeability μ_r, dielectric loss $\tan\delta_e$ and magnetic loss $\tan\delta_u$. Keysight's Advanced Design System (ADS) is used to compute the frequency response of the microstrip line from $100MHz$ to $12GHz$ and the range of all the parameters is shown in table 7.3.

Table 7.3 Range of the parameters used to simulate different combinations for training data set.

Parameter Name (unit)	Minimum	Step Size	Maximum
L (mm)	1	5	100
W (mm)	0.1	0.5	10
ϵ_r	1.1	0.5	10.1
H (mm)	1.27	-	-
μ_r	1	-	-
$\tan\delta_e$	0.01	-	-
$\tan\delta_u$	0.01	-	-

The setup done in ADS for microstrip line simulation is straight forward, it requires a definition

of MSUB (microstrip substrate), MLIN (microstrip line), S parameters (S-parameter simulator) and two 50-ohm ports as shown in figure 7.10(a). The microstrip substrate requires details on permittivity, permeability, dielectric loss tangent, substrate and conductor thickness and metal conductivity. The schematic of the synthetic microstrip lines with dimensions is illustrated in figure 7.9 The microstrip line frequency response simulation gets completed in 0.45 seconds for a single combination of parameters. Figure 7.10(b) shows the frequency response on a Roger4350 board.

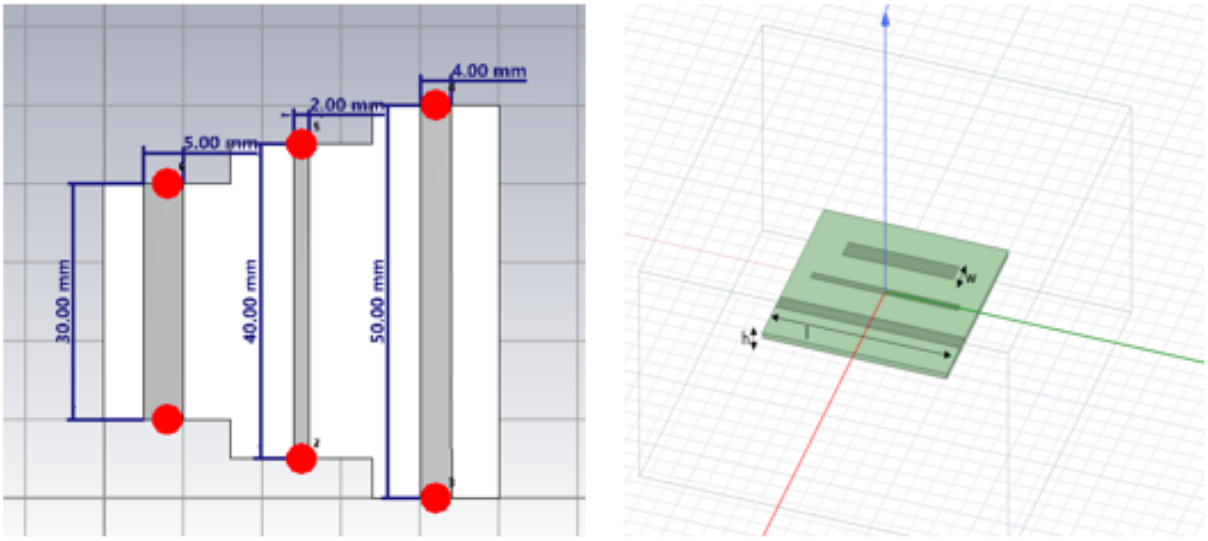

Figure 7.9 (a) Schematic of synthetic micro strip line, (b) dimensionsal representation of the microstrip lines.

For the experimental test data, 11 different Rogers boards such as Rogers 3039730, Rogers 3107150, Rogers 3039851, Rogers 3008049, Rogers 3032153, Rogers 3028557, Rogers 3068926, Rogers 2007685, Rogers 3040756, Rogers 3029210, Rogers 3033859 are considered. On each substrate, three different transmission lines of varying width and lengths such as 2mm width 40 mm length, 5 mm width 30 mm length, 4 mm width 50 mm length are fabricated. The dimensions of 33 different microstrip lines fabricated on 11 different Rogers boards are shown in figure 7.11. Reflection and transmission coefficients (S parameters) are measured for each transmission line by

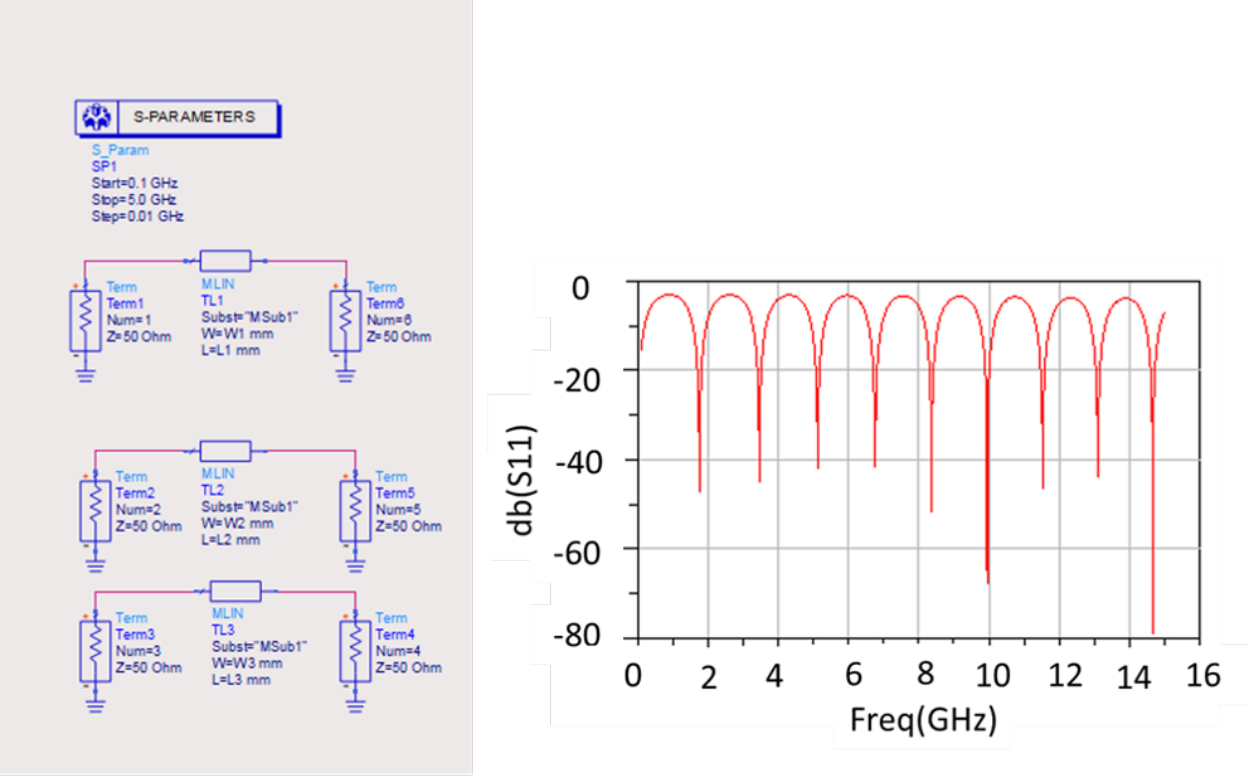

Figure 7.10 (a) ADS setup for microstrip line, (b) Simulated frequency response on RO4350 board.

connecting to E8362B PNA Series Vector Network Analyzer manufactured by Agilent Technologies in a broad frequency range of 100 MHz to 12 GHz. The electrical properties and dimensions of each substrate considered for experiments are reported in table 7.2.

Table 7.4 Different substrates considered in experiments DT: dielectric thickness and CT: copper thickness, DK: dielectric constant.

	Substrate	**Dimensions (Units DT: mm, CT:um)**
1	Rogers 3039730(R03010 DK =10.2)	DT=1.27mm CT=17.5um
2	Rogers 3107150 (RT/ droid 5870 DK =2.33)	DT=1.57mm CT=17.5um
3	Rogers 3039851 (R03006 DK=6.15)	DT=1.27mm CT=17.5um
4	Rogers 3008049 (R04 350B DK =3.33)	DT= 0.1mm CT=17.5um
5	Rogers 3032153(RT/droid 6010LM(DK=10.2)	DT=01.27mm CT=35um
6	Rogers 3028557 (RT/duroid 5880 DK=2.2)	DT =1.57mm CT=17.5um
7	Rogers 3068926 (TMM10 DK=9.20)	DT=1.27mm CT=35um
8	Rogers 2007685 (R04350B DK=3.48)	DT=1.52mm CT=17.5um
9	Rogers 3040756 (RT/duroid 6006 DK=6.15)	DT=1.27mm CT=35um
10	Rogers 3029210(RT/duroid 6010LM DK=10.2)	DT=1.27mm CT=17.5um
11	Rogers 3033859 (R04835 DK=3.48)	DT=1.52mm CT=17.5um

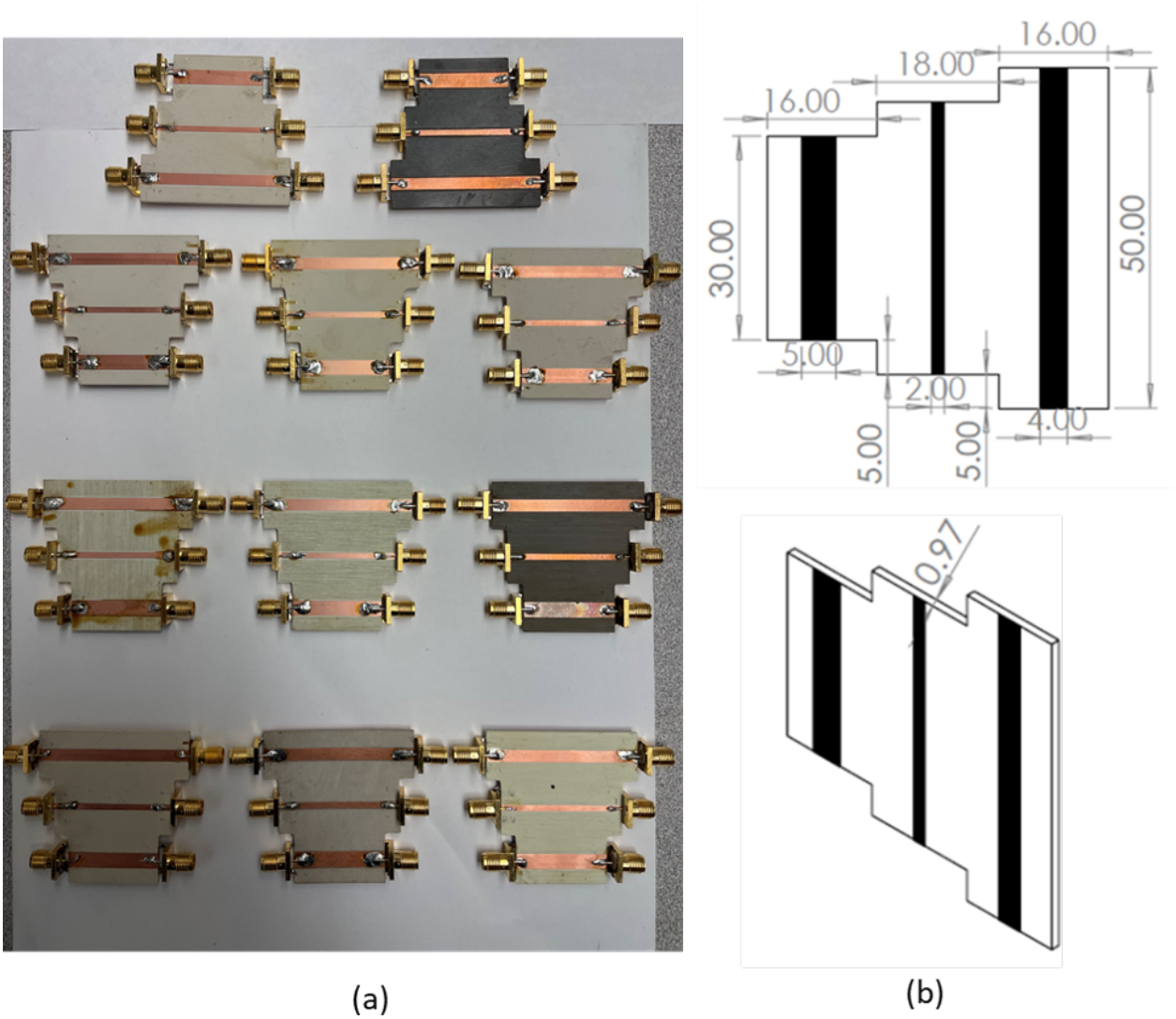

Figure 7.11 (a) 33 microstrip lines fabricated on 11 different substrates for experimental test set, (b) dimensions of the microstrip lines.

7.2.3 Material Characterization using Physics-based Model

The conventional Physics based model takes into consideration mathematical equations for electrical parameter retrieval of substrate from the measured S parameters. This exact mathematical formulae based parameter estimation method is discussed in the section which requires the S-parameter, length (L), width (W) of transmission line and height (H) of substrate. Figure 7.12 shows the E and H field distribution of the microstrip line (153). The electromagnetic field in a microstrip line are not only passing through the substrate, but it has an effective permittivity $\epsilon_{\mathtt{eff}}$ and permeability $\mu_{\mathtt{eff}}$ (154). The propagation constant for the microstrip line is as shown in

equation 7.9.

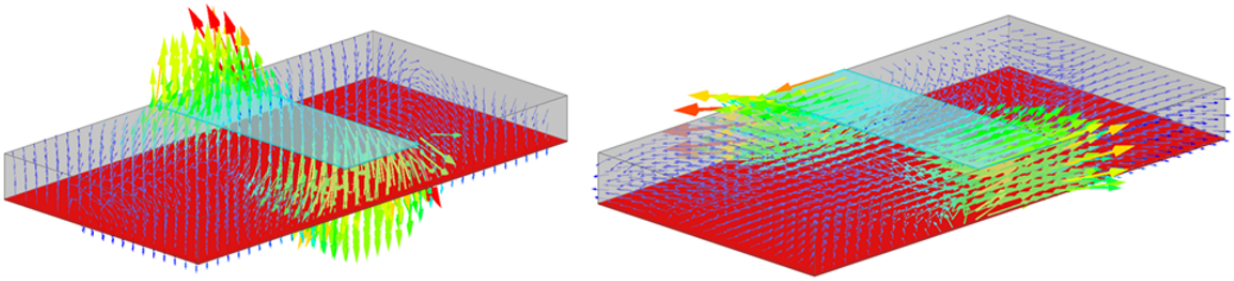

Figure 7.12 E and H-field distribution of microstrip line.

$$\gamma = j\beta = jk_0(\epsilon_{\mathtt{eff}}\mu_{\mathtt{eff}})^{1/2}, \text{ where, } k_0 = 2\pi c^{-1} f \ . \tag{7.4}$$

The phase constant can be estimated from the measured frequency response by converting the measured S-parameters to ABCD-parameters as done in equation 7.5:

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \text{S to } \overrightarrow{\text{ABCD}} \begin{bmatrix} \cos\beta l & jZ_0 \sin\beta l \\ jY_0 \sin\beta l & \cos\beta l \end{bmatrix},$$

$$\beta = L^{-1} \arccos\left((1+S_{11})(1-S_{22}) + (S_{12}S_{21})\right) 2{S_{21}}^{-1} \tag{7.5}$$

The characteristic impedance Z_0 of the microstrip transmission line is given by equation (7.3):

$$Z_0 = \{(1+S_{11})(1-S_{22}) + S_{12}S_{21}\} \cdot (2S_{21}\sin\beta l)^{-1} \tag{7.6}$$

The relation of characteristic impedance Z_0 with effective permittivity $\epsilon_{\mathtt{eff}}$ and permeability $\mu_{\mathtt{eff}}$ are shown for below two conditions: (a) when $W/H < 1$:

$$Z_0 = ({\mu_0\epsilon_0}^{-1}{\mu_{\mathtt{eff}}\epsilon_{\mathtt{eff}}}^{-1})^{1/2}(2\pi)^{-1}\ln(8HW^{-1} + W(4H)^{-1})$$

$$2\epsilon_{\mathtt{eff}} = \epsilon_r + 1 + (\epsilon_r - 1)[(1+12H/W)^{1/2} + 0.04(1-W/H)^2]$$

$$2\mu_{\mathtt{eff}} = \frac{1}{\mu_r} + 1 + \left(\frac{1}{\mu_r} - 1\right)[(1+12H/W)^{1/2} + 0.04(1-W/H)^2]^{-1} \tag{7.7}$$

(b) when $W/H > 1$, we have:

$$Z_0 = \frac{\sqrt{\mu_0{\epsilon_0}^{-1}\mu_{\mathtt{eff}}{\epsilon_{\mathtt{eff}}}^{-1}}}{1.393 + W * H^{-1} + 2 * 3^{-1}\ln(W * H^{-1} + 1.444)}$$

$$2\epsilon_{\mathtt{eff}} = (\epsilon_r + 1) + (\epsilon_r - 1)(1+12H/W)^{1/2}$$

$$2\mu_{\mathtt{eff}} = {\mu_r}^{-1} + 1 + ({\mu_r}^{-1} - 1)(1+12H/W)^{-1/2}. \tag{7.8}$$

Finally by rearranging the above equations (7.7), (7.8) we get the values of relative permeability and relative permittivity of substrate as: when $W * H^{-1} < 1$:

$$\epsilon_r = \frac{(1 + 12H/W)^{-1/2} + 0.04(1 - W/H)^2 - 1 + 2\epsilon_{\texttt{eff}}}{(1 + 12H/W)^{-1/2} + 0.04(1 - W/H)^2 + 1},$$
$$\mu_r = \frac{(1 + 12H/W)^{-1/2} + 0.04(1 - W/H)^2 + 1}{(1 + 12H/W)^{-1/2} + 0.04(1 - W/H)^2 - 1 + 2/\mu_{eff}}$$

and, when $W/H^{-1} > 1$, we have:

$$\epsilon_r = \frac{(1 + 12H/W)^{-1/2} - 1 + 2\epsilon_{\texttt{eff}}}{(1 + 12H/W)^{-1/2} + 1}$$
$$\mu_r = \frac{(1 + 12H/W)^{-1/2} + 1}{(1 + 12H/W)^{-1/2} - 1 + 2/\mu_{\texttt{eff}}}.$$

7.2.4 Data-driven Predictive Models

The objective of this work is to predict the dielectric constants of the experimental substrates where the covariates are (a) the S parameters which are smooth function of frequencies (b) height and width of the lines. The machine learning model is trained on the simulation data. Functional data analysis (FDA) is one of the fastest growing areas in the field of data science where the observations are functions defined on some continuous domain and the observed data are a sample of functions taken from discrete grid of population size (159).In this work we have considered the real (Re) of the reflection coefficients $Re(S_{11})$, imaginary (Im) of the reflection coefficients $Im(S_{11})$ and the magnitude of the reflection coefficient in db scale (dB) as covariates which are a function of frequencies. Similarly, the same components are selected from the transmission coefficients S_{21}. Since it is a symmetric framework hence $S_{11} = S_{22}$ and $S_{21} = S_{12}$. Hence we have selected the S_{11} and S_{21} parameters. The other scalar covariates serving as metadata are the height of substrate (H), width (W), and length (L) of the transmission lines respectively. The dependent variable or the response is the dielectric constants which are scalars and needs to be predicted. Hence this forms a functional regression (FR) which consists of scalar responses with functional and scalar covariates (160).

Curve Fitting using Loess. The S parameters collected from simulations and experiments can be different. As we are considering the parameters as a function of frequencies hence the

frequencies in both the training and test set have to be the same. For that, the simulated components of *S* parameters are smoothed first by locally estimated scatterplot smoothing (loess) (161; 162). Local polynomial regression is fitted locally on the covariates (real, imaginary, and magnitude of the S parameters) against the frequencies. Then this fit is applied on the test frequency set to make the training simulated data to be in the same granularity as the experimental test data. Figure 7.13 shows one such covariate, the real component of the reflection coefficient of simulated data on a Rogers 3039730 board at different sub sampling rate fitted against the experimental test frequencies. The figure shows that the simulated curve has been successfully smoothed and can retain the original shape even on choosing lesser number of sampling points.

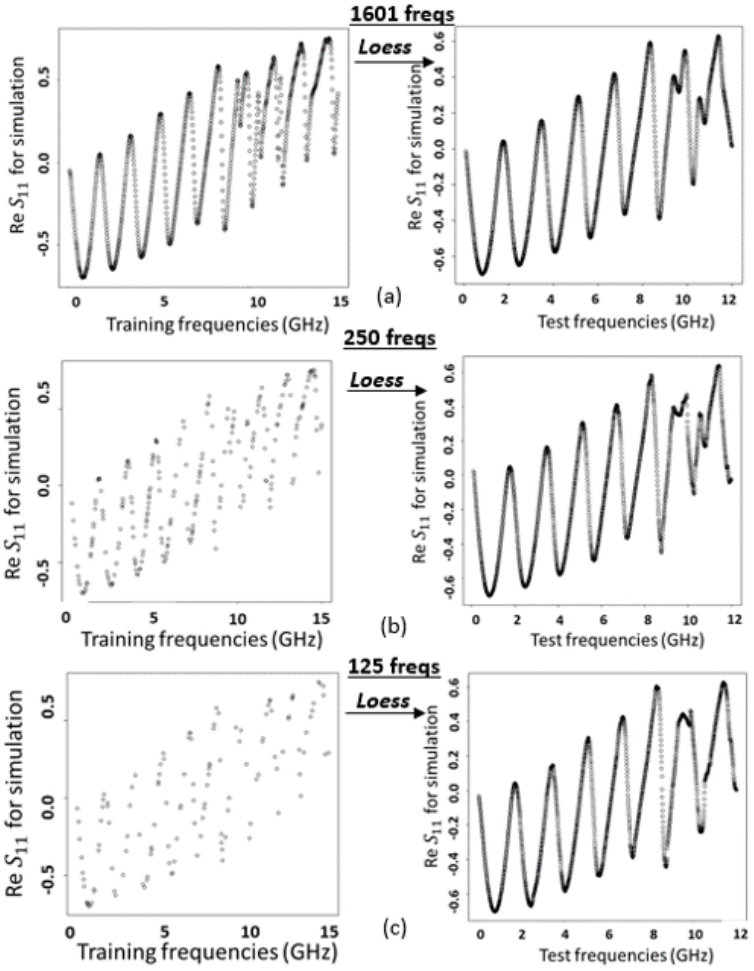

Figure 7.13 Reduced data structure after converting all data in the same granularity by curve fitting.

The data are then structured in a fashion as shown in figure 7.14 with the real component of the

S_{11} stacked across 1601 frequencies, then the imaginary and the magnitude of the S_{11} are stacked. Same we have performed for the S_{21} parameters. The last three columns are the H, W,L respectively and Y is the dependent variable.

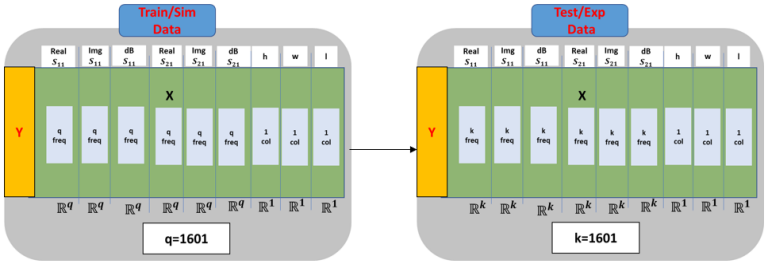

Figure 7.14 Data Structure considered in the paper.

Dimension Reduction. In the data, the number of observations (p) $1601 * 6 + 3 = 9609$ are much greater than the number of samples (n) ($p >> n$). For test data, there are $11 * 3 = 33$ rows for 11 different substrates where each substrate has 3 different microstrip lines embedded on it. Considering all the columns will make the model highly unstable. We have developed dimensionality reduction technique and feature selection for our dataset.

While performing principal component based regression, we are reducing the columns from $q = 1601$ for each feature component to first 3 principal components. These are the linear combination of columns that maximize the variance. The choice of principal component is a tuning parameter and first 3 principal components are chosen as it gives the least training error of 1.1 root mean square error (RMSE) where RMSE = $\sqrt{\text{Average}(\hat{y} - y)^2}$ and y is the dependent variable or the response and $\hat{y}$ the predicted value. Table 7.5 shows that $k = 3$ gives us the minimum training error as the training error does not decreases with increase in the value of k. We can not increase the value of k beyond 5 as that will result in more number of columns than rows in our reduced experimental test set. The training error is evaluated by dividing the simulated training set in the ratio of 3:1 into train and validation and run for several iterations. 1601 columns of each component are reduced to 3 principal components. This is done separately for all the six set of features (a) real of reflection coefficient, (b) imaginary of reflection coefficient, (c) magnitude (dB) of reflection coefficient, (d) real of transmission coefficient, (e) imaginary of transmission

coefficient, (f) magnitude (db) of transmission coefficient.

$$P_{n*3} = \sum_{j=1}^{1601} w_j c_j \text{such that} \sum w_j^2 = 1$$

Then a linear regression model is fitted on the simulated training data as used in equation 3.5

$$Y_i = \alpha + \sum_{j=1}^{3} \beta_1 R_{ij} + \sum_{j=1}^{3} \beta_2 I_{ij} + \sum_{j=1}^{3} \beta_3 dB_{ij} + \beta_4 W_i + \beta_5 L_i \tag{7.9}$$

where R, I, and dB represent the real, imaginary and magnitude of S parameters.

Figure 7.15 illustrates the data format after reduction by dominant principal component.

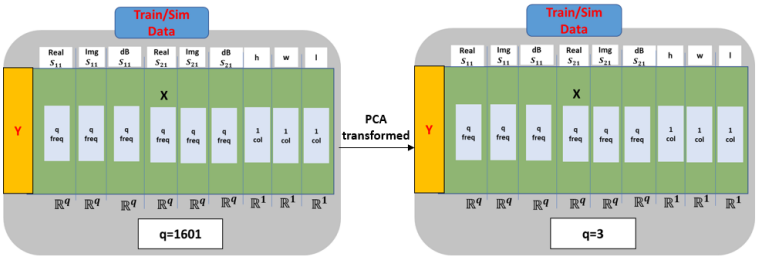

Figure 7.15 Schematic of data format on reduced by dominant principal component.

Table 7.5 RMSE of the training error when applying linear regression on different choices of principal components.

	Number of Principal Components (k)	Training Error
1	1	2.24
2	2	1.50
3	3	1.104
4	4	1.101
5	5	1.09

Dimensionality reduction is performed in the same manner on experimental data.

$$P_{n*3} = \sum_{j=1}^{1601} w_j b_j \text{such that} \sum u_j^2 = 1$$

The regression coefficients obtained from fitting the training data are used to predict the reduced test matrix. The columns are reduced from 9609 to $3 * 6 + 2 = 20$.

Grouping by Principal components enable better visualization as it enables lesser intra class distance and larger inter class distance among the different response categories. Figure 7.16 shows that when plotting the different S components against the frequencies, they can not be discriminated

based on the response variable. Plotting the response variable against first two principal components (mid plots in column a and b) shows the features to be easily discriminated based on 10 categories of the response variable. For better visualization we have binned the response variable in 4 groups of 1.1 to 3.1, 4.1 to 6.1,7.1 to 9.1 and 10.1 (bottom plots in column a and b). Figure 7.16 illustrates that second principal component bears more weightage for magnitude of the reflection coefficient S_{11} and first principal component is more significant for magnitude of the transmission coefficient S_{21}.

Plotting the response of the training data against the first two principal components for the real and imaginary S paramaters in figure 7.17 show the imaginary components to be least significant among magnitude, real and imaginary components of the S parameters.

The principal component takes the weightage of all the features. For this reason we implemented lasso based shrinkage model to observe the important features. Lasso minimizes the residual sum of squares subject to the sum of the absolute value of regression coefficients being less than a constant. It produces a regularization path over a grid of values for the tuning penalty parameter λ. Lasso is implemented using glmnet package (163) where different penalty parameters are used to obtain model with varying number of features. The best value of $\lambda = 0.128$ is obtained by applying k-fold cross validation on the training set and the training error is reported for different multiples of the best λ. Table 7.6 shows the training error obtained for different values of λ and the number of features chosen in each scenario. For $\lambda = 0.128$, the least training error is obtained when 1571 out of 9609 features are selected by lasso model. However, PCA obtained this training error only with 20 feature combinations.

Table 7.6 RMSE of the training set and the number of features chosen by lasso model.

	RMSE	Number of features
1	0.964	1571
2	1.1	445
3	1.03	616
4	1.27	345
5	2.08	73
6	2.45	44

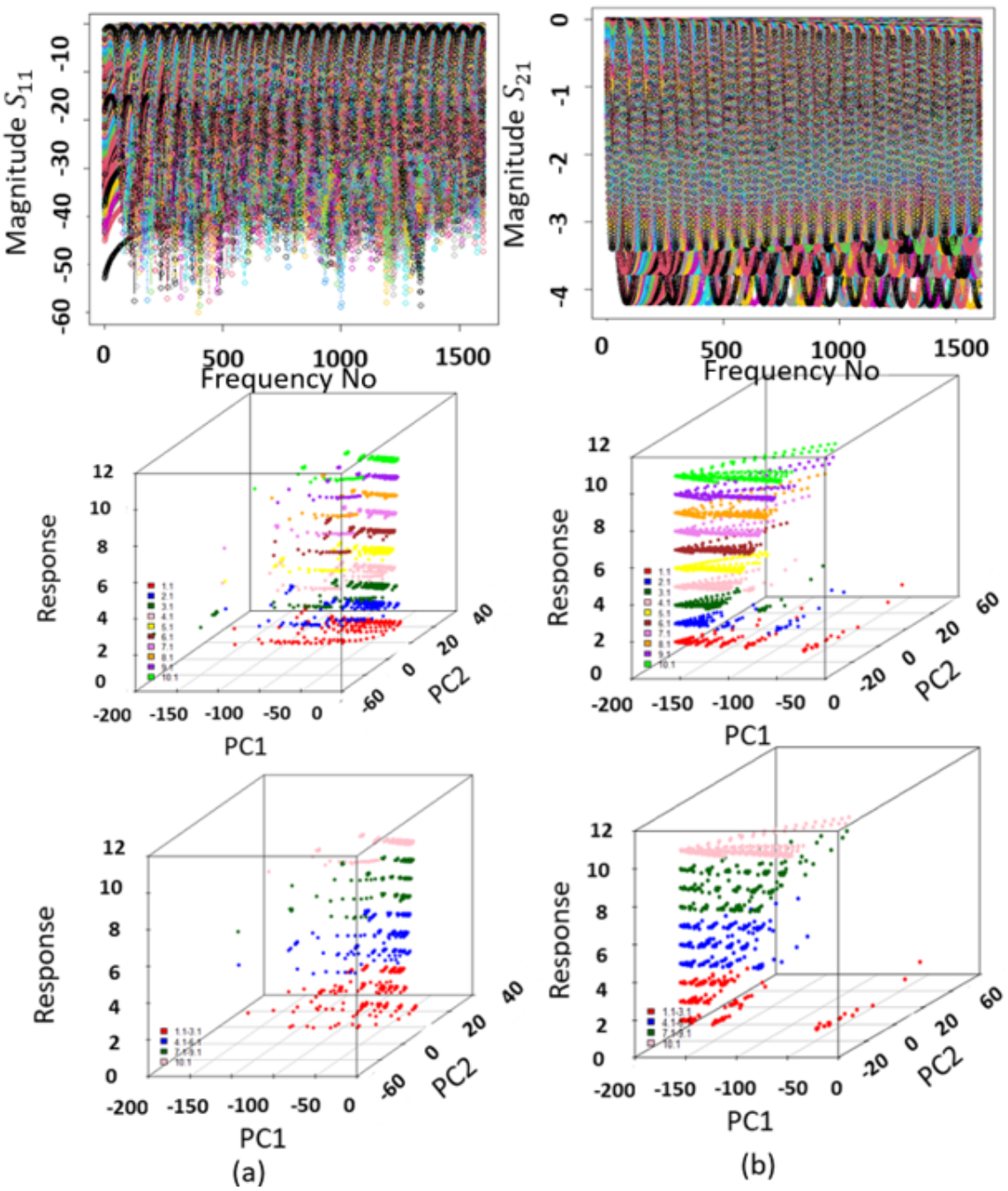

Figure 7.16 Better discrimination of the covariates based on principal components. (a) represents the magnitude of reflection coefficients whereas (b) represents the magnitude of the transmission coefficients.

Figure 7.18 shows as we decrease the number of features, the lasso model is selecting random features from the six sets. This is happening, as there is high correlation among the features across the frequencies. This approach of using lasso to select a few principal informative frequencies

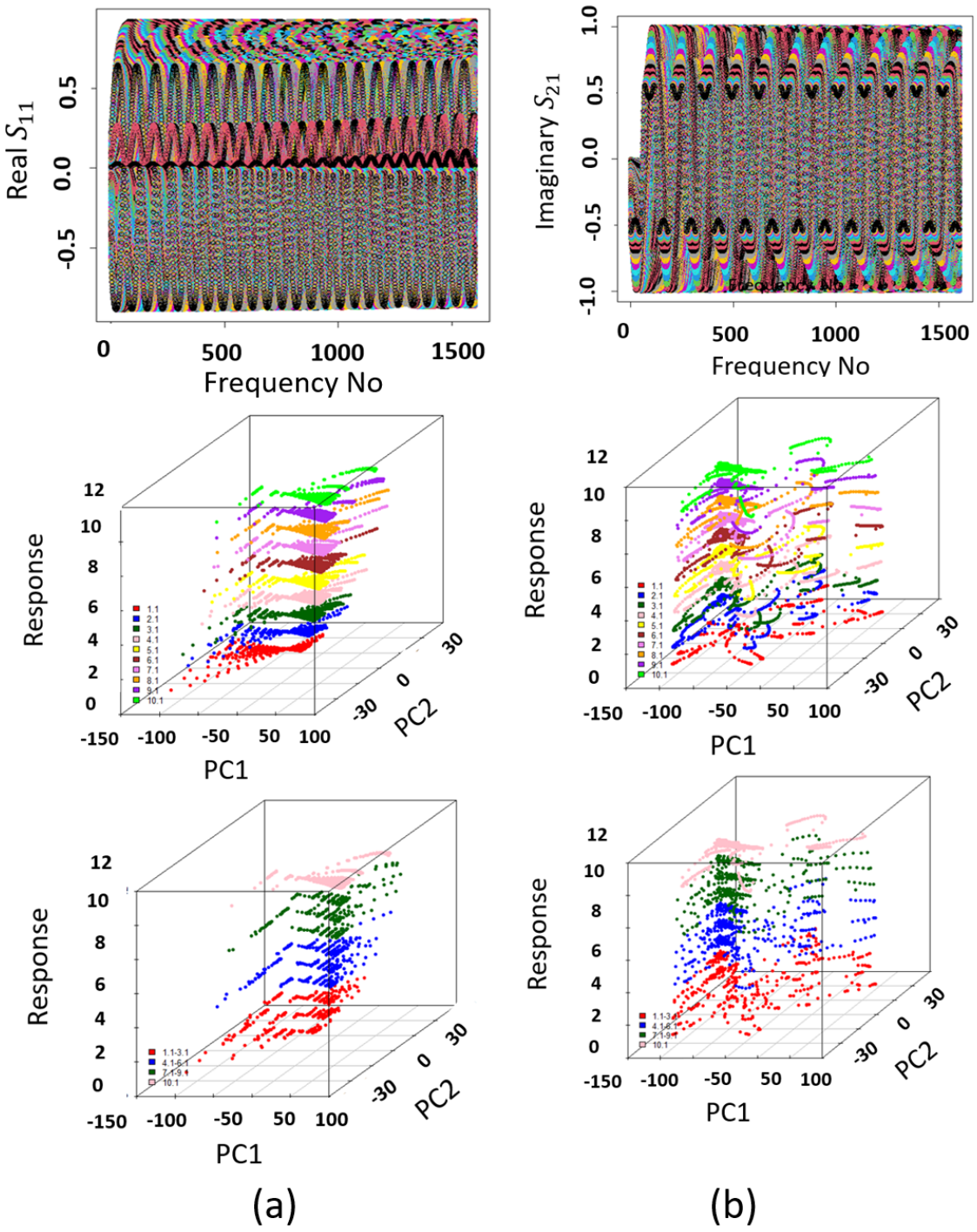

Figure 7.17 Better discrimination of the covariates based on principal components. (a) represents the real of reflection coefficients S_{11} whereas (b) represents the imaginary component S_{21} of the transmission coefficients.

leads to spurious features as lasso model is randomly choosing some features as 1 and others as 0. Hence, we focus on principal component in our analysis.

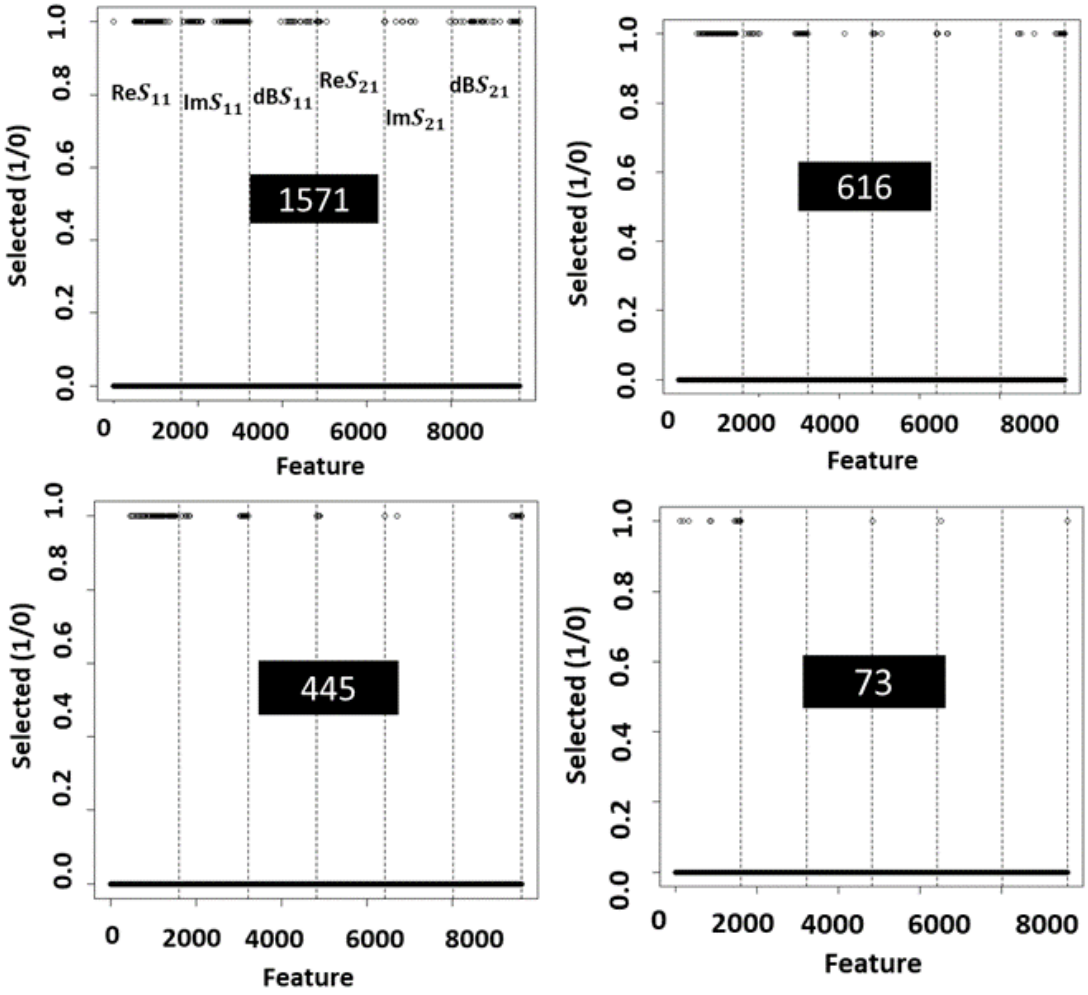

Figure 7.18 Feature selection by lasso for different values of λ.

7.2.5 Registration aided Regression

Linear regression model is fitted after applying principal component analysis on the simulated training set. It gives the optimum training error of 1.10 RMSE. Figure 7.19 (a) shows the histogram plot of the permittivity of the training set where the true and the estimated permittivity get superimposed in majority of the cases. Figure 7.19 (b) shows violin plot distribution of the estimated permittivity against the true permittivity. The permittivities are binned into 10 groups as shown in the X axis. The mean of each violin plot distribution is marked by white dot. From the violin plot

it is observed that for lower permittivity, the predictions are a bit overestimated.

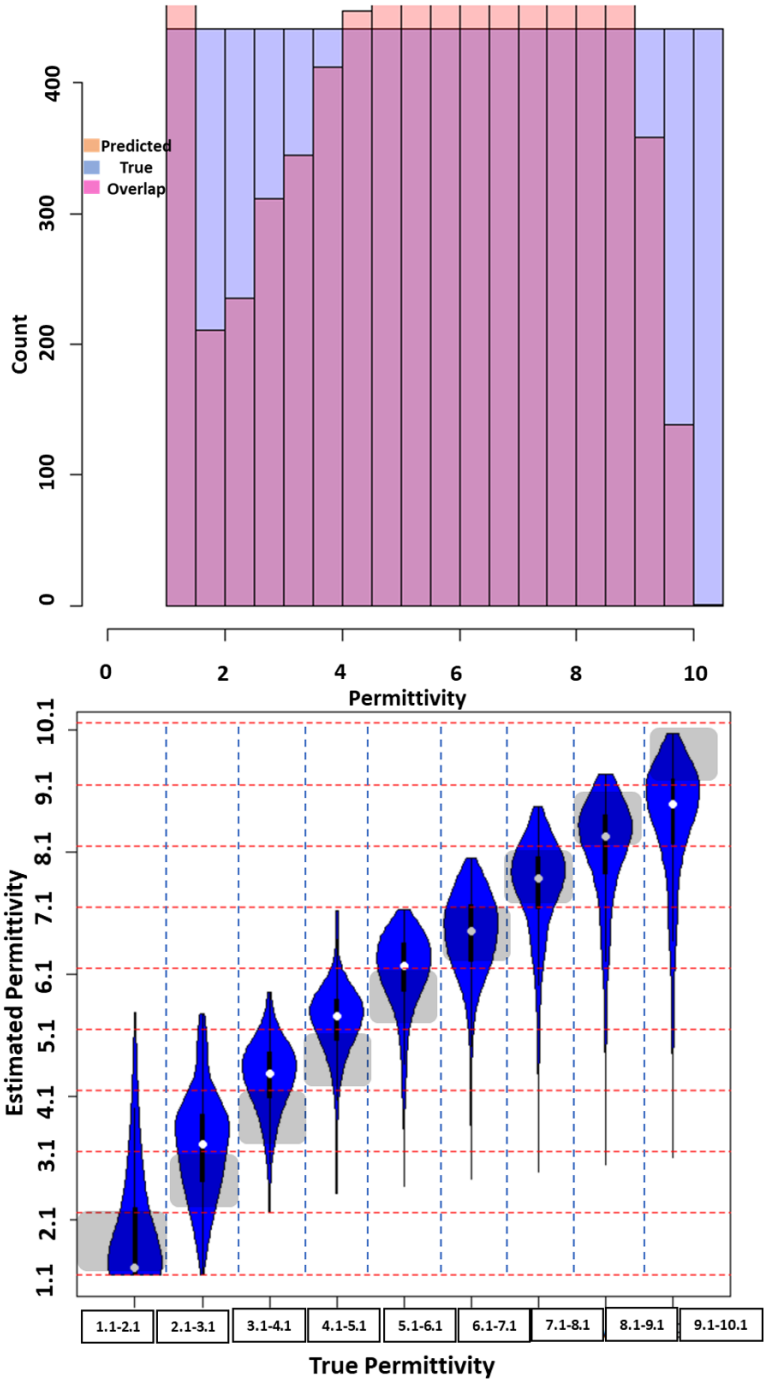

Figure 7.19 Plots of the training error on simulated set: (a) Histogram plot showing the estimated permittivity (light pink), true permittivity (light blue) and the overlap in (dark pink) , (b) Estimated vs True permittivity on training set.

For permittivity prediction on the experimental test set, a naive linear regression is first applied after performing the frequency or column reduction by PCA. Dimensionality reduction is applied

separately on the training and test set as discussed above. After reduction, the dimensions of the train and test matrix become $8379 * 20$ and $32 * 20$ respectively. A linear regression model is fitted by equation 4.10 on the training set. The same model that we learn from the fit is used to predict on the test set. However as the data sources in the train and test set are different, hence simply fitting the same model on test set results in misregistration and thus results in high RMSE by the naive model.

Figure 7.20 shows the difference in alignment of S parameters in the simulation and experimental data using same material substrate (Rogers 3039851) and line dimensions in simulations and experiments.

As the curves in simulation and experiments are different, there is a need to do proper registration/ alignment (164) while applying the regression coefficients learned from simulations on experiments. Knowledge of a portion of the test data as calibration helps in the accurate transfer of information from large simulation database to fewer test data.

Here we have considered 10 rows out of the experimental data for calibration. We have used intercept correction based registration to transfer information. It is the simplest method to transfer information in the linear model setup. The linear model trained on simulation data is used on experimental data by allowing only one parameter, the intercept parameter to change in the model. The intercept parameter is tuned on calibration data set. As we are in linear model setup, there exists a closed form expression for the δ coefficient which is the difference in the means of the observed and predicted permittivity in the calibration data set. With this tuned registered coefficient, we then use the registered linear model for prediction on the experimental data. We formally present the approach below. Let the model based on simulation data be:

$$\texttt{Simulated Data:} \quad \hat{Y}_i = \alpha + \sum_{k=1}^{K} \beta_k x_{ik}, \tag{7.10}$$

where, $i = 1, \ldots, n_s$ and n_s is the sample size of the simulated data. The model on the response for the experimental data is:

$$\texttt{Experimental Data:} \quad \hat{Y}_j = \delta + \alpha + \sum_{k=1}^{K} \beta_k z_{jk}; \tag{7.11}$$

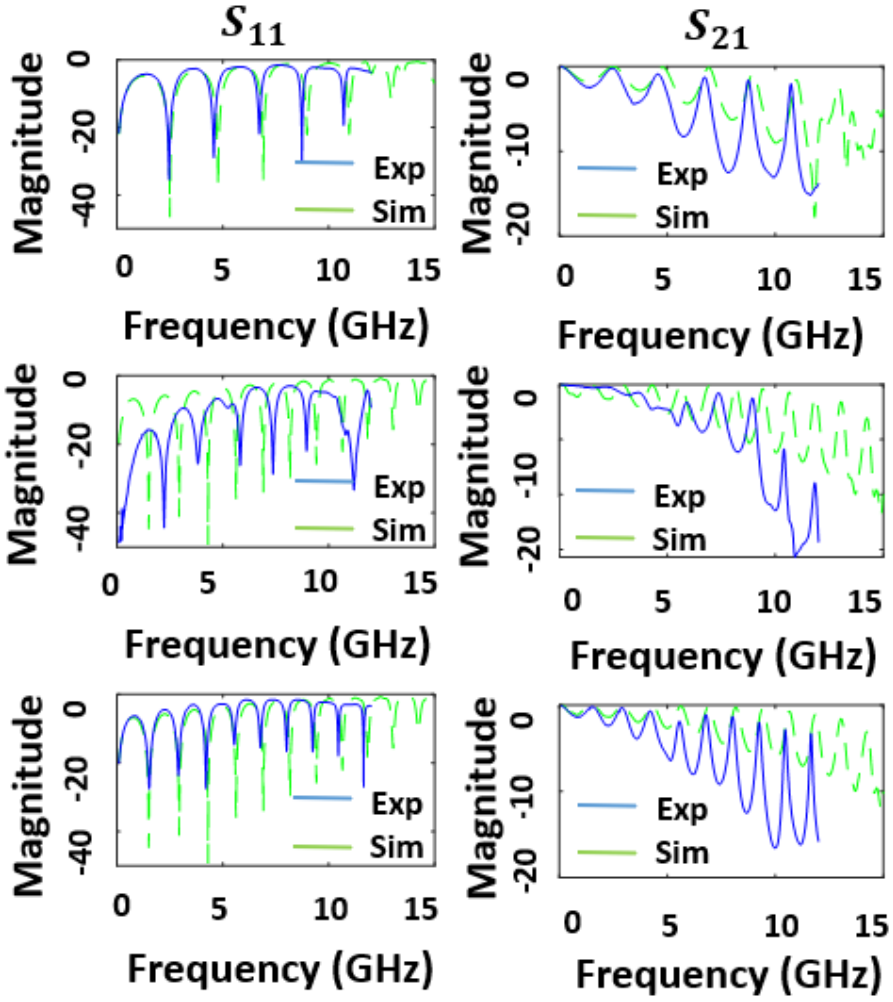

Figure 7.20 Plots of the reflection and transmission coefficients (magnitude in dB) for both simulation and experiment of Rogers 3039851: (upper) Reflection and transmission coefficients when W= 5mm and L= 30 mm, (mid) Reflection and transmission coefficients when W= 2mm and L= 40 mm, (bottom) Reflection and transmission coefficients when W= 4mm and L= 50 mm.

where, $j = 1, \ldots, n_e$ and n_e is the sample size of the experimental data. Now, the registered coefficient δ is learned from the calibration (C) data as follows. Say D_j is the observed permittivity and w_j are different features of S parameter curves of the calibration data. In equation (7.12) the permittivity of the calibration set is predicted by fitting the model on the simulation data.

$$\texttt{Calibration:} \quad \hat{D}_j[\delta] = \delta + \alpha + \sum_{k=1}^{K} \beta_k w_{jk} \tag{7.12}$$

where, $j = 1, \ldots, n_C$ and n_C is the sample size of the calibration data. Next, the residuals are then

evaluated as:

$$\hat{\delta} = \texttt{argmin}_{\delta} \sum_{j \in C} (D_j - \hat{D}_j[\delta])^2$$

and thus,

$$\hat{\delta} = \texttt{argmin} \sum_{j \in C} (D_j - (\delta + \alpha + \sum_{k=1}^{K} \beta_k w_{jk}))^2.$$

Thus, $\hat{\delta}$ is the difference in mean of the observed and predicted values in the calibration set as shown in equation (7.13)

$$\hat{\delta} = \frac{1}{n_C} \sum_{j \in C} (D_j - \alpha + \sum_{k=1}^{K} \beta_k w_{jk}) \tag{7.13}$$

This registered coefficient is then used as correction factor while performing prediction on test set.

Table 7.7 Performance of the registered regression methods vs non registered ones on both test and training sets by intercept registration.

Methods	Experimental Data		Simulated Data
	Test Errors		Training Error
	With Registration	Without Registration	
LR on PCs	3.21	6.8	1.1
RF on PCs	2.72	3.6	1.04
LR on Energy features	6.3	15.2	1.29

In table 7.7, we report the performance of (a) linear regression (LR) (b) random forest (RF) regression (c) linear regression method that uses energy (L_2 norm) of the frequency sweep for each of S parameters. The intercept calibration in all the models is similar to the procedure in (7.11)–(7.13). The first two methods are based on PCs. The energy based linear regression method only has 7 coefficients and does not involve PCs. All the three methods have very low training RMSE on the simulated data. We found that the test RMSE on the simulated data is also similar which is expected as the simulated data set was large. Direct application of these models on the experimental data produced very high experimental test errors. Using registration greatly improved the performance of the PC based linear regression model as well as that of PC based RF regression model. Next, based on the registration both parametric (linear regression) and non parametric

models (Random Forest) is developed. Random forests combine tree predictors that depend on values of a random vector sampled independently and identically for all trees in the forest. This methodology was proposed by Breiman (165) and become very popular owing to its simplicity and high accuracy. Random Forest is mostly used with classification trees. Random Forests grow multiple classification trees, and classify a new object from an input vector by putting it down each of the trees in the forest. Then each tree gives a classification, and the majority of "votes" determines a class of a prediction. More specifically, each tree in random forest grows as follows. If a number of observations in the training set is n, it takes a sample of n observations at random with replacement, from the original data. This creates a training set for growing the tree. Then, if there are M input predictors, a number $m < M$ is specified such that at each node, m variables are selected at random out of the M and the best split on these m divides the node. The value of m is being held constant during the forest growing. Then each tree grows to the largest extent possible with no pruning applied.

As such PC based LR improved by 50% due to registration, abet from a really high predictive error rate without registration. The predictive error of the PC based RF regression model on the experimental data improved by 25% due to the registration procedure. It has the best performance among the concerned methods. Considering that we had only very few experimental samples, its predictive accuracy was encouraging. The energy based LR model had very high predictive error even after registration which demonstrates that all regression models are not capable of properly transferring information from simulated data to experimental data via registration.

Next, we have performed registration by data augmentation where we have pooled the information from the calibration data into the training set and have created a augmented training data. Here the the calibration data also contributes to the change in slope as illustrated in schematic of figure 7.21.

$$\texttt{Simulated Data:} \quad \hat{Y}_{i,sim} = \mathbf{X}'_{i,sim}\beta \tag{7.14}$$

$$\texttt{Calibration Data:} \quad \hat{Y}_{j,cal} = \mathbf{X}'_{j,exp}\beta + \alpha \tag{7.15}$$

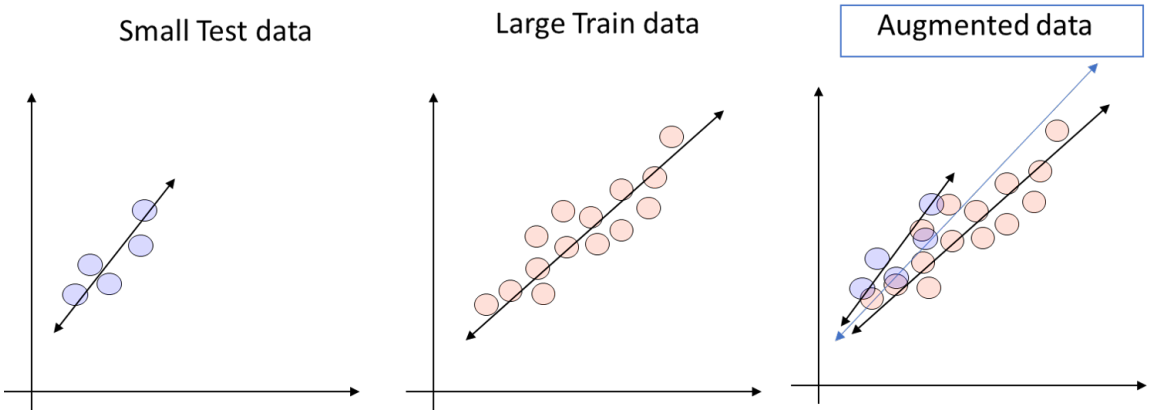

Figure 7.21 Schematic of the data augmentation based registration framework.

Now the pooled model constituting the training data is as follows:

$$\texttt{Training Data:} \quad \begin{bmatrix} \mathbf{Y}_{sim} \\ \mathbf{Y}_{cal} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{sim} \\ \mathbf{X}_{cal} \end{bmatrix} \beta + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \times \alpha \tag{7.16}$$

Table 7.8 shows that performing data augmentation based registration the error on test set is reduced by significant margins in comparison the non-calibrated regression.

Table 7.8 Performance of the registered regression methods vs non registered ones on test sets by data augmentation.

3 Methods	Experimental Data	
	Test Errors	
	Registration by Data Augmentation	Without Registration
LR on PCs	2.98	6.8
RF on PCs	2.79	3.6
LR on Energy features	5.6	15.2

Portions of this chapter appeared in the publication (166).

7.2.6 Discussion

Novel registration aided curve fitting model for microstrip transmission line based material characterization has been developed. The covariates based on reflection and transmission coeffi-

cients are treated as functional data against the frequency sweeps. The source of training and test data being different, there is a need to perform registration to transfer the information. The result shows that registration based regression can be an efficient tool to transfer information and to do estimation on scarce experimental test sets.

Portions of this section has been accepted in ICPHM 2023 and awaiting publication.

CHAPTER 8

ENHANCED DEFECT DETECTION IN NDE USING REGISTRATION AIDED HETEROGENOUS DATA FUSION

8.1 Introduction

Due to advancements in sensing technologies, the smart nondestructive evaluation (NDE) community is currently experiencing a significant increase in the use of multiple sensing methods. It has become crucial to develop robust and fully automated fusion algorithms that can comprehend various physics-based inspection technologies. The concept of "data fusion" has been actively researched in the field of noninvasive inspection since the late 1990s (167; 37). In practical scenarios, establishing a single NDE inspection procedure that replicates the optimal, less noisy conditions of laboratory settings, where these methods were initially created and validated, often poses challenges. Uncertainties such as scanning speed, lift-off, probe drift, low sensor sensitivity, fluctuations in material surface, and other scanning noises (12; 36; 116) introduce uncertainties in the collected NDE data. The signal variations caused by micrometer-level cracks and corrosion are weak and easily overshadowed by background fluctuations and structural noises (28; 29; 99; 168). Consequently, leaks and breakages in pipe structures can remain undiagnosed, resulting in catastrophic fatalities and significant property and life losses (61; 62). Therefore, there is a need for accurate defect detection through inline inspection (ILI) that provides reliable information about the shape, size, and location of structural defects.

The commonly used NDE approaches includes magnetic flux leakage (MFL), Eddy current (EC), Ultrasonic testing (UT), thermography, electromagnetic acoustic transducer (EMAT), pulsed eddy current (PEC) and microwave testing (26; 27; 30; 31). Among them, MFL and EC are widely used for detection of surface and near surface defects in metallic conductive pipe walls. In this context, miniaturized in-house MFL and EC probes have been designed, which are inexpensive, highly flexible, and easy to implement. They eliminate the need for cumbersome hardware setups used in conventional MFL pipeline inspection gauges (PIGs) (51). The working principle of MFL and EC is quite straightforward. In MFL the ferromagnetic wall of the pipe is magnetized close to

saturation either by utilizing a permanent rare earth magnet or a coil wound around a ferromagnetic yoke. When defects are present, they cause a reduction in the thickness of the pipe wall, which in turn increases the reluctance and magnetic flux density near the defects. Consequently, a greater portion of the magnetic flux leaks into the surrounding area, and this leakage can be measured using either a Hall effect sensor (for both static and dynamic fields) or a coil (for dynamic fields) (105; 106; 119). The eddy current (EC) considers the electromagnetic (EM) induction based on the inductive properties of the alternating current (AC) in detecting anomaly. EC gets generated in conducting specimen on application of time varying magnetic field which flows in opposite direction to that of current generating it (169). Both these NDE methods enjoy the following advantages over competing NDE methods for inspecting conductive samples: (a) When using inexpensive thermal cameras for imaging, thermographic measurements have limited sensitivity and are not resilient to minor temperature variations (1). X-ray imaging is an expensive method that necessitates careful screening to mitigate the harmful effects of ionizing radiation (33), (b) Laser profilometry and structure light-based optical inspections are unsuitable for examining ferromagnetic conducting pipe walls due to the occurrence of unwanted reflections (99), (c) Ultrasonic testing requires a coupling medium and physical contact with the material under test (MUT) (32; 4). Although air-coupled ultrasonic testing eliminates the need for coupling, it necessitates long sound pulses for excitation, making precise timing measurements challenging. Additionally, its application is not suitable for inspecting ferromagnetic metallic pipe walls due to the high impedance of metal structures (5), (d) Microwave imaging operates at high frequencies in the gigahertz range, which necessitates complex data acquisition circuitry and struggles to penetrate deeply into conducting materials [28] whereas capacitive sensing displays non-linearity, as the output voltage is inversely proportional to the distance between the sensing and driving electrodes (12).

In recent times, considerable effort has been dedicated to fusing the results obtained from multiple sensor inspections, as this leads to improved defect detection, interpretation, and ultimately, better decision-making. The fusion of sensing modalities provides two types of information: complementary and redundant, as depicted in figure 8.1. Redundant information refers to the shared

information among multiple sensors, which increases the reliability of defect detection. Complementary information, on the other hand, is unique to each inspection method and enhances the accuracy of the detection system. Image fusion can be performed at three levels: signal or pixel level, object level, and decision level. Signal-level fusion involves combining information associated with each pixel obtained from multi-sensor data into a single fused image, representing fusion at the lowest level. In object-level image fusion, features, object labels, and descriptor information are extracted (170). Decision or symbol-level fusion combines probabilistic decision information obtained from experts or decision makers based on the processed results from individual sources (171). In this study, we have implemented heterogeneous image fusion at the pixel level to develop an automated detection system.

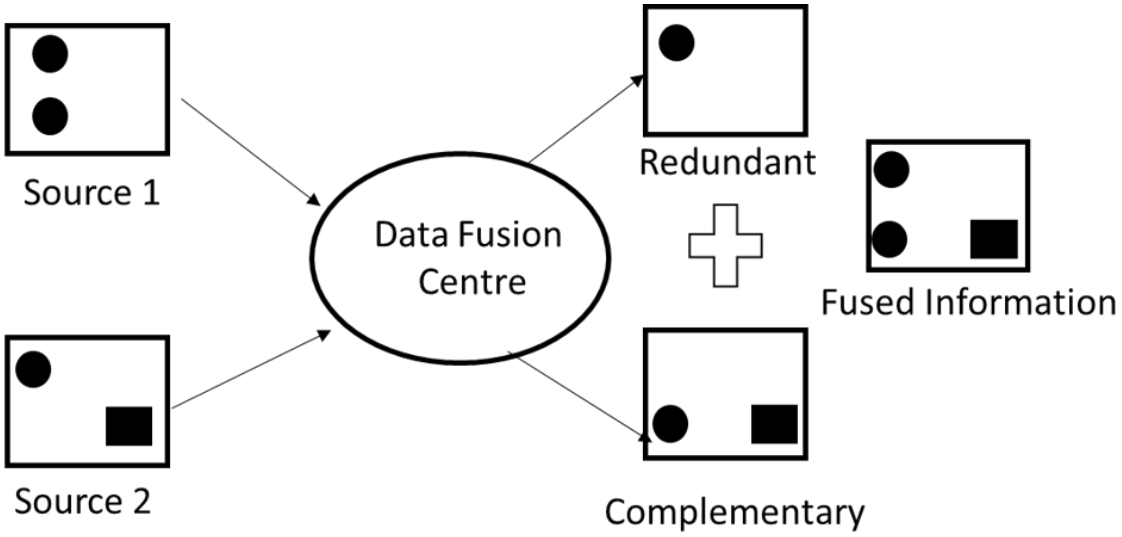

Figure 8.1 Schematic showing the complementary and redundant fusion information.

Heterogeneous fusion involves gathering data from various sensing sources. In the case of magnetic flux leakage (MFL), the sensor information obtained is related to magnetic flux leakage, while in eddy current (EC) testing, the obtained sensor information is complex impedance. Furthermore, in heterogeneous fusion, the issue of misalignment needs to be taken into account due to the different footprints and phenomenological properties of the sensors. On the other hand, homogeneous fusion utilizes data obtained from the same inspection modality but at different frequencies, as demonstrated in (172; 173). Heterogeneous fusion is a powerful tool

for leveraging complementary information compared to homogeneous fusion. In the field of NDE, considerable amount of literature exists applying fusion using conventional techniques like weighted average, Hadamard product, average, difference, Dempster-Shafer rule of combination etc. (174; 173; 175; 176; 132; 177; 178; 179; 180; 181; 182; 183; 184; 168). In (174; 168) a wavelet based fusion method was proposed to fuse the data obtain from EC, MFL and thermographic methods, thereby providing an improved sensitivity-based analysis for crack detection. In (173) the authors developed a generalized additive model (GAM) based fusion model utilizing the EC and Pulse eddy current (PEC) data to detect the hidden corrosion defects in aircraft lap joints. Ultrasonic inspection based on pulse echo immersion technique and induction thermography were used in (175; 176) to evaluate the adhesive bonded quality in composite-epoxy single lap joints based on naïve data fusion algorithms such as average, difference, weighted average and Dempster-Shafer rule. In (37; 177) homogenous fusion techniques were applied to combine the axial and tangential components of MFL signal. There, the fusion process utilized a wavelet basis function neural network (WBFNN), resulting in the generation of complexed value MFL signal bearing additional information. In (178) defect profiles were extracted by fusion of 3D MFL signals using RBFNN. Multi-frequency EC images were combined using discrete cosine transform (DCT) in (179) whereas in (180) Bayesian analysis, Dempster-Shafer theory were used to combine EC and infrared thermal images. Combination of Ultrasonic (UT) and Eddy Current (EC) non-destructive evaluation (NDE) images was achieved using various approaches such as Bayesian analysis, different morphological methods, and minimum mean square error in (181; 183; 182). Authors in (184) fused UT images with EC images obtained from aluminum plates using the AND operation. The defect boundary was extracted from the UT image, while depth information was obtained from the ET image thereby resulting in improved defect characterization. In (185; 186), multi-frequency EC images were fused using neural networks, specifically the multilayer perceptron (MLP) and radial basis function (RBF). These networks learned the fusion weights through the backpropagation algorithm, enhancing the fusion process. Different geometric transforms and ensemble classifier based on the Learn++ algorithm was utilized in (187; 188) to fuse data obtained from MFL and

UT methods thereby aiming to improve the detection of surface defects.

However, to the best of the authors' knowledge, the existing literature on heterogeneous fusion does not address the issue of registration. In heterogeneous fusion, since the data sources are different, blindly fusing them would result in misaligned fused images. This study demonstrates that registration is a common problem in experimentally generated nondestructive evaluation (NDE) data. Additionally, different physics-based sensing methods have different operational trade-offs. Fusion between EC and MFL is beneficial if both techniques provide defect characteristics that complement each other. However, when one of the methods fails, performing fusion would yield suboptimal results (29). The current literature does not mention how fusion would be effective under these boundary conditions. This paper aims to address this gap by automatically determining when to perform fusion based on a data-driven screening rule. Translational registration (189; 190) is then carried out after screening, where the MFL source images are anchored, and based on that, the EC offset images are translated. This approach provides a simple method for registration based on the chi-squared distance. The chi-square algorithm is applied to find the optimal translation parameters for aligning the offset source EC image with the anchored source MFL image. Another registration method utilizing Optimal Transport (OT) (191; 192) is developed, treating the MFL and EC source images as probability densities. A transport plan is established to transfer densities from the stronger (source) to weaker (target) source densities. The transport plan is subject to a cost function that minimizes the distance between the source and target densities (193; 194). Therefore, OT offers an efficient platform for registration in boundary NDE conditions. Pixel-level fusion algorithms are then applied to the registered images. In transform domain-based fusion, each registered source image is represented as base and detailed layers (195; 196). The base layers are decomposed into basis, and the coefficients are evaluated and fused. In spatial domain-based fusion, the spatial closeness of pixels in registered images is considered using kernel smoothing (197; 198). The results demonstrate that this novel registered fusion algorithm is advantageous in noisy conditions and enables better identification of defect dimensions compared to unregistered fusion. Figure 8.2 illustrates the framework of the proposed study. In order to provide a disciplined

analysis of the work, we further divide the study into following components:

- In section 8.2, the experimental and simulation set up of both the MFL and EC NDE techniques and description of NDE data generation under various uncertainties is given. Cheap EC and MFL probes are designed for the purpose of generating data.
- Thereafter, in section 8.3, a data driven screening rule and an improved NDE theory-based fusion is discussed.
- We show the need of performing registration especially on the experimental data in Section 8.4. Detailed analysis of two modes of registration a> Translational registration, b> Optimal Transport (OT) based registration and their properties are discussed in this section.
- In section 8.5, it is demonstrated that performing fusion based on the registered images provide better fused results without artifacts in comparison to that without registration. For fusion we have used sparse representation-based pixel level fusion. We showed our method of registration also worked while using fusion based on bilateral filtering and wavelet decomposition.
- This chapter concludes with a discussion in section 8.6 .

8.2 MFL and EC Data Generation

Source data for fusion is generated by both synthetical and experimental means. Synthetic data is obtained by designing 3D finite element model using COMSOL MULTIPHYSICS 6.0 modeling software. Experimental data is generated by developing in-house cost effective MFL and EC probes. The simulation and experimental setups are discussed in the subsequent subsections.

8.2.1 Simulation & Experimental design of MFL based NDE sensing

The MFL simulation model based on the magneto static problem governed by conventional Maxwell's equation is developed. The benchmark settings as discussed in (28) is used to construct and validate the model. The magnetic circuit consists of a ferromagnetic yoke, two ferromagnetic couplings and two permanent magnets of opposite polarities as in [61,62]. Figure 8.3 provides the schematic of the simulation model where the distance between two magnet poles are 145 mm. The

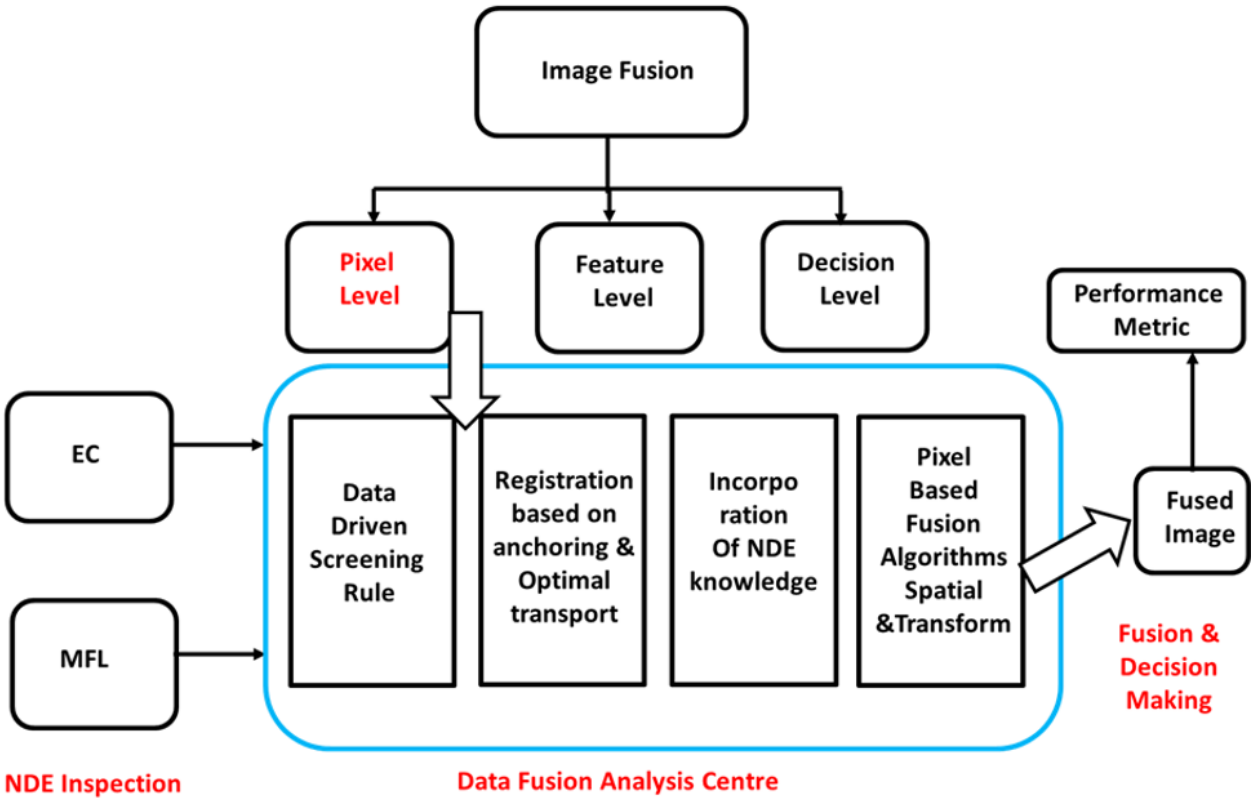

Figure 8.2 Schematic of the entire fusion setup.

existence of the defects increases the magnetic resistance which in turn distorts the magnetic field. The bent magnetic flux, which leaks out of the material surface form a magnetic leakage field that is collected by the 3D Hall effect sensors placed at the middle of the permanent magnets as marked. Dimension of the model parameters is stated in Table 8.1. For faster computation extremely, fine triangular mesh is used near the defects, whereas on the rest of the surface, tetrahedral mesh is employed.

Table 8.1 Dimensions of the different parts of MFL model (L=length, W=Width, H= Height, all the dimensions are reported in mm).

Term	Length(l)	Width (w)	Height (h)	Material
Yoke	320	80	30	Structural Steel
Magnet	30	80	40	NdFeB
Brush	30	80	20	Structural Steel
Specimen	600	400	10	Stainless Steel 405 Annealed

Figure 8.3: (a) Schematic of the conventional 3D MFL model in COMSOL, (b) MFL inspection on a pipe segment The experimental setup is constructed using the following components:

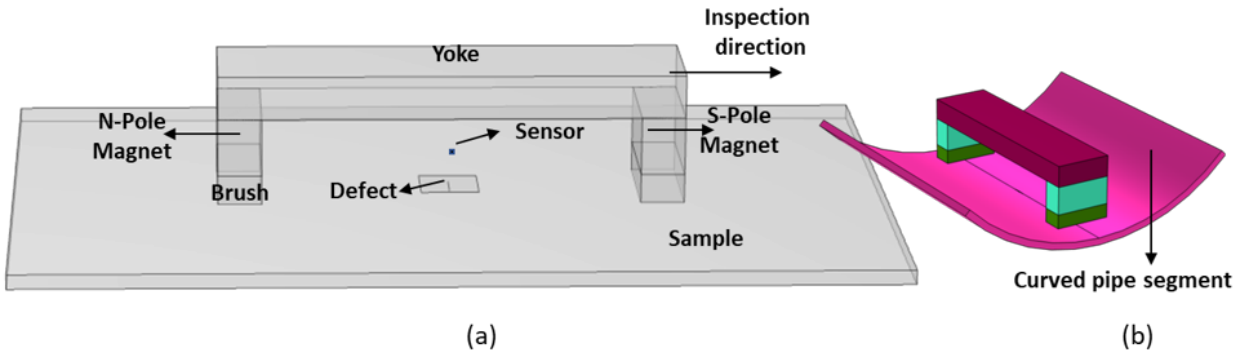

Figure 8.3 (a) Schematic of the conventional 3D MFL model in COMSOL, (b) MFL inspection on a pipe segment.

- MFL probe based on permanent magnets
- TLE493D-A2B6 3-axis Hall effect sensor
- Scanning robot arm for moving the probe along the material being tested
- Data acquisition system (DAS) and corresponding code for generating images
- Other associated units

A schematic of the complete experimental setup is depicted in Figure 8.4. The MFL probe holder accommodates two cubical permanent magnets made of NdFeB material. Two different holders are designed to fit magnets of different dimensions. Figure 4 (b) displays a holder measuring $80 \times 72 \times 30mm$, while a larger holder is shown in Figure 8.4 (e). Inside the holder, two cubical permanent magnets with dimensions of 25 mm are positioned with opposite polarities facing the material under test (MUT). Additionally, larger cuboidal permanent magnets measuring 40 mm*4 mm*20 mm are utilized. The Hall sensor, with dimensions of $25mm \times 14mm$, is positioned in the middle of the holder between the magnets with opposite polarities. For scanning purposes, an Aerotech AGS1000 programmable XYZ scanner is employed, while the TLE493D-A2B6 three-axis Hall effect sensor manufactured by Infineon is used for measuring the axial, radial, and tangential components (in X, Y, Z directions) of the magnetic field leakage. The TLE sensor consists of three main blocks: (a) a power mode control system comprising oscillators, (b) the main sensing unit consisting of Hall probes, a multiplexer, and an ADC controller, and (c) an I2C interface with register files. The sensing unit of the sensor is connected in sequence to a multiplexer and

then to an Analog-to-Digital Converter (ADC), which is connected to the Data Acquisition System (DAS). The DAS incorporates a National Instrument Data Acquisition Card PCIe-6341, where the magnetic flux signal is sampled and digitized by a routine, and the output is recorded.

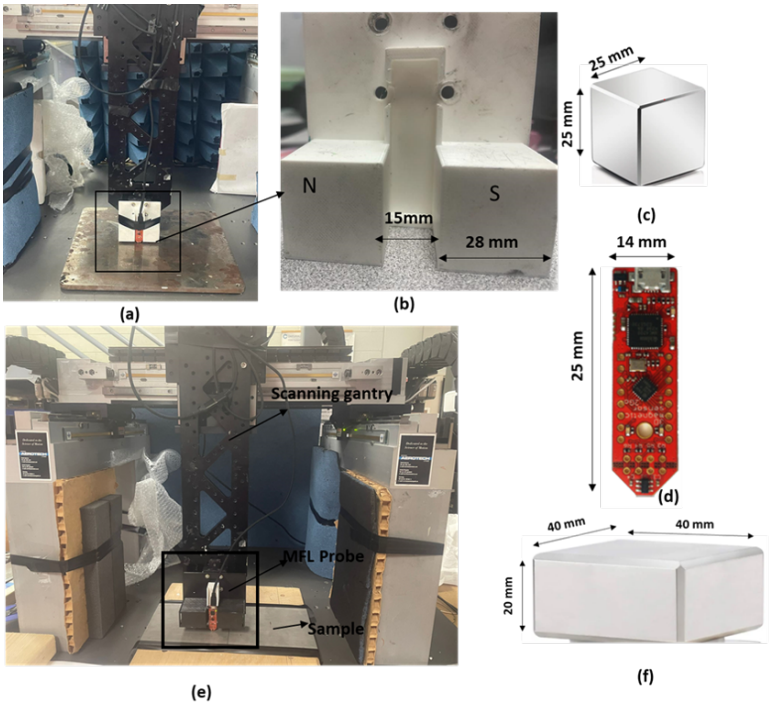

Figure 8.4 (a) Schematic illustration of MFL probe scanning the MUT, (b) Dimensions of the probe holder, (c) Permanent magnet used in sensing, (d) Infenion TLE493D-A2B6 sensor used in sensing, (e) Entire setup with gantry, (f) Dimension of larger permanent magnet.

8.2.2 Simulation & Experimental design of EC based NDE sensing

Eddy current (EC) is extensively employed in non-destructive evaluation (NDE) inspections of metallic components due to its reliable inspection capability, non-contact operation, and sensitivity to defect detection. The EC technique utilizes the principles of electromagnetic induction and the inductive properties of alternating current (AC) to identify anomalies (199; 200; 201). When a time-varying magnetic field is applied, eddy currents are induced within the conducting specimen. The simulated EC model in this study is based on a benchmark model (201). In this model, an air-cored coil is excited with a 1-ampere alternating current, generating an alternating magnetic field at a frequency of 100 Hz. By scanning a conducting specimen, opposing currents to those in the coil are induced, resulting in the generation of eddy currents. Defects are detected by monitoring the change in coil impedance. The schematic of the developed 3D model in COMSOL is depicted in

Figure 8.5. A quarter model is utilized based on its symmetrical characteristics, which significantly reduces computation time and resource usage compared to meshing the entire specimen.

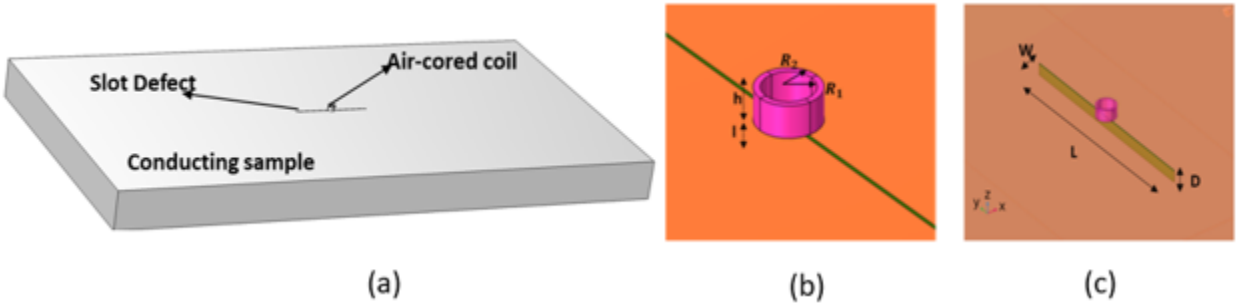

Figure 8.5 (a) Schematic of the simulated EC setup, (b) Different parameters of the air cored coil, (c): Schematic for the measurement of the impedance change due to crack like opening.

The different parameters of the air cored coil such as outer radius, inner radius, number of turns are given in Table 8.2.

Table 8.2 Different dimensions of the EC simulation setup.

Coil	Dimension
Inner Radius	0.8 mm
Outer Radius	1 mm
Coil height h	1 mm
Number of Turns N	4000
Lift-off	1 mm, 3mm, 5mm, 7mm

In the experimental setup an array probe with 18 channels are chosen for rapid inspection. There are two rows each containing 9 coils and the area of each coil is $34mm \times 3.45mm$. The distance between each neighboring coil on a row is 4mm, and the rows are shifted by 2mm with a gap of 3.45mm between the two rows. To obtain the output data from each coil an Ectane 2 testing instrument that allows for multiplexing each channel within the operating frequency range. An input voltage of 6V and an operating frequency of 500 Hz is used to conduct the experiment. Figure 8.6 (a) provides the entire schematic of the different components used in the EC setup, 8.6 (b) provides the diagram of the EC coil sensing topology. Real time data acquisition is done by the software corresponding to Ectane 2 where the results of scan are first interpolated and then stored in .csv format which are subjected to further processing using Matlab. Absolute (ABS) mode is

chosen where the output per coil corresponds to the data pixel as shown in Figure 8.6(b). For post processing the obtained data is at first interpolated along X and Y axis. The real and imaginary components of the image scan is thus obtained. In order to avoid bias due to the "null" regions along the edges usually caused by array configuration, the data is cropped along the edges.

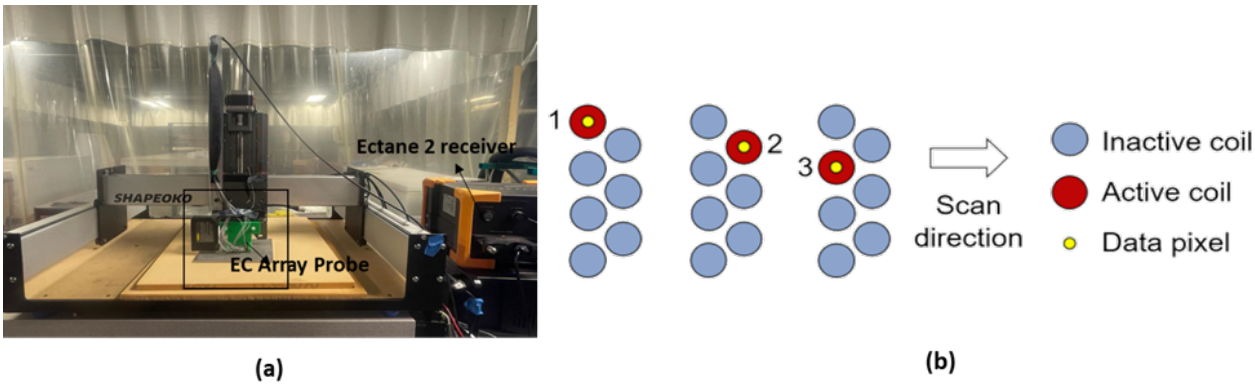

Figure 8.6 (a) Experimental scanning setup based on EC, (b) Absolute sensing topology based switching sequence for EC measurements.

Table 8.3 presents the defect dimensions, uncertainties resulting from lift-off variations, and noise levels associated with the experimental data utilized in the subsequent sections involving registration and fusion techniques.

8.3 Data driven screening rule

Fusion improve defect detection under noisy regimes and various lift offs when both the MFL and EC captures the defect characteristics that are complementary to each other. Then we can integrate these characteristics and provide better defect diagnosis. In (168; 174) it is shown that there lies trade off between heterogenous MFL and EC methods: MFL fail to detect shallow defects whereas EC cannot detect thin defects. In these boundary conditions when one of the methods fails, then fusing images even by the state-of-the-art fusion algorithms lead to sub-optimal detection. In these cases, it is better not to fuse and use the inspection method that is most appropriate for the situation. An automated data-based screening rule is first proposed that assigns weight as 1 to the method that is appropriate for the situation in these limiting conditions.

The data-based rule for deciding the fusion weights is as follows: For both the EC and MFL, at first readings from cases without defects are considered and the standard deviations of the signal

Table 8.3 Design of the defects and conditions incorporated in the experimental study where l: length, w: width and d: depth.

Defect No.	**Dimensions (l*w*d)**	**Conditions**
Defect 1	1cm*0.5mm*0.5mm	Lift off: 1mm
Defect 2	1cm*0.5mm*0.2mm	Lift off: 1mm
Defect 3	2cm*1.2mm*3.8mm	Lift off: 2mm
Defect 4	2cm*1.2mm*3.8mm	Lift off: 5mm
Defect 5	2cm*1mm*2mm	Lift off: 3mm
Defect 6	8mm*8mm*5mm	Lift off: 1mm
Defect 7	8mm*8mm*5mm	Lift off: 5mm
Defect 8	8mm*8mm*5mm	Lift off: 3mm
Defect 9	2cm*1cm*5mm	Lift off: 2mm
Defect 10	2cm*1cm*5mm	Lift off: 2mm, Noise: SNR 10 dB
Defect 11	2cm*1cm*5mm	Lift off: 2mm, Noise: SNR 30 dB
Defect 12	2cm*1cm*5mm	Lift off: 7mm
Defect 13	2cm*1cm*5mm	Lift off: 5mm
Defect 14	2cm*1cm*5mm	Lift off: 9mm
Defect 15	2cm*2cm*5mm	Lift off: 2mm
Defect 16	3cm*2.5cm*5mm	Lift off: 1mm
Defect 17	3cm*2.5cm*5mm	Lift off: 7mm
Defect 18	3cm*2.5cm*5mm	Lift off: 9mm
Defect 19	3cm*2.5cm*5mm	Lift off: 1mm, Noise: SNR 30 dB
Defect 20	2cm*2cm*5mm	No Lift off
Defect 21	2cm*2cm*5mm	Lift off: 5mm

interpreted by both the methods $\bar{\sigma}_E$ and $\bar{\sigma}_M$ are computed (202). Next, for a new reading the significant perturbations in readings w.r.t to the non-defective cases are checked by comparing standard deviations σ_E and σ_M with that of $\bar{\sigma}_E$ and $\bar{\sigma}_M$ respectively. The tuning parameter chosen for checking the standard deviations is c=1.28. The criteria for assigning weights are as follows:

1. The screening rule chooses EC only when $\sigma_E > c\bar{\sigma}_E$ and $\sigma_M \leq c\bar{\sigma}_M$.

2. The rule will choose MFl only when $\sigma_M > c\bar{\sigma}_M$ and $\sigma_E \leq c\bar{\sigma}_E$.

3. Assign weights to both the methods if $\sigma_E > c\bar{\sigma}_E$ and $\sigma_M > c\bar{\sigma}_M$.

Figure 8.7 shows the data based screening rule where in 8.7(a) the screening rule gives weight as 1 to the EC source where the defect is of dimension width:0.5mm, height:0.05mm. In figure 8.7(b) the rule selects MFL for defect of dimension width:0.01mm, height: 0.5mm, whereas both the MFL and EC will be selected in third condition where a defect of width: 0.1mm and height:

0.1mm is considered. In figure 8.7, the data is generated by synthetical means at lift off of 3mm. In the later experimental cases, the axial component of the magnetic flux leakages B_x in the MFL data are considered. However, if contemplating the tangential component B_z, then performing blind fusion with EC source results in erroneous fused images as shown in Figure 8.8 (e). Figure 8.8 illustrates that NDE based phenomenological knowledge can provides better fusion. B_z contains a plateau and trough with the defect in the mid region as demonstrated in 8.8(a). The central ridge region is extracted by contour plot based extraction as in 8.8(b) and defect localization based on B_z is thereby shown in figure 8.8(c). On performing fusion of reconstructed B_z with EC source results in better fused result indicated in8.8(f). The SNR of the source MFL is 3.33, source EC is 16.32 whereas after fusion the SNR of the fused image gets increased to 22.35. The fusion algorithm is discussed in the later sections.

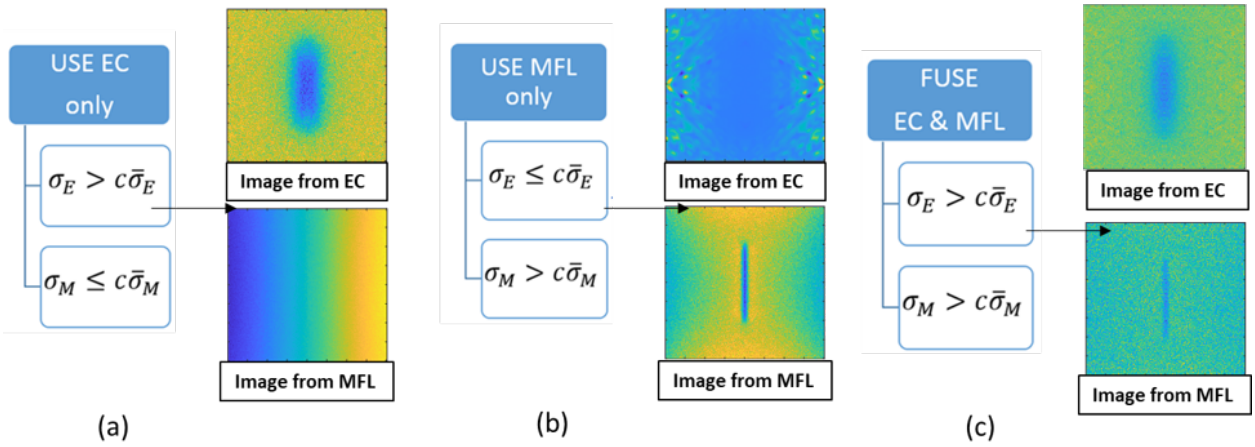

Figure 8.7 Screening rule choosing EC criteria, (b) choosing MFL criteria, (c) choosing both EC and MFL criteria respectively.

8.4 Registration

The inspection data sources being heterogeneous, there lies a need for proper alignment of the source images especially in experimental data. In this section, we have shown that performing fusion without proper registration results in artifacts in fused images. Offset arises while conducting raster scans by different sensors giving rise to misregistration. Registration is a process of establishing a common geometric frame for data coming from different modalities. Here two kinds of registration strategies are implemented, (a) Translational registration by anchoring, (b) Registration by Optimal

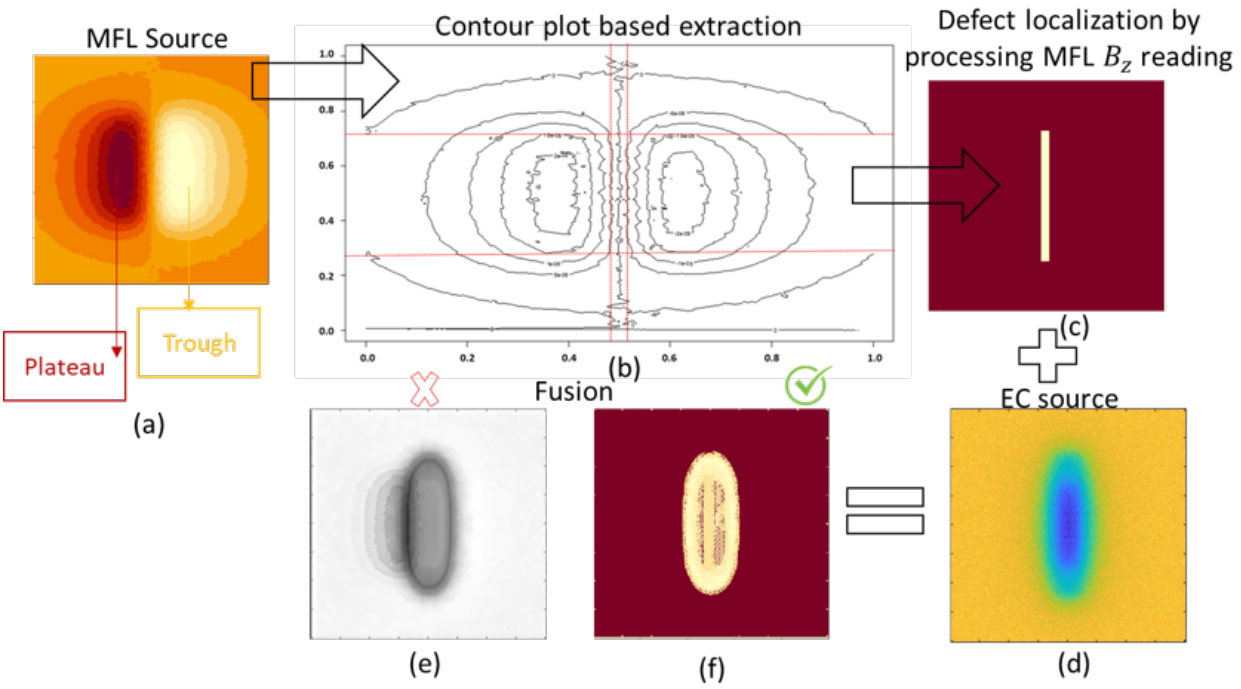

Figure 8.8 Fusion based on NDE based phenomenological knowledge.

Transport. In translational registration the MFL source images are anchored and the offset of the EC image w.r.t to anchored MFL image is obtained in both x and y coordinates. We have adopted a discrete Fourier transform (DFT) based up sampling method for the purpose of registration as this is both time and memory efficient. At first, an initial estimate of the cross-correlation peak is obtained by fast Fourier transform (FFT) method. Then the initial shift estimation is improved by up sampling the DFT in a small neighborhood of that estimate by matrix multiplication in an iterative fashion as in (190). Thus, all the image points of the target EC image get registered with respect to the referenced MFL image. In Figure 8.9, we demonstrate the efficacy of translational registration on Defect 9 case of Table 8.3. Figure 8.9(a), (b) shows the source images. Since there is an offset in the EC image, performing blind fusion by sparse representation (discussed in next section) results in misaligned incorrect fusion as shown in 8.9(c). Figure 8.9 (d) is the results of translational registration of EC image with the referenced MFL image. Figure 8.9(e) and (f) demonstrate fusion of the registered EC image with MFL image by sparse fusion. Post processed result 8.9(f) does not show any overlapping erroneous patterns like that in 8.9 (c). SNR of the fused image after registration is 85.56 which is significantly higher than that of source images,

MFL:40.33, EC:6.97.

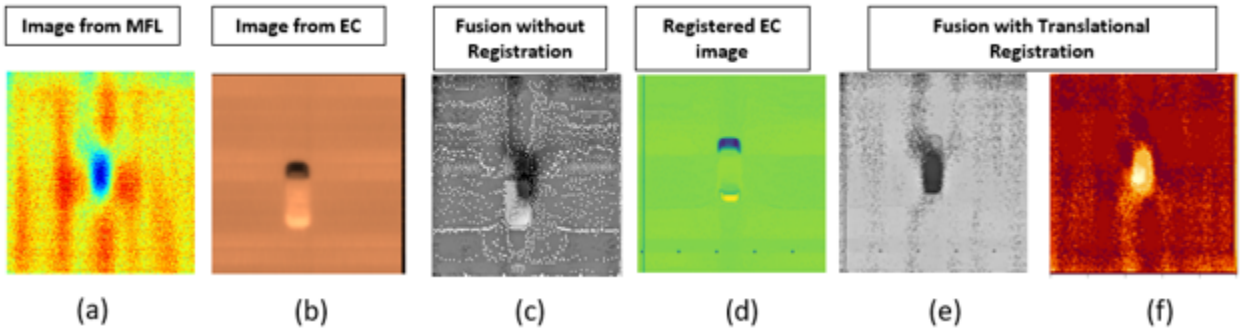

Figure 8.9 Benefits of translational registration(a) Inspection by MFL source, (b) By EC source, (c) Fusion of (a) and (b) by SR, (d) Registered EC source by translational registration, (e) fusion between (a) and (d), (f) post processed result of (e).

In this section, we have developed optimal transport (OT) based image registration, where we are transporting the mass between two distributions of inspection modalities. The transformation plan (T) between the source and the target distribution are chosen based on minimization of Kantorovich-Wasserstein or Earth Mover's Distance (EMD) under a mass preservation constraint (190; 191; 192; 193; 194; 203). The objective is to choose that transport map which minimizes the overall cost of transportation i.e to minimize the l_2 distance between the source and target distribution in embedded space. OT does not need any specific landmark for registration and thus has advantage over other registration methods. In multi modal NDE fusion, Optimal transport is mostly effective in performing registration in boundary conditions when one of the methods fails in defect detection. Under these circumstances where screening rule suggest not to fuse, optimal transport (OT) transfer mass density from the stronger inspection source (source) to the weaker source (target). Then performing fusion, on these registered source images results in better defect localization and SNR. The setup of OT is as follows: OT transformation (T) is defined from a reference density function $\mu_s(x)$ to target density $\mu_t(x)$. Now, assuming source and target density functions to be positive on domains $\Omega_{\mu_s}, \Omega_{\mu_t} \subseteq R^d$ respectively and after normalization such that:

$$\int_{\Omega_{\mu_s}} \mu_s(x)\, dx = \int_{\Omega_{\mu_t}} \mu_t(x)\, dx = 1$$

The minimizer of the distance functional is unique, provided the source and target densities have

finite second order moments. The minimum transport map *min(T)* is as follows:

$$min(T) = \int_{R^d} |x - T(x)|^2 \mu_s(x)\, dx$$
$$\text{s.t.} \int_B \mu_t(y)\, dy = \int_{T^{-1}(B)} \mu_s(x)\, dx \forall B \subseteq R^d$$

when T is smooth and bijective then the above equation can be simplified as Jacobian equation as:

$$\mu_s(x) = |\det((x))|\mu_t(T(x))$$

The minimizer to the above problem gives us the unique optimal transport map between the successful modality to unsuccessful one. The advantage of optimal transport is that, it uses gray scale data from both the images and place them in same normalized scale. Moreover, no landmark is needed for successful registration. Figure 8.10 illustrates the schematic of OT where (a) represents the transport (T) of scenario 'Defect 9' from Table 8.3. Here EC is considered as source and MFL the target. In Figure 8.10 (b) we have chosen random 1000 points from each image source for better visualization and have plotted the uniform distribution of the samples. Figure 8.10(c) shows the OT distance from the source (in blue marks) to target (in red marks).

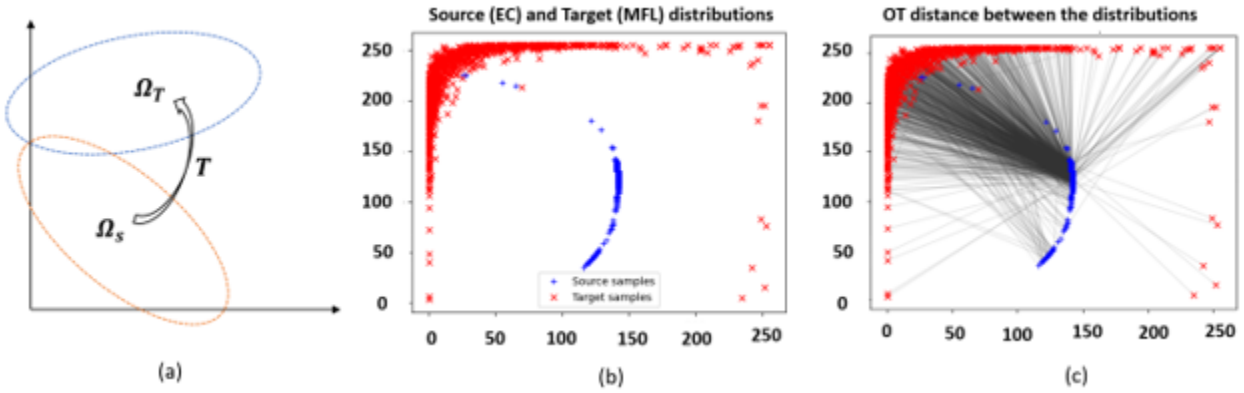

Figure 8.10 Schematic of the OT framework: (a) Transfer from source (EC) to target (MFL) for scenario 'Defect9', (b) 1000 random samples are chosen and uniform distribution is shown, (c) OT distance between the source and target samples.

Figure 8.11 shows the improvement of using OT as registration for Figure 8.9 over translational registration. Here information is transferred from MFL to EC. Panel A (d) and (e) shows the improved fused result on fusion with OT based registration whereas by fusion with translational registration still shows some mis registration. In Panel B the forward optimal transport map is

illustrated where (d) shows clear displacement in EC source image due to transport from MFL source. Panel C demonstrates the geodesic or the optimal path connecting the two distributions in embedded space. The steps of transferring information following the minimal path is shown in this panel. The SNR of the fused image after OT based registration increased to 172.92. The tuning parameters of the iterative OT algorithm is as follows: learning rate of 0.001 with a decay of 0.01. The maximum number of iterations to run is 10,000 and the minimum tolerance of the cost function used is 0.00001.

Similarly, for 'Defect 16' scenario of Table 8.3, OT based registration results in better fusion than the unregistered fusion. Steps similar to Figure 11 are illustrated in Figure 8.12.

8.5 Pixel based fusion

8.5.1 Sparse representation-based fusion

In this section transform domain-based fusion, method in the form of sparse representation is adopted on the registered source images. In transform domain-based methods, the information contained in the original images are merged after converting into some other domain such as discrete wavelet transforms (147), discrete cosine transforms (204), sparse representation (205) etc. In the convolutional sparse representation (SR) the sparsity of the entire source image is obtained and the source image is decomposed into base and detailed layers (195). The schematic of the fusion framework is shown in Figure 8.13. The fusion framework consists of two segments (a) the data fusion analysis center to perform screening judgement and registration, (b) sparse representation based fusion by decomposing into basis and detailed coefficients.

For a set of K heterogeneous source images ($K = 2$) as the sensing modalities are EC and MFL and a set of dictionary filters $d_m, m \in 1\cdot, M$ the CSR method consist of the following steps.

1. Decompose each image into two layers: Base and Detailed. Base layer obtained by solving a Tikhonov Regularization problem.

2. Subtraction of the base layer from original image to obtain the detailed layer. Decomposing the detailed layer into basis and calculate the coefficients under sparsity using convolutional

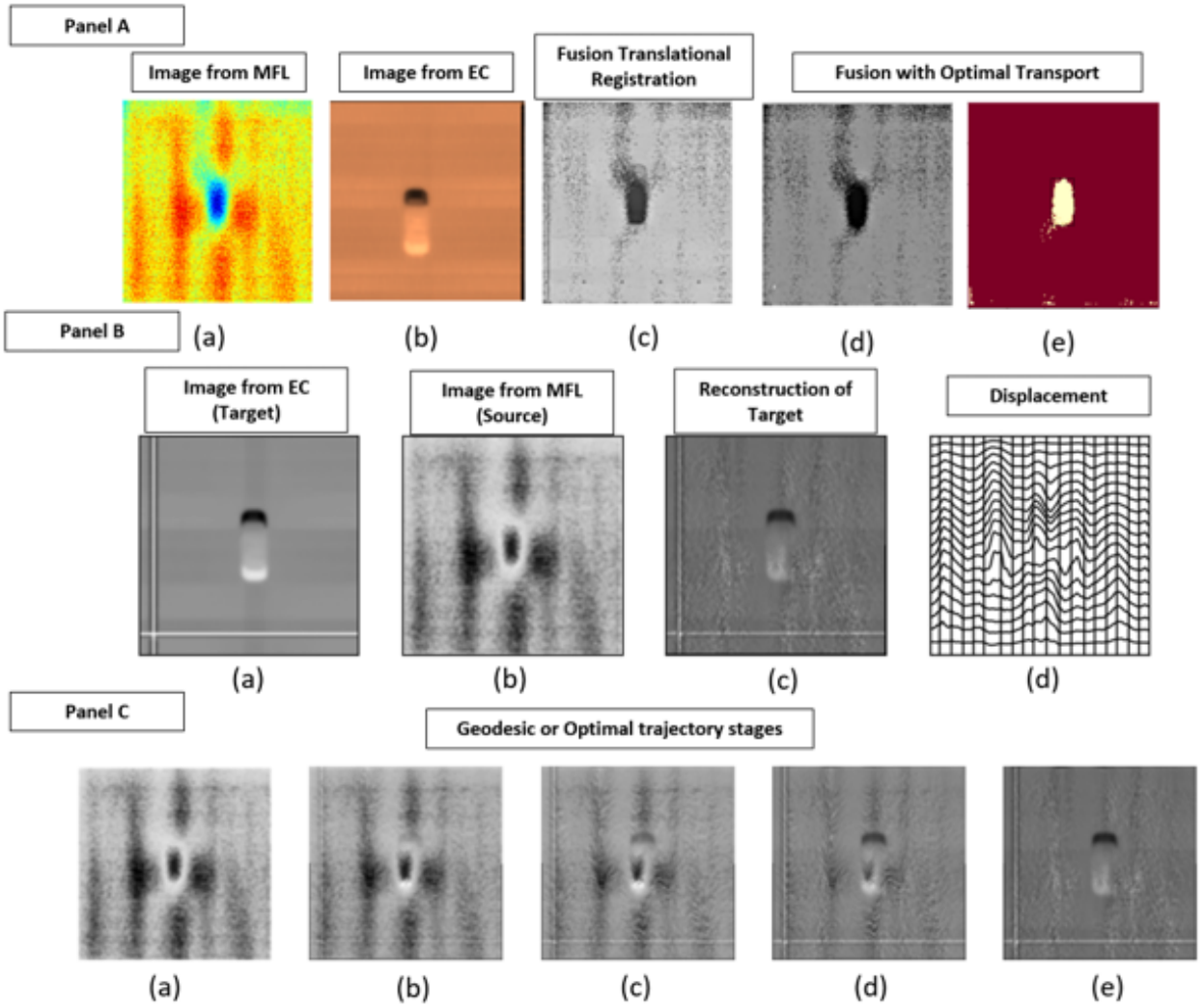

Figure 8.11 Panel A: (a) Source MFL image, (b) Source EC image, (c) result on applying fusion after performing translational registration, (d) result in applying fusion based on OT, (e) post processed results on (d), Panel B: (a) EC image which forms the target distribution, (b) MFL image acting as source distribution in OT, (e) Reconstruction of target after OT, (d) Displacement showing transfer of mass from source to target, Panel C: Optimal trajectory stages showing transfer of mass from source to target in five stages. Analysis is done based on 'Defect 9' of Table 8.3.

basis pursuit denoising (CPDN) algorithm.

3. Activity level maps of the coefficients are hereby obtained.

4. Finally, the fused image is constructed by fusion of both base and detail layers.

The base layer for a source image $I^b_{k=1,2}$ is obtained by solving the Tikhonov regularization problem using the fast Fourier transform of the source data $I_{k=EC,MFL}$. The optimization problem for

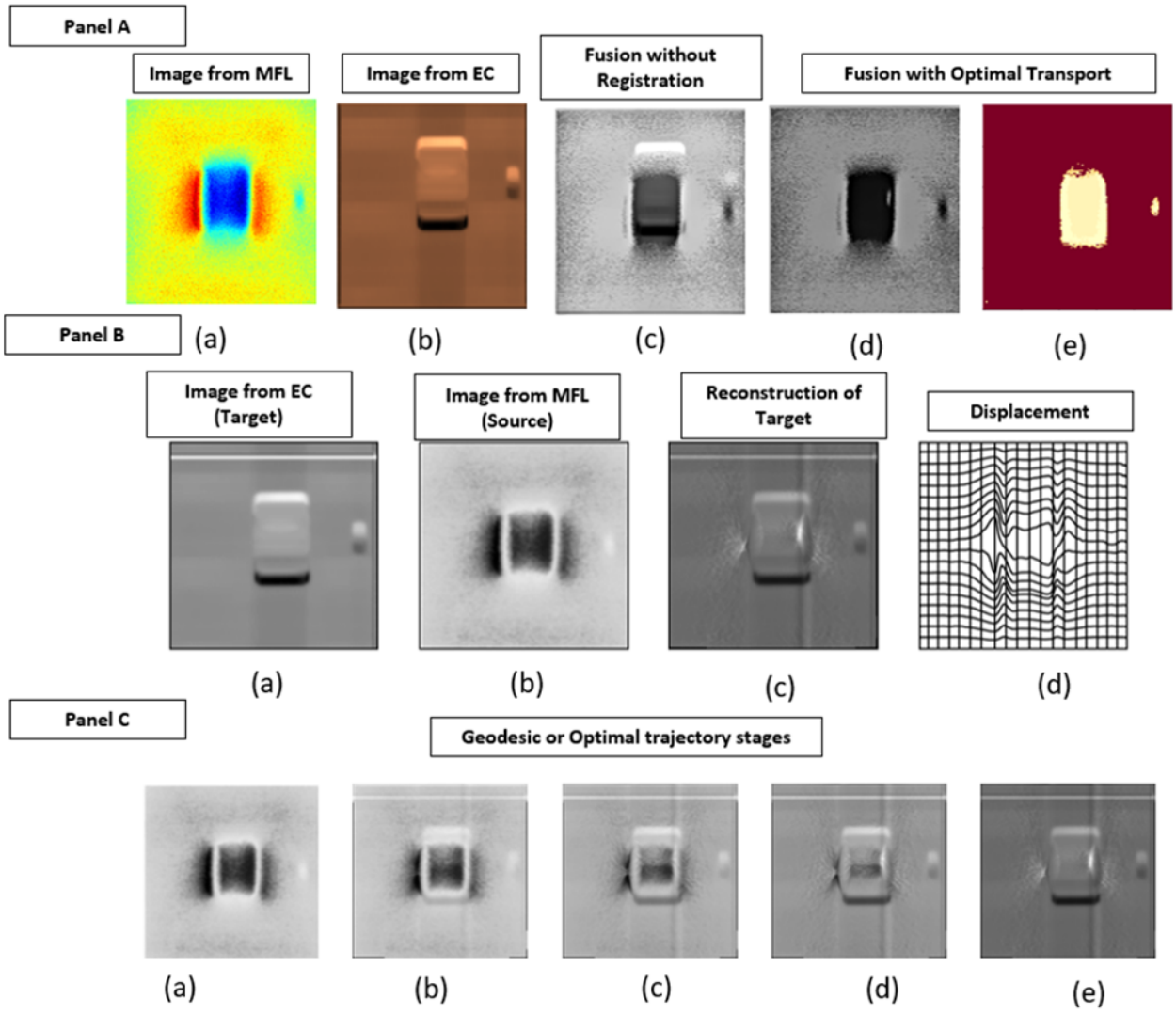

Figure 8.12 Panel A: (a) Source MFL image, (b) Source EC image, (c) result on applying fusion after performing translational registration, (d) result in applying fusion based on OT, (e) post processed results on (d), Panel B: (a) EC image which forms the target distribution, (b) MFL image acting as source distribution in OT, (e) Reconstruction of target after OT, (d) Displacement showing transfer of mass from source to target, Panel C: Optimal trajectory stages showing transfer of mass from source to target in five stages. Analysis is done based on 'Defect 16' of Table 8.3.

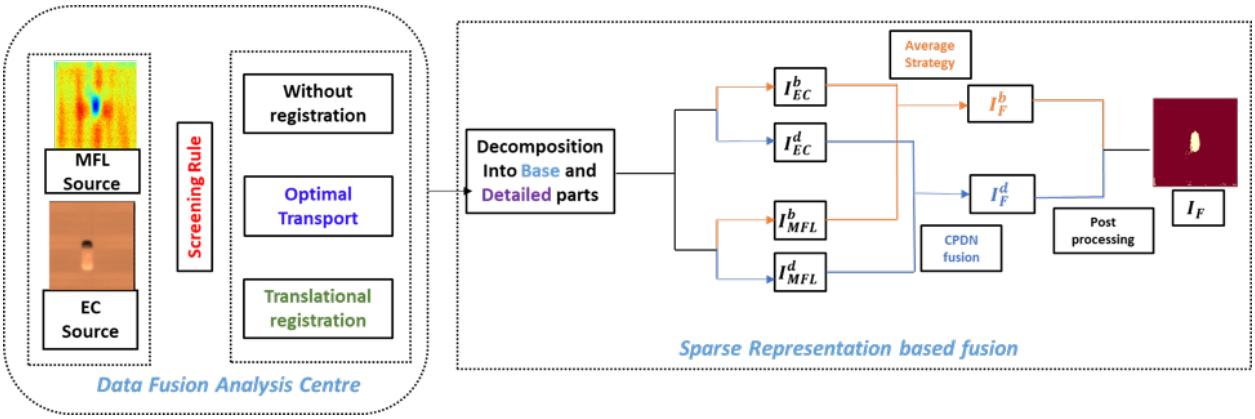

Figure 8.13 Entire framework of the sparse representation-based fusion conducted in this paper.

obtaining the base layer is as follows:

$$\arg\min_{I_k^b} \|I_k - I_k^b\|_F^2 + \alpha(\|g_i * I_k^b\|_F^2 + \|g_j * I_k^b\|_F^2)$$

where the regularization parameter is α here and $g_i = [-11], g_j = [-11]^\intercal$ are the horizontal and vertical gradient operators respectively. By solving regularization problem, base layers I_{EC}^b and I_{MFL}^b are obtained. The base layers are simply averaged to obtain the fused based layer I_F^b for every pixel (i,j)

$$I_F^b(i,j) = \frac{1}{2}\sum_{k=1}^{2} I_k^b(i,j)$$

Then the detailed layers are derived by subtracting the source image from base layers as $I_k^d = I_k - I_k^b$

Now to fuse the detailed coefficients, its sparse coefficients $x_{k,m}, m \in 1, \cdot, M$ are obtained by solving the convolutional basis pursuit denoising (CPDN) problem using alternating direction method of multipliers (ADMM) algorithm as developed in (206).

$$\arg\min_{x_{k,m}} \frac{1}{2}\|\sum_{m=1}^{M} d_m * x_m - I_k^d\|_2^2 + \lambda \sum_{m=1}^{M} \|x_{k,m}\|_1$$

The activity level detailed coefficients $x_{f,m}$ are then obtained by considering the l_1 norm of the obtained coefficients at every location (i,j). The final fused detailed layer is reconstructed as:

$$I_f^d = \sum_{m=1}^{M} d_m * x_{f,m}$$

Then finally by adding the detailed and base layer coefficients the fused image is obtained as:

$$I_f = I_f^b + I_f^d$$

Figure 8.14 shows the result of applying sparse representation-based fusion on synthetical data where the defect is of dimension 1 cm in length, 0.3mm in width and 0.5 mm in depth at a lift off of 3mm. The data being synthetically generated there is no need of registration prior to applying fusion. Figure 8.14 (d) shows the post processed fused result to be better localized and of 7 times higher SNR (83.2) than the source images (EC: 11.86, MFL:5.26).

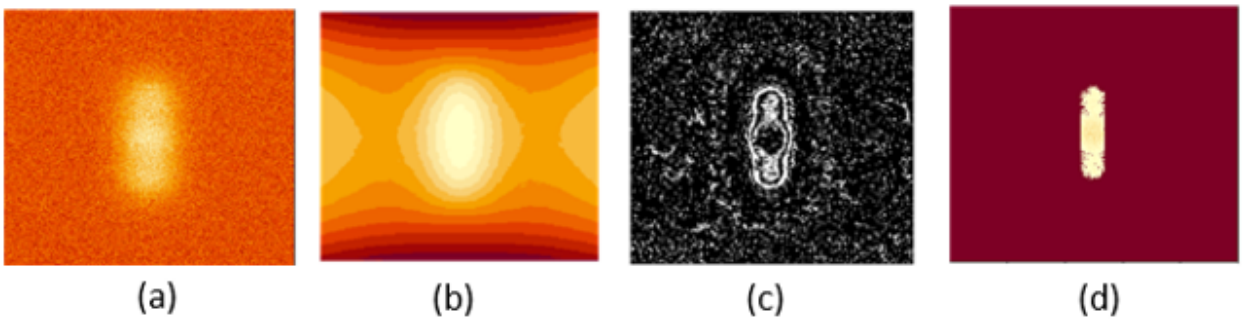

Figure 8.14 CSR fusion applied on simulation data showing better results (a) source EC, (b) source MFL, (c) fused result in gray scale, (d) post processed result.

Figure 8.15 illustrates a scenario described as 'Defect 1' in Table 8.3, wherein the MFL source image is not correctly detected and an offset is observed in the EC image. It is clear from Figure 15(d) that performing CSR fusion without registration leads to incorrect reconstruction. However, Figure 8.15(e) demonstrates improved image reconstruction after employing optimal transport-based reconstruction. Similarly, in Figure 8.16, we demonstrate the advantage of registration aided fusion on 'Defect 5' which shows significant improvement in defect localization.

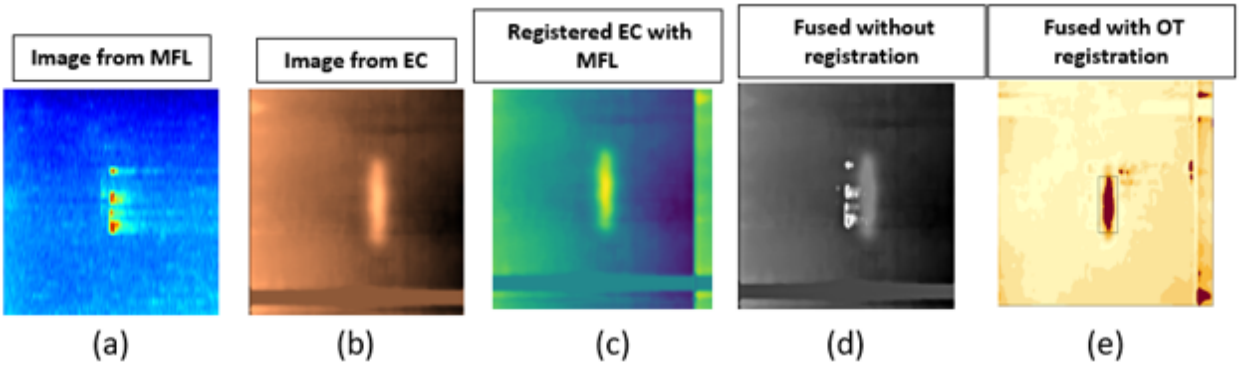

Figure 8.15 (a) Experimental MFL data based on Defect 1, (b) Experimental EC data, (c) EC registered by MFL, (d) Fused result without registration, (e)fused after OT based registration.

Table 8.4 presents the results of registration aided fusion applied on 21 experimental scenarios

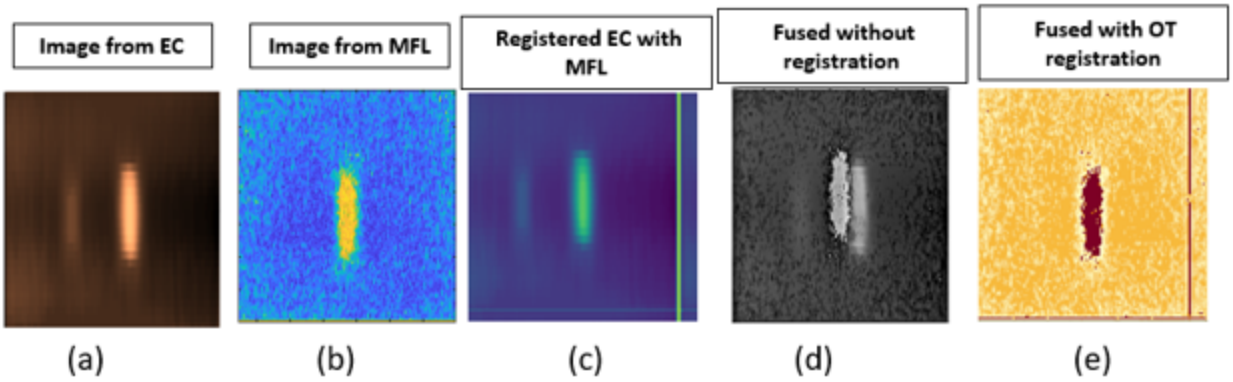

Figure 8.16 (a) Experimental EC data based on Defect 5, (b) Experimental MFL data, (c) EC registered by MFL, (d) Fused result without registration, (e)fused after OT based registration.

of Table 8.3. For all the scenarios, we have witnessed significant improvement in SNR than the SNR in fusion without registration and the SNR of source images. Here the SNR is defined as:

$$\text{SNR} = \frac{E(\text{defective area})^2}{E(\text{background})^2}$$

Table 8.4 Performance of the registration aided fusion VS non-registered fusion and source images in terms of SNR.

Defect No.	MFL Source SNR	EC Source SNR	SNR of fusion without registration	SNR of fusion with OT registration
Defect 1	20.01	18.2	17.23	29.33
Defect 2	1.87	64.35	58.56	129.27
Defect 3	8.44	3.76	5.62	12.02
Defect 4	1.13	1.2	1.53	2.5
Defect 5	4.7	5.06	10.12	22.23
Defect 6	23.08	8.03	25.19	85.18
Defect 7	6.514	0.017	1.24	8.82
Defect 8	15.27	0.045	7.43	35.06
Defect 9	40.33	6.97	51.57	172.94
Defect 10	7.63	0.25	4.12	8.23
Defect 11	39.97	6.02	35.87	164.72
Defect 12	9.55	0.97	3.72	15.12
Defect 13	8.75	0.05	3.8	9.01
Defect 14	6.5	0.97	2.5	3.69
Defect 15	3.25	1.63	2.12	6.37
Defect 16	21.29	3.9	49.18	59.14
Defect 17	2.33	0.38	1.56	2.82
Defect 18	1.05	0.012	0.47	1.51
Defect 19	43.01	2.6	51.23	73.59
Defect 20	5.47	3.78	6.37	9.32
Defect 21	2.01	0.637	1.18	2.4

8.5.2 Implementation of other common fusion algorithms

Wavelet decomposition-based fusion is a popular method for combining images from the EC and MFL techniques in the transform domain. This approach involves decomposing each source

image using an orthogonal basis, specifically employing the two-dimensional Discrete Wavelet Transform (DWT) based on Mallat's pyramidal algorithm (147). The decomposition is performed at a level of 5, resulting in an approximation image (CA) and three sets of detailed coefficient images (CV, CH, CD). To fuse the coefficients, a selection rule is applied, which can be Minimum, Maximum, or the mean of the coefficients. This leads to nine possible combinations: Min-Mean, Min-Max, Min-Min, Mean-Mean, Mean-Min, Mean-Max, Max-Min, Max-Mean, and Max-Max. For demonstration purposes, the Min-Max combination is used to showcase the superior results achieved by registration-aided fusion. After selecting the fusion combination, an Inverse Discrete Wavelet Transform (IDWT) is applied to reconstruct the fused image using the fused coefficients (CA_F, CV_F, CH_F, CD_F). Figure 8.17 provides a schematic of the fusion algorithm based on wavelet decomposition.

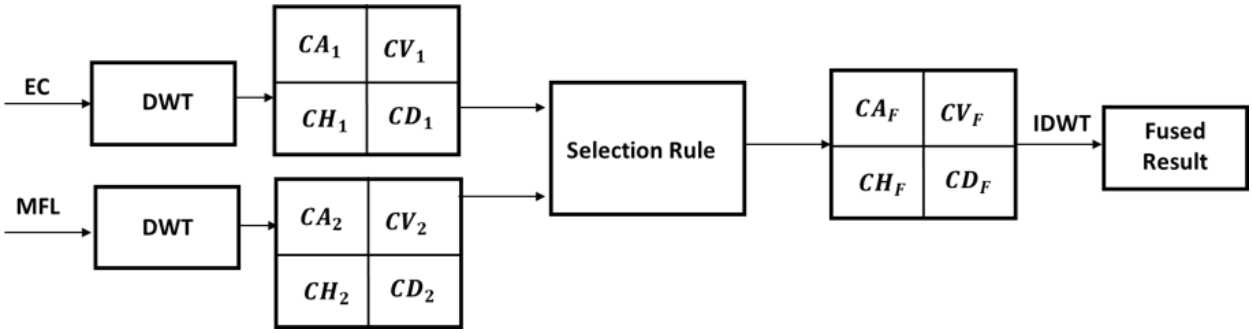

Figure 8.17 Schematic of the wavelet based fusion rule where the EC, MFL images are subjected to DWT. On the obtained coefficients, fusion based on selection rule is applied. Finally, on the fused coefficients IDWT is applied to obtain the fused image.

Applying OT based registration on 'Defect 16', shows better fusion by wavelet decomposition than on unregistered source images as evident from Figure 8.18 .

The effectiveness of the registration aided fusion is tested against a spatial domain based fusion called cross bilateral filter (CBF) (198). CBF is a kernel smoothing method which uses spatial closeness of pixels. In this technique, smoothing of EC source image is done by weights multiplied by corresponding kernel weights from MFL source. Similarly, MFL is smoothed using information from EC image. The two smoothed images are then combined to generate the fused image. Figure 8.19 illustrates the fusion results obtained by applying CBF to 'Defect 17' and 'Defect 18'.

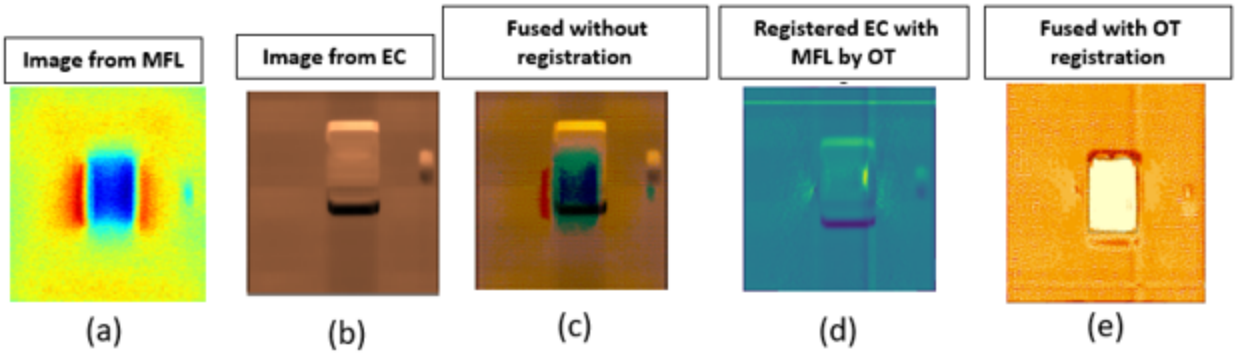

Figure 8.18 Schematic of the wavelet based fusion rule where the EC, MFL images are subjected to DWT. On the obtained coefficients, fusion based on selection rule is applied. Finally, on the fused coefficients IDWT is applied to obtain the fused image.

It is witnessed that using registered sources yields superior outcomes compared to fusion without registration. In Panel B, it is shown that at lift off of 7mm, EC fails to detect the defect as the SNR is very low. However, on transferring information from MFL source to EC by OT and then performing fusion based on CBF produced better defect identification.

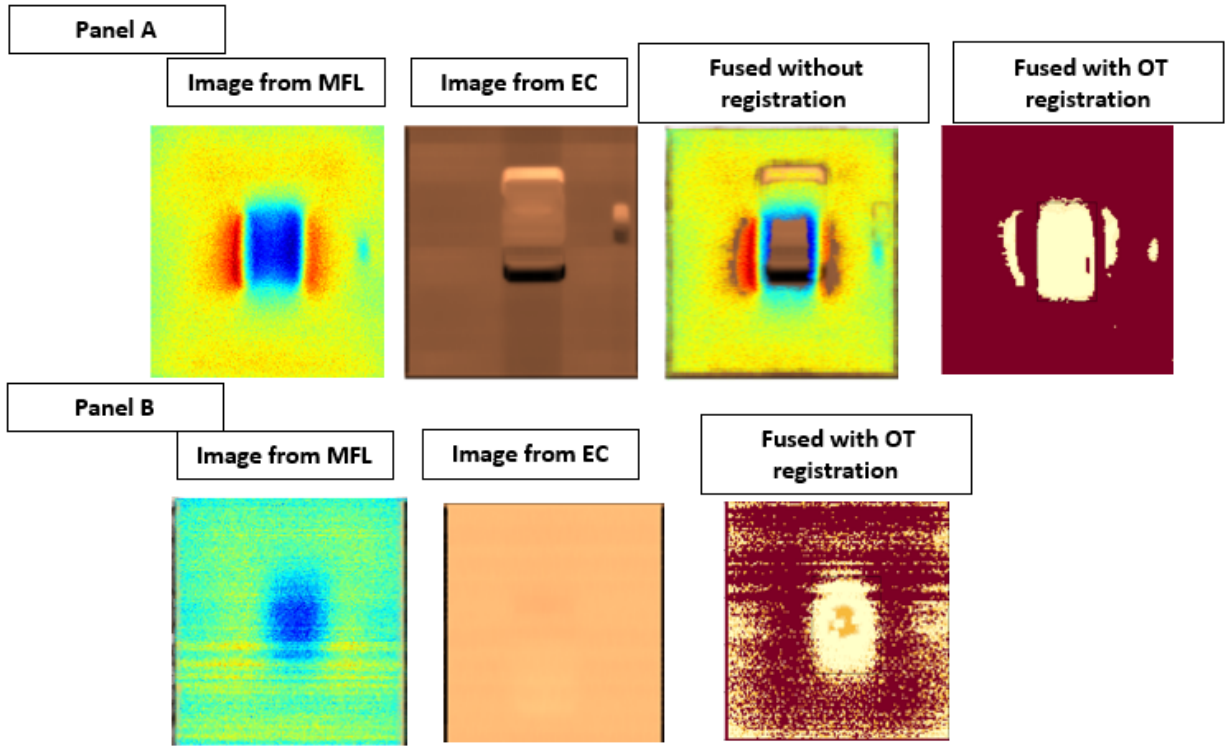

Figure 8.19 CBF based fusion works better in both Panel A (for defect 17) and Panel B (Defect 18) on registered images.

We provide some extra results of the fused images of Table 8.3 in the appendix section.

8.6 Discussion

A fully automated registration aided NDE based fusion framework has been developed that is shown to work well for defect detection. This framework demonstrates effective fusion of information from Eddy Current (EC) and Magnetic Flux Leakage (MFL) methods, resulting in a significant increase in Signal-to-Noise Ratio (SNR) and improved characterization of defects in conductive metallic surfaces. To collect experimental data, we have designed inexpensive, flexible, and miniaturized NDE probes based on MFL and EC techniques. These probes account for various uncertainties encountered during data collection. Our scanning experiments focus on detecting submillimeter and millimeter-sized crack and slit-like defects. Given the heterogeneity of the MFL and EC methods, a trade-off exists between these sensing techniques. We establish that performing fusion in boundary conditions, where one of the inspection methods fails, leads to sub-optimal fused images. To address this, we have developed a data-driven screening rule based on signal strength, which determines when fusion is appropriate. We emphasize the importance of registration when dealing with different inspection sources. Fusion without registration produces erroneous reconstructions with artifacts. To tackle this, we propose two novel image registration approaches: translational registration and optimal transport. Optimal transport allows information to be transferred from the stronger source to the weaker source based on a developed transport rule. Optimal transport exhibits promising results, particularly in boundary conditions, as it yields improved SNR. Sparse representation-based fusion is implemented on the source data collected from 21 scanning experiments. Our results illustrate that fusion performed on registered sources yields better defect identification compared to unregistered fusion. Incorporating NDE knowledge enhances the fusion framework, providing improved results. In future work, we aim to implement this fusion framework on other NDE modalities, exploring its applicability beyond EC and MFL. Additionally, we see potential for studying alternative fusion algorithms, expanding the scope of research in this area.

Portions of this chapter appeared in the publication (202).

CHAPTER 9

GENERAL CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

9.1 General Conclusion

This dissertation demonstrates new-age AI integrated NDE frameworks from data acquisition to data evaluation, thereby providing accurate,reliable, fast and smart defect diagnostics. This thesis promises to make a substantial impact, leading to the advancement of NDE 4.0 and NDE 5.0. The key contribution are hereby enlisted:

1. In chapter 2, a robust NDE detection methodology utilizing short waves based capacitive sensing is provided. A key aspect of the developed spatially adaptive denoising algorithm is its noise type agnosticism. The proposed method effectively utilizes the spatial contiguity of defective scan points, enabling accurate defect identification even in capacitive sensing data affected by uncertainties stemming from lift-off, probe tilt, and low sampling rates. A disciplined analysis based on Gaussian mixture models is also employed to accurately detect defective scan points in the sample, while maintaining low false positive misclassification rates for non-defective scan points. This enables precise identification of defect characteristics such as location, size, and shape.

2. Chapter 3 introduces a defect detection and classification procedure using laser profilometry and a crossed layered hierarchical bilinear pooling (HBP) deep learning architecture, along with other hybrid deep learning frameworks, to analyze collected images. The HBP architecture enables fine-grained visualization and effectively detects and distinguishes subtle deformities caused by different defect types in circular laser ring images. The proposed procedure outperforms conventional machine learning tools and deep learning architectures as it captures the spatio-temporal features in an efficient manner. This endoscopic methodology offers an automatic and cost-effective tool for defect localization and classification, suitable for inspecting large polymer-based pipeline systems. The laser scanning and analysis framework exhibits high flexibility in detecting defects across pipelines with varying radii

and light intensities and thus can be employed in various online non-destructive evaluation (NDE) methods that rely on image analysis of polymer materials.

3. The kriging-based method in chapter 4 can accurately identify defective areas with significantly fewer MFL readings compared to exhaustive line scans. This makes the proposed NDE method fast, cost-effective, and scalable for inspecting defects in large pipelines.

4. In chapter 5, a novel framework by applying bivariate function based estimation to the noisy data obtained from the later inspections using fast and cheap scanning has been investigated and demonstrated. In order to properly recognize the defect under noisy condition, the developed algorithms dynamically update the grid by the auxiliary information obtained from baseline inspection. The location and the size of the defect is used as a transfer learning to update the size of the defect, thereby establishing a dynamic defect tracking framework.

5. In chapter 6, another robust framework for dynamic defect detection and tracking in NDE data is developed, enabling fast and accurate analysis for periodic inspection of metallic pipelines and other inspections. The framework incorporates cheap, flexible, and miniaturized MFL and EC sensors, along with dynamic monitoring algorithms that are resilient to various noise types and uncertainties. Transfer learning-based binned hypothesis testing (TLBH) is introduced for defect growth recognition in later noisy periodic scans, demonstrating high coverage and low overestimation compared to other algorithms. The choice of optimal binning parameter is done by transfer learning where the location and size of the defects from detailed baseline inspections are transferred.

6. Chapter 7 consists of miscellaneous works where in one work we have developed a spatially adaptive framework based on nearest neighbor and wavelets to distinguish coincident defects from the harmless benign ones. In the next section, a data driven inverse approach is developed to perform material characterization based upon frequency responses. Several registration strategies are proposed to transfer information from large simulated data sets to acute experimental data sets. The results show that this novel registration aided regression

can be a promising tool to solve inverse problems in NDE getting rid of physics based requirements.

7. In Chapter 8, a fully automated fusion framework based on registration aid in NDE is developed, effectively combining information from Eddy Current (EC) and Magnetic Flux Leakage (MFL) methods for enhanced defect detection. Miniaturized NDE probes are designed for experimental data collection, considering uncertainties. The fusion process is optimized through data-driven screening, image registration techniques (translational and optimal transport), and sparse representation-based fusion, resulting in improved defect identification and increased Signal-to-Noise Ratio (SNR) in heterogeneous inspection scenarios. The incorporation of NDE knowledge further enhances the fusion framework's performance.

9.2 Future Work

This work has the potential to be used in future researches. Some possible directions are indicated below:

1. The performance of the defect defection algorithm in chapter 2 is sensitive to the threshold and neighborhood size parameters. Larger threshold value will increase coverage but also result in higher false discovery rates. The threshold value needs to be set based on applications.Fixed neighborhood shapes and sizes are considered. However data driven choices of neighborhood size can further improve this methodology. This can be incorporated by using Gaussian kernel density based filters instead of the square grids and then to adaptively choose the kernel bandwidth. There is plan to develop theoretical support for selecting these hyper-parameter in future. Our method would need some manual adaptation under varying weather conditions and humidity levels. We need the RF source to be set near the resonant frequency of the probe which was estimated to be 5 MHz for our lab experiments. If the medium between the sample and the probe is much different from our experimental setup, then the RF source needs to manually reset to the corresponding resonant frequency. The problem can be solved by using multiple probes each operating at different frequencies. However, it would require multi-parametric estimation. We aim to study it in our future

works. Another potentially useful direction of research would be to study the applicability of the spatially adaptive estimator developed in this paper for analyzing NDE data collected with uncertainties by other NDE methods.

2. In chapter 3 the prescribed method being based on 2D optical imaging of the inspected surface can not be used for estimating the depth of the defects for which 3D profiling can be useful. An interesting future work will involve extending our methodology to the 3D profile sensors. Another interesting research direction would be to progressively track the growth of erosive defects in polymer gas pipelines. This needs precise estimation of defect sizes. Using temporal correlations based tracking procedures and deep learning networks on laser profilometry data collected by the method developed in this chapter can be promising.

3. Compared to MFL readings from extensive scans there is always some information loss in the Kriging based reconstructed signal as proposed in chapter 4. The success of the procedure rests on the fact that this information loss does not hamper detection of the presence and location of the defects in the metallic surface. One future direction is to study the improvements due to non uniform sampling designs. Another interesting work will be the application of kriging in dynamic defect tracking.

4. It will be interesting to see how the dynamic defect tracking algorithms will feature in studying the interaction of multiple threats. Another potential useful direction of research is to see how TLBH and Mixture regression algorithms work on other NDE modalities where the image format is different. Role of multi modal fusion in performing defect tracking is another great problem to attack.

5. For the material characterization work, the future work can be to develop other modes of registration. It will be interesting to validate the registration aided regression models on other complicated electromagnetic models.

6. It would be fascinating to apply the optimal transport and deep learning based fusion algorithms on other heterogeneous NDE modalities.

BIBLIOGRAPHY

[1] C. Meola, S. Boccardi, G. Carlomagno, N. Boffa, F. Ricci, G. Simeoli, and P. Russo, "Impact damaging of composites through online monitoring and non-destructive evaluation with infrared thermography," *Ndt & E International*, vol. 85, pp. 34–42, 2017.

[2] Y. Xiong, Z. Liu, C. Sun, and X. Zhang, "Two-dimensional imaging by far-field superlens at visible wavelengths," *Nano letters*, vol. 7, no. 11, pp. 3360–3365, 2007.

[3] A. N. Abdalla, K. Ali, J. K. Paw, D. Rifai, and M. A. Faraj, "A novel eddy current testing error compensation technique based on mamdani-type fuzzy coupled differential and absolute probes," *Sensors*, vol. 18, no. 7, p. 2108, 2018.

[4] D. Palumbo, R. Tamborrino, U. Galietti, P. Aversa, A. Tatì, and V. Luprano, "Ultrasonic analysis and lock-in thermography for debonding evaluation of composite adhesive joints," *Ndt & E International*, vol. 78, pp. 1–9, 2016.

[5] J. Buckley and H. Loertscher, "Frequency considerations in air-coupled ultrasonic inspection," *Insight*, vol. 41, no. 11, pp. 696–9, 1999.

[6] R. Gongzhang, M. Li, T. Lardner, and A. Gachagan, "Robust defect detection in ultrasonic nondestructive evaluation (nde) of difficult materials," in *2012 IEEE International Ultrasonics Symposium*, pp. 467–470, IEEE, 2012.

[7] C. Mandache, M. Brothers, and V. Lefebvre, "Time domain lift-off compensation method for eddy current testing," *NDT. net*, vol. 10, no. 6, pp. 1–7, 2005.

[8] T. Shibata, H. Hashizume, S. Kitajima, and K. Ogura, "Experimental study on ndt method using electromagnetic waves," *Journal of materials processing technology*, vol. 161, no. 1-2, pp. 348–352, 2005.

[9] J. W. Wilson and G. Y. Tian, "3d magnetic field sensing for magnetic flux leakage defect characterisation," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 48, no. 6, pp. 357–359, 2006.

[10] S. Haugland, "Fundamental analysis of the remote-field eddy-current effect," *IEEE Transactions on Magnetics*, vol. 32, no. 4, pp. 3195–3211, 1996.

[11] G. G. Diamond and D. A. Hutchins, "A new capacitive imaging technique for ndt," in *Proceedings of European Conference on NDT*, pp. 1–8, 2006.

[12] S. Mukherjee, D. Kumar, L. Udpa, and Y. Deng, "Robust defect detection under uncertainties using spatially adaptive capacitive imaging," *Journal of Applied Physics*, vol. 131, no. 23, p. 234901, 2022.

[13] C. P. Vetter, L. A. Kuebel, D. Natarajan, and R. A. Mentzer, "Review of failure trends in the us natural gas pipeline industry: an in-depth analysis of transmission and distribution system incidents," *Journal of Loss Prevention in the Process Industries*, vol. 60, pp. 317–333, 2019.

[14] A. McMurray, C. County, M. Glendive, R. Earth, A. Creek, A. Lake, and N. Tioga, "Annex j1: Recent pipeline accidents involving crude oil in canada and the united states," *Guidance for the Environmental Public Health Management of Crude Oil Incidents*, p. 131, 2018.

[15] J. B. Harley, "Predictive guided wave models through sparse modal representations," *Proceedings of the IEEE*, vol. 104, no. 8, pp. 1604–1619, 2015.

[16] S. Sabeti and J. B. Harley, "Spatio-temporal undersampling: Recovering ultrasonic guided wavefields from incomplete data with compressive sensing," *Mechanical Systems and Signal Processing*, vol. 140, p. 106694, 2020.

[17] T. Di Ianni, L. De Marchi, A. Perelli, and A. Marzani, "Compressive sensing of full wave field data for structural health monitoring applications," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 62, no. 7, pp. 1373–1383, 2015.

[18] N. Shipway, T. Barden, P. Huthwaite, and M. Lowe, "Automated defect detection for fluorescent penetrant inspection using random forest," *NDT & E International*, vol. 101, pp. 113–123, 2019.

[19] N. Shipway, P. Huthwaite, M. Lowe, and T. Barden, "Using resnets to perform automated defect detection for fluorescent penetrant inspection," *NDT & E International*, vol. 119, p. 102400, 2021.

[20] A. Osman, U. Hassler, and V. Kaftandjian, "Automatic classification of three-dimensional segmented computed tomography data using data fusion and support vector machine," *Journal of Electronic Imaging*, vol. 21, no. 2, p. 021111, 2012.

[21] P. Zhu, Y. Cheng, P. Banerjee, A. Tamburrino, and Y. Deng, "A novel machine learning model for eddy current testing with uncertainty," *ndt & e International*, vol. 101, pp. 104–112, 2019.

[22] N. Munir, H.-J. Kim, J. Park, S.-J. Song, and S.-S. Kang, "Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions," *Ultrasonics*, vol. 94, pp. 74–81, 2019.

[23] M. A. Sutton, J. J. Orteu, and H. Schreier, *Image correlation for shape, motion and deformation measurements: basic concepts, theory and applications*. Springer Science & Business Media, 2009.

[24] N. J. Vickers, "Animal communication: when i'm calling you, will you answer too?," *Current biology*, vol. 27, no. 14, pp. R713–R715, 2017.

[25] E. J. Heravi, H. H. Aghdam, and D. Puig, "Classification of foods using spatial pyramid convolutional neural network.," in *CCIA*, pp. 163–168, 2016.

[26] D. E. Bray and R. K. Stanley, *Nondestructive evaluation: A tool in design, manufacturing, and service*. CRC press, 2018.

[27] N. Gloria, M. Areiza, I. Miranda, and J. Rebello, "Development of a magnetic sensor for detection and sizing of internal pipeline corrosion defects," *NDT & e International*, vol. 42, no. 8, pp. 669–677, 2009.

[28] S. Mukherjee, X. Huang, L. Udpa, and Y. Deng, "Nde based cost-effective detection of obtrusive and coincident defects in pipelines under uncertainties," in *2019 Prognostics and System Health Management Conference (PHM-Paris)*, pp. 297–302, IEEE, 2019.

[29] S. Mukherjee, X. Huang, V. T. Rathod, L. Udpa, and Y. Deng, "Defects tracking via nde based transfer learning," in *2020 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 1–8, IEEE, 2020.

[30] R. Wagner, O. Goncalves, A. Demma, and M. Lowe, "Guided wave testing performance studies: comparison with ultrasonic and magnetic flux leakage pigs," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 55, no. 4, pp. 187–196, 2013.

[31] E. Mohseni, H. Habibzadeh Boukani, D. Ramos França, and M. Viens, "A study of the automated eddy current detection of cracks in steel plates," *Journal of Nondestructive Evaluation*, vol. 39, no. 1, pp. 1–12, 2020.

[32] T. Gan, D. Hutchins, D. Billson, and D. Schindel, "The use of broadband acoustic transducers and pulse-compression techniques for air-coupled ultrasonic imaging," *Ultrasonics*, vol. 39, no. 3, pp. 181–194, 2001.

[33] H. Kriesz, "Radiographic ndt—a review," *NDT International*, vol. 12, no. 6, pp. 270–273, 1979.

[34] T. Venkatsubramanian and B. Unvala, "An ac potential drop system for monitoring crack length," *Journal of physics E: scientific instruments*, vol. 17, no. 9, p. 765, 1984.

[35] D. Jiles, "Review of magnetic methods for nondestructive evaluation," *NDT international*, vol. 21, no. 5, pp. 311–319, 1988.

[36] E. Mohseni and M. Viens, "Sensitivity of eddy current signals to probe's tilt and lift-off while scanning semi-elliptical surface notches-a finite element modeling approach," 2017.

[37] J. Lim, *Data fusion for NDE signal characterization*. Iowa State University, 2001.

[38] G. Y. Tian and A. Sophian, "Reduction of lift-off effects for pulsed eddy current ndt," *NDT*

& E International, vol. 38, no. 4, pp. 319–324, 2005.

[39] H. Hoshikawa, K. Koyama, and H. Karasawa, "A new eddy current surface probe without lift-off noise," in *Proc. 10th APCNDT*, pp. 275–8575, 2001.

[40] A. McNab and J. Thomson, "An eddy current array instrument for application on ferritic welds," *NDT & E International*, vol. 28, no. 2, pp. 103–112, 1995.

[41] B. Rao, B. Raj, T. Jayakumar, and P. Kalyanasundaram, "An artificial neural network for eddy current testing of austenitic stainless steel welds," *Ndt & E International*, vol. 35, no. 6, pp. 393–398, 2002.

[42] D. He and M. Yoshizawa, "Dual-frequency eddy-current nde based on high-tc rf squid," *Physica C: Superconductivity*, vol. 383, no. 3, pp. 223–226, 2002.

[43] Y. Fu, M. Lei, Z. Li, Z. Gu, H. Yang, A. Cao, and J. Sun, "Lift-off effect reduction based on the dynamic trajectories of the received-signal fast fourier transform in pulsed eddy current testing," *NDT & E International*, vol. 87, pp. 85–92, 2017.

[44] Y. He, M. Pan, F. Luo, and G. Tian, "Reduction of lift-off effects in pulsed eddy current for defect classification," *IEEE Transactions on Magnetics*, vol. 47, no. 12, pp. 4753–4760, 2011.

[45] T. Chen and N. Bowler, "A rotationally invariant capacitive probe for materials evaluation," *Materials Evaluation*, vol. 70, no. 2, pp. 161–172, 2012.

[46] B. Auld, A. Clark, S. Schaps, and P. Heyliger, "Capacitive probe array measurements and limitations," *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 1063–1070, 1993.

[47] T. Chen and N. Bowler, "Design of interdigital spiral and concentric capacitive sensors for materials evaluation," in *AIP Conference Proceedings*, vol. 1511, pp. 1593–1600, American Institute of Physics, 2013.

[48] X. Yin, C. Li, Z. Li, W. Li, and G. Chen, "Lift-off effect for capacitive imaging sensors," *Sensors*, vol. 18, no. 12, p. 4286, 2018.

[49] M. Morozov, W. Jackson, and S. Pierce, "Capacitive imaging of impact damage in composite material," *Composites Part B: Engineering*, vol. 113, pp. 65–71, 2017.

[50] D. Kumar, S. Karuppuswami, Y. Deng, and P. Chahal, "A wireless shortwave near-field probe for monitoring structural integrity of dielectric composites and polymers," *NDT & E International*, vol. 96, pp. 9–17, 2018.

[51] S. Mukherjee, X. Huang, L. Udpa, and Y. Deng, "A kriging based fast and efficient method for

defect detection in massive pipelines using magnetic flux leakages," in *ASME International Mechanical Engineering Congress and Exposition*, vol. 84669, p. V014T14A010, American Society of Mechanical Engineers, 2020.

[52] S. Mukherjee, X. Huang, L. Udpa, and Y. Deng, "A kriging-based magnetic flux leakage method for fast defect detection in massive pipelines," *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, vol. 5, no. 1, 2022.

[53] W. Sun and T. T. Cai, "Oracle and adaptive compound decision rules for false discovery rate control," *Journal of the American Statistical Association*, vol. 102, no. 479, pp. 901–912, 2007.

[54] O. Muralidharan, "An empirical bayes mixture method for effect size and false discovery rate estimation," *The Annals of Applied Statistics*, pp. 422–438, 2010.

[55] Y. Benjamini, "Discovering the false discovery rate," *Journal of the Royal Statistical Society: series B (statistical methodology)*, vol. 72, no. 4, pp. 405–416, 2010.

[56] B. Efron, *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*, vol. 1. Cambridge University Press, 2012.

[57] T. T. Cai, W. Sun, and Y. Xia, "Laws: A locally adaptive weighting and screening approach to spatial multiple testing," *Journal of the American Statistical Association*, pp. 1–14, 2021.

[58] E. J. Candes, C. A. Sing-Long, and J. D. Trzasko, "Unbiased risk estimates for singular value thresholding and spectral estimators," *IEEE transactions on signal processing*, vol. 61, no. 19, pp. 4643–4657, 2013.

[59] J. Josse, S. Sardy, and S. Wager, "denoiser: A package for low rank matrix estimation," *arXiv preprint arXiv:1602.01206*, 2016.

[60] J. Josse and S. Sardy, "Adaptive shrinkage of singular values," *Statistics and Computing*, vol. 26, no. 3, pp. 715–724, 2016.

[61] K. Fukushima, R. Maeshima, A. Kinoshita, H. Shiraishi, and I. Koshijima, "Gas pipeline leak detection system using the online simulation method," *Computers & chemical engineering*, vol. 24, no. 2-7, pp. 453–456, 2000.

[62] R. Fullenbaum, J. Fallon, and B. Flanagan, "Oil & natural gas transportation & storage infrastructure: Status, trends, & economic benefits," *Washington: IHS Global Inc*, 2013.

[63] N. Merah, F. Saghir, Z. Khan, and A. Bazoune, "Effect of temperature on tensile properties of hdpe pipe material," *Plastics, rubber and composites*, vol. 35, no. 5, pp. 226–230, 2006.

[64] R. W. Tucker Jr, S. W. Kercel, and V. K. Varma, "Characterization of gas pipeline flaws

using wavelet analysis," in *Sixth International Conference on Quality Control by Artificial Vision*, vol. 5132, pp. 485–493, International Society for Optics and Photonics, 2003.

[65] O. Duran, K. Althoefer, and L. D. Seneviratne, "State of the art in sensor technologies for sewer inspection," *IEEE Sensors journal*, vol. 2, no. 2, pp. 73–81, 2002.

[66] O. Duran, K. Althoefer, and L. D. Seneviratne, "Pipe inspection using a laser-based transducer and automated analysis techniques," *IEEE/ASME Transactions on mechatronics*, vol. 8, no. 3, pp. 401–409, 2003.

[67] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2017.

[68] S. Iyer and S. K. Sinha, "A robust approach for automatic detection and segmentation of cracks in underground pipeline images," *Image and Vision Computing*, vol. 23, no. 10, pp. 921–933, 2005.

[69] T. Vidal-Calleja, J. V. Miró, F. Martín, D. C. Lingnau, and D. E. Russell, "Automatic detection and verification of pipeline construction features with multi-modal data," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3116–3122, IEEE, 2014.

[70] C. Ajmi, J. Zapata, J. J. Martínez-Álvarez, G. Doménech, and R. Ruiz, "Using deep learning for defect classification on a small weld x-ray image dataset," *Journal of Nondestructive Evaluation*, vol. 39, no. 3, pp. 1–13, 2020.

[71] S. Tang and Z. Chen, "Scale–space data augmentation for deep transfer learning of crack damage from small sized datasets," *Journal of Nondestructive Evaluation*, vol. 39, no. 3, pp. 1–18, 2020.

[72] I. Virkkunen, T. Koskinen, O. Jessen-Juhler, and J. Rinta-Aho, "Augmented ultrasonic data for machine learning," *Journal of Nondestructive Evaluation*, vol. 40, no. 1, pp. 1–11, 2021.

[73] C. Yu, X. Zhao, Q. Zheng, P. Zhang, and X. You, "Hierarchical bilinear pooling for fine-grained visual recognition," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 574–589, 2018.

[74] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. De Geus, "Malicious software classification using transfer learning of resnet-50 deep neural network," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 1011–1014, IEEE, 2017.

[75] R. D. Roberts, "Laser profilometry as an inspection method for reformer catalyst tubes," *Materials evaluation*, vol. 57, no. 4, pp. 420–422, 1999.

[76] G. Teza, A. Galgaro, and F. Moro, "Contactless recognition of concrete surface damage from laser scanning and curvature computation," *NDT & E International*, vol. 42, no. 4, pp. 240–249, 2009.

[77] J. E. Romaine and E. I. Chaleff, "Object dimensioning using scanned laser and linear ccd triangulation," in *Optical Scanning Systems: Design and Applications*, vol. 3131, pp. 232–240, International Society for Optics and Photonics, 1997.

[78] Y.-w. Yu, G.-f. Yin, Y. Yin, *et al.*, "Defect recognition for radiographic image based on deep learning network," *Chin. J. Sci. Instrum*, vol. 35, pp. 2012–2019, 2014.

[79] D. Gamdha, S. Unnikrishnakurup, K. J. Rose, M. Surekha, P. Purushothaman, B. Ghose, and K. Balasubramaniam, "Automated defect recognition on x-ray radiographs of solid propellant using deep learning based on convolutional neural networks," *Journal of Nondestructive Evaluation*, vol. 40, no. 1, pp. 1–13, 2021.

[80] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for lvcsr using rectified linear units and dropout," in *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 8609–8613, IEEE, 2013.

[81] R. Wirahadikusumah, D. M. Abraham, T. Iseley, and R. K. Prasanth, "Assessment technologies for sewer system rehabilitation," *Automation in Construction*, vol. 7, no. 4, pp. 259–270, 1998.

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

[83] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[84] M. E. Taylor and P. Stone, "Transfer learning for reinforcement learning domains: A survey.," *Journal of Machine Learning Research*, vol. 10, no. 7, 2009.

[85] Y. Gong, J. Luo, H. Shao, K. He, and W. Zeng, "Automatic defect detection for small metal cylindrical shell using transfer learning and logistic regression," *Journal of Nondestructive Evaluation*, vol. 39, no. 1, pp. 1–13, 2020.

[86] D. Xie, L. Zhang, and L. Bai, "Deep learning in visual computing and signal processing," *Applied Computational Intelligence and Soft Computing*, vol. 2017, 2017.

[87] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.

[88] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. "

O'Reilly Media, Inc.", 2008.

[89] K. Team, "Keras documentation: Keras applications," *Inglés. URL: https://keras.io/api/applications*, 2020.

[90] X. Zhang, H. Xiong, W. Zhou, W. Lin, and Q. Tian, "Picking deep filter responses for fine-grained image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1134–1142, 2016.

[91] H. Zhang, T. Xu, M. Elhoseiny, X. Huang, S. Zhang, A. Elgammal, and D. Metaxas, "Spda-cnn: Unifying semantic part detection and abstraction for fine-grained recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1143–1152, 2016.

[92] T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear cnn models for fine-grained visual recognition," in *Proceedings of the IEEE international conference on computer vision*, pp. 1449–1457, 2015.

[93] T.-Y. Lin and S. Maji, "Improved bilinear pooling with cnns," *arXiv preprint arXiv:1707.06772*, 2017.

[94] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, pp. 618–626, 2017.

[95] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2921–2929, 2016.

[96] H. Gao, L. Dou, W. Chen, and J. Sun, "Image classification with bag-of-words model based on improved sift algorithm," in *2013 9th Asian Control Conference (ASCC)*, pp. 1–6, IEEE, 2013.

[97] N. Shipway, P. Huthwaite, M. Lowe, and T. Barden, "Performance based modifications of random forest to perform automated defect detection for fluorescent penetrant inspection," *Journal of Nondestructive Evaluation*, vol. 38, no. 2, pp. 1–11, 2019.

[98] H. R. Sofaer, J. A. Hoeting, and C. S. Jarnevich, "The area under the precision-recall curve as a performance metric for rare binary events," *Methods in Ecology and Evolution*, vol. 10, no. 4, pp. 565–577, 2019.

[99] S. Mukherjee, R. Zhang, M. Alzuhiri, V. V. Rao, L. Udpa, and Y. Deng, "Inline pipeline inspection using hybrid deep learning aided endoscopic laser profiling," *Journal of Nondestructive Evaluation*, vol. 41, no. 3, pp. 1–13, 2022.

[100] G. Matheron, "Principles of geostatistics," *Economic geology*, vol. 58, no. 8, pp. 1246–1266, 1963.

[101] F. Jiang, S. Liu, S. Xin, and H. Zhang, "Electromagnetic nondestructive testing model and surface magnetic field analysis for circumferential cracks on metal rod," *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, vol. 2, no. 4, 2019.

[102] M. Afzal and S. Udpa, "Advanced signal processing of magnetic flux leakage data obtained from seamless gas pipeline," *Ndt & E International*, vol. 35, no. 7, pp. 449–457, 2002.

[103] A. Sophian, G. Y. Tian, and S. Zairi, "Pulsed magnetic flux leakage techniques for crack detection and characterisation," *Sensors and Actuators A: Physical*, vol. 125, no. 2, pp. 186–191, 2006.

[104] M. Prabhakaran, P. Karuppasamy, A. Abudhahir, and B. Rao, "Modeling of mfl system to inspect ferro magnetic tubes," *Int. J. Comput. Appl. Technol.*, vol. 975, p. 8887, 2013.

[105] P. Ivanov, V. Zhang, C. Yeoh, H. Udpa, Y. Sun, S. Udpa, and W. Lord, "Magnetic flux leakage modeling for mechanical damage in transmission pipelines," *IEEE Transactions on magnetics*, vol. 34, no. 5, pp. 3020–3023, 1998.

[106] F. Al-Naemi, J. P. Hall, and A. J. Moses, "Fem modelling techniques of magnetic flux leakage-type ndt for ferromagnetic plate inspections," *Journal of Magnetism and Magnetic Materials*, vol. 304, no. 2, pp. e790–e793, 2006.

[107] J. Aguila-Munoz, J. H. Espina-Hernandez, J. A. Perez-Benitez, F. Caleyo, and J. Hallen, "Crack detection in steel using a gmr-based mfl probe with radial magnetization," in *CONIELECOMP 2013, 23rd International Conference on Electronics, Communications and Computing*, pp. 104–108, IEEE, 2013.

[108] L. Chen, P.-W. Que, and T. Jin, "A giant-magnetoresistance sensor for magnetic-flux-leakage nondestructive testing of a pipeline," *Russian Journal of Nondestructive Testing*, vol. 41, no. 7, pp. 462–465, 2005.

[109] D. J. Rogers and L. Sedda, "Statistical models for spatially explicit biological data," *Parasitology*, vol. 139, no. 14, pp. 1852–1869, 2012.

[110] D. Zimmerman, C. Pavlik, A. Ruggles, and M. P. Armstrong, "An experimental comparison of ordinary and universal kriging and inverse distance weighting," *Mathematical Geology*, vol. 31, no. 4, pp. 375–390, 1999.

[111] E. J. Pebesma and C. G. Wesseling, "Gstat: a program for geostatistical modelling, prediction and simulation," *Computers & Geosciences*, vol. 24, no. 1, pp. 17–31, 1998.

[112] X. Du and L. Leifsson, "Efficient uncertainty propagation for mapod via polynomial chaos-based kriging," *Engineering Computations*, 2019.

[113] E. J. Pebesma, "Multivariable geostatistics in s: the gstat package," *Computers & geosciences*, vol. 30, no. 7, pp. 683–691, 2004.

[114] E. Pebesma and R. S. Bivand, "S classes and methods for spatial data: the sp package," *R news*, vol. 5, no. 2, pp. 9–13, 2005.

[115] G. Piao, J. Guo, T. Hu, H. Leung, and Y. Deng, "Fast reconstruction of 3-d defect profile from mfl signals using key physics-based parameters and svm," *NDT & E International*, vol. 103, pp. 26–38, 2019.

[116] G. Piao, J. Guo, T. Hu, and H. Leung, "The effect of motion-induced eddy current on high-speed magnetic flux leakage (mfl) inspection for thick-wall steel pipe," *Research in Nondestructive Evaluation*, vol. 31, no. 1, pp. 48–67, 2020.

[117] A. L. Pullen, P. C. Charlton, N. R. Pearson, and N. J. Whitehead, "Magnetic flux leakage scanning velocities for tank floor inspection," *IEEE Transactions on Magnetics*, vol. 54, no. 9, pp. 1–8, 2018.

[118] C. Xu, C. Wang, F. Ji, and X. Yuan, "Finite-element neural network-based solving 3-d differential equations in mfl," *IEEE Transactions on Magnetics*, vol. 48, no. 12, pp. 4747–4756, 2012.

[119] A. Khodayari-Rostamabad, J. P. Reilly, N. K. Nikolova, J. R. Hare, and S. Pasha, "Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection," *IEEE Transactions on magnetics*, vol. 45, no. 8, pp. 3073–3084, 2009.

[120] M. Afzal, R. Polikar, L. Udpa, and S. Udpa, "Adaptive noise cancellation schemes for magnetic flux leakage signals obtained from gas pipeline inspection," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, vol. 5, pp. 3389–3392, IEEE, 2001.

[121] A. Joshi, L. Udpa, S. Udpa, and A. Tamburrino, "Adaptive wavelets for characterizing magnetic flux leakage signals from pipeline inspection," *IEEE transactions on magnetics*, vol. 42, no. 10, pp. 3168–3170, 2006.

[122] P. Ramuhalli, L. Udpa, and S. S. Udpa, "Neural network-based inversion algorithms in magnetic flux leakage nondestructive evaluation," *Journal of applied physics*, vol. 93, no. 10, pp. 8274–8276, 2003.

[123] H. Zuoying, Q. Peiwen, and C. Liang, "3d fem analysis in magnetic flux leakage method," *Ndt & E International*, vol. 39, no. 1, pp. 61–66, 2006.

[124] J. A. Parra-Raad and S. Roa-Prada, "Multi-objective optimization of a magnetic circuit for magnetic flux leakage-type non-destructive testing," *Journal of Nondestructive Evaluation*, vol. 35, no. 1, pp. 1–12, 2016.

[125] B. Muthén and K. Shedden, "Finite mixture modeling with mixture outcomes using the em algorithm," *Biometrics*, vol. 55, no. 2, pp. 463–469, 1999.

[126] L. Scrucca, M. Fop, T. B. Murphy, and A. E. Raftery, "mclust 5: clustering, classification and density estimation using gaussian finite mixture models," *The R journal*, vol. 8, no. 1, p. 289, 2016.

[127] Y. Bengio, "Deep learning of representations for unsupervised and transfer learning," in *Proceedings of ICML workshop on unsupervised and transfer learning*, pp. 17–36, JMLR Workshop and Conference Proceedings, 2012.

[128] T. Benaglia, D. Chauveau, D. R. Hunter, and D. S. Young, "mixtools: an r package for analyzing mixture models," *Journal of statistical software*, vol. 32, pp. 1–29, 2010.

[129] S. Mukherjee, X. Huang, L. Udpa, and Y. Deng, "Dynamic tracking of defects in pipelines via nde based transfer learning," *Review of Progress in Quantitative Nondestructive Evaluation*, 2019.

[130] H. M. Kim and G. S. Park, "A new sensitive excitation technique in nondestructive inspection for underground pipelines by using differential coils," *IEEE Transactions on Magnetics*, vol. 53, no. 11, pp. 1–4, 2017.

[131] Z. Li, X. Huang, O. Elshafiey, S. Mukherjee, and Y. Deng, "Fem of magnetic flux leakage signal for uncertainty estimation in crack depth classification using bayesian convolutional neural network and deep ensemble," in *2021 International Applied Computational Electromagnetics Society Symposium (ACES)*, pp. 1–4, IEEE, 2021.

[132] J. Lim, S. S. Udpa, L. Udpa, and M. Afzal, "Multisensor fusion for 3-d defect characterization using wavelet basis function neural networks," in *AIP Conference Proceedings*, vol. 557, pp. 679–686, American Institute of Physics, 2001.

[133] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[134] P. H. Westfall and S. S. Young, *Resampling-based multiple testing: Examples and methods for p-value adjustment*, vol. 279. John Wiley & Sons, 1993.

[135] G. Casella and R. L. Berger, *Statistical inference*. Cengage Learning, 2021.

[136] E. Gosset and B. Louis, "The binning analysis: Towards a better significance test," *Astrophysics and space science*, vol. 120, pp. 263–306, 1986.

[137] C. Zhang and B. Cheng, "Binning methodology for nonparametric goodness-of-fit test," *Journal of Statistical Computation and Simulation*, vol. 73, no. 1, pp. 71–82, 2003.

[138] G. Katz, P. Piantanida, R. Couillet, and M. Debbah, "On the necessity of binning for the distributed hypothesis testing problem," in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 2797–2801, IEEE, 2015.

[139] S. Watanabe, "On sub-optimality of random binning for distributed hypothesis testing," in *2022 IEEE International Symposium on Information Theory (ISIT)*, pp. 2708–2713, IEEE, 2022.

[140] M. S. Rahman and A. B. Wagner, "On the optimality of binning for distributed hypothesis testing," *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6282–6303, 2012.

[141] W. Sun, B. J. Reich, T. Tony Cai, M. Guindani, and A. Schwartzman, "False discovery control in large-scale spatial multiple testing," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 77, no. 1, pp. 59–83, 2015.

[142] L. Deng, K. He, and X. Zhang, "Powerful spatial multiple testing via borrowing neighboring information," *arXiv preprint arXiv:2210.17121*, 2022.

[143] T. T. Cai, W. Sun, and Y. Xia, "Laws: A locally adaptive weighting and screening approach to spatial multiple testing," *Journal of the American Statistical Association*, vol. 117, no. 539, pp. 1370–1383, 2022.

[144] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE transactions on information theory*, vol. 13, no. 1, pp. 21–27, 1967.

[145] S. Cost and S. Salzberg, "A weighted nearest neighbor algorithm for learning with symbolic features," *Machine learning*, vol. 10, no. 1, pp. 57–78, 1993.

[146] S. Mallat, *A wavelet tour of signal processing*. Elsevier, 1999.

[147] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[148] L. Evers and T. Heaton, "Locally adaptive tree-based thresholding," *Journal of Computational and Graphical Statistics*, vol. 18, no. 4, pp. 961–977, 2009.

[149] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the american statistical association*, vol. 90, no. 432, pp. 1200–1224, 1995.

[150] I. M. Johnstone and B. W. Silverman, "Empirical bayes selection of wavelet thresholds," *The Annals of Statistics*, vol. 33, no. 4, pp. 1700–1752, 2005.

[151] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *Journal of the ACM (JACM)*, vol. 45, no. 6, pp. 891–923, 1998.

[152] L. Evers and T. Heaton, "Locally adaptive tree-based thresholding using the treethresh package in r," *Journal of Statistical Software*, vol. 78, pp. 1–22, 2017.

[153] D. M. Pozar, *Microwave engineering*. John wiley & sons, 2011.

[154] L. Chen, C. Ong, C. Neo, V. Varadan, and V. Varadan, "Microwave theory and techniques for materials characterization," *Microwave Electronics: Measurement and Materials Characterization*, pp. 37–141, 2004.

[155] N. Soleimani, R. Trinchero, and F. Canavero, "Application of different learning methods for the modelling of microstrip characteristics," in *2020 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)*, pp. 1–3, IEEE, 2020.

[156] K. Sharma and G. P. Pandey, "Designing a compact microstrip antenna using the machine learning approach," *Journal of Telecommunications and Information Technology*, 2020.

[157] N. T. Tokan and F. Gunes, "Support vector characterisation of the microstrip antennas based on measurements," *Progress In Electromagnetics Research B*, vol. 5, pp. 49–61, 2008.

[158] Y. Zhu, Y. Xie, N. Bai, and X. Sun, "Inverse design of a microstrip meander line slow wave structure with xgboost and neural network," *Electronics*, vol. 10, no. 19, p. 2430, 2021.

[159] J. S. Morris, "Functional regression," *Annual Review of Statistics and Its Application*, vol. 2, pp. 321–359, 2015.

[160] A. Antoniadis and T. Sapatinas, "Estimation and inference in functional mixed-effects models," *Computational statistics & data analysis*, vol. 51, no. 10, pp. 4793–4813, 2007.

[161] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: an approach to regression analysis by local fitting," *Journal of the American statistical association*, vol. 83, no. 403, pp. 596–610, 1988.

[162] B. A. Brumback and J. A. Rice, "Smoothing spline models for the analysis of nested and crossed samples of curves," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 961–976, 1998.

[163] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of statistical software*, vol. 33, no. 1, p. 1, 2010.

[164] J. O. Ramsay and B. W. Silverman, "Functional data analysis," *İnternet Adresi: http*, 2008.

[165] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[166] S. Mukherjee, D. Kumar, O. Elshafiey, L. Udpa, and Y. Deng, "Accurate material characterization of wideband rf signals via registration-based curve fitting model using microstrip transmission line," in *2023 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 139–145, IEEE, 2023.

[167] X. Gros, *NDT data fusion*. Elsevier, 1996.

[168] R. Heideklang and P. Shokouhi, "Multi-sensor image fusion at signal level for improved near-surface crack detection," *Ndt & E International*, vol. 71, pp. 16–22, 2015.

[169] S. Udpa and L. Udpa, "Eddy current testing-are we at the limits," *Proceedings of 16th WCNDT, Canada, Montreal, 31st August-3rd September*, 2004.

[170] M. Sasikala and N. Kumaravel, "A comparative analysis of feature based image fusion methods," *Information Technology Journal*, vol. 6, no. 8, pp. 1224–1230, 2007.

[171] Q. Tao and R. Veldhuis, "Threshold-optimized decision-level fusion and its application to biometrics," *Pattern Recognition*, vol. 42, no. 5, pp. 823–836, 2009.

[172] I. Elshafiey, M. Alkanhal, and A. Algarni, "Image fusion based enhancement of eddy current nondestructive evaluation," *International journal of applied electromagnetics and mechanics*, vol. 28, no. 1-2, pp. 291–296, 2008.

[173] Z. Liu, P. Ramuhalli, S. Safizadeh, and D. S. Forsyth, "Combining multiple nondestructive inspection images with a generalized additive model," *Measurement Science and Technology*, vol. 19, no. 8, p. 085701, 2008.

[174] R. Heideklang and P. Shokouhi, "Application of data fusion in nondestructive testing (ndt)," in *Proceedings of the 16th International Conference on Information Fusion*, pp. 835–841, IEEE, 2013.

[175] B. Yilmaz, A. Ba, E. Jasiuniene, H.-K. Bui, and G. Berthiau, "Evaluation of bonding quality with advanced nondestructive testing (ndt) and data fusion," *Sensors*, vol. 20, no. 18, p. 5127, 2020.

[176] P. Daryabor and M. Safizadeh, "Image fusion of ultrasonic and thermographic inspection of carbon/epoxy patches bonded to an aluminum plate," *Ndt & E International*, vol. 90, pp. 1–10, 2017.

[177] Q. Song, "Data fusion for mfl signal characterization," in *Applied Mechanics and Materials*, vol. 44, pp. 3519–3523, Trans Tech Publ, 2011.

[178] M. Mina, J. Yim, S. Udpa, L. Udpa, W. Lord, and K. Sun, "Two dimensional multi-frequency

eddy current data fusion," in *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 2125–2132, Springer, 1996.

[179] X. Gros, J. Bousigue, and K. Takahashi, "Ndt data fusion at pixel level," *Ndt & E International*, vol. 32, no. 5, pp. 283–292, 1999.

[180] D. Horn and W. R. Mayo, "Nde reliability gains from combining eddy-current and ultrasonic testing," *NDT & e International*, vol. 33, no. 6, pp. 351–362, 2000.

[181] J. Yim, S. Udpa, M. Mina, and L. Udpa, "Optimum filter based techniques for data fusion," in *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 773–780, Springer, 1996.

[182] Y.-W. Song, *NDE data fusion using morphological approaches*. Iowa State University, 1997.

[183] Y.-W. Song and S. Udpa, "A new morphological algorithm for fusing ultrasonic and eddy current images," in *1996 IEEE Ultrasonics Symposium. Proceedings*, vol. 1, pp. 649–652, IEEE, 1996.

[184] L. Udpa, "Neural networks for nde," in *Proc. IV Int. Workshop: Advances Signal Process. Nondestructive Eval. Mater. Quebec City, QC, Canada*, 2001.

[185] J. Yim, S. Udpa, L. Udpa, M. Mina, and W. Lord, "Neural network approaches to data fusion," in *Review of Progress in Quantitative Nondestructive Evaluation*, pp. 819–826, Springer, 1995.

[186] J. A. Oagaro and S. Mandayam, "Multi-sensor data fusion using geometric transformations for gas transmission pipeline inspection," in *2008 IEEE Instrumentation and Measurement Technology Conference*, pp. 1734–1737, IEEE, 2008.

[187] D. Parikh, M. T. Kim, J. Oagaro, S. Mandayam, and R. Polikar, "Ensemble of classifiers approach for ndt data fusion," in *IEEE Ultrasonics Symposium, 2004*, vol. 2, pp. 1062–1065, IEEE, 2004.

[188] R. Heideklang and P. Shokouhi, "Decision-level fusion of spatially scattered multi-modal data for nondestructive inspection of surface defects," *Sensors*, vol. 16, no. 1, p. 105, 2016.

[189] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE transactions on Computers*, vol. 100, no. 2, pp. 179–186, 1972.

[190] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics letters*, vol. 33, no. 2, pp. 156–158, 2008.

[191] R. Chartrand, B. Wohlberg, K. Vixie, and E. Bollt, "A gradient descent solution to the monge-kantorovich problem," *Applied Mathematical Sciences*, vol. 3, no. 22, pp. 1071–

1080, 2009.

[192] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, "Optimal mass transport for registration and warping," *International Journal of computer vision*, vol. 60, pp. 225–240, 2004.

[193] S. Ferradans, N. Papadakis, J. Rabin, G. Peyré, and J.-F. Aujol, "Regularized discrete optimal transport," in *Scale Space and Variational Methods in Computer Vision: 4th International Conference, SSVM 2013, Schloss Seggau, Leibnitz, Austria, June 2-6, 2013. Proceedings 4*, pp. 428–439, Springer, 2013.

[194] S. Kolouri, A. B. Tosun, J. A. Ozolek, and G. K. Rohde, "A continuous linear optimal transport approach for pattern analysis in image datasets," *Pattern recognition*, vol. 51, pp. 453–462, 2016.

[195] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE signal processing letters*, vol. 23, no. 12, pp. 1882–1886, 2016.

[196] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 301–315, 2015.

[197] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pp. 839–846, IEEE, 1998.

[198] B. Shreyamsha Kumar, "Image fusion based on pixel significance using cross bilateral filter," *Signal, image and video processing*, vol. 9, no. 5, pp. 1193–1204, 2015.

[199] A. Kyrgiazoglou and T. Theodoulidis, "Simulation of eddy current non destructive testing using comsol® multiphysics," in *Proceedings of the COMSOL Conference, Rotterdam*, 2017.

[200] T. Takagi, M. Hashimoto, S. Arita, S. Norimatsu, T. Sugiura, and K. Miya, "Experimental verification of 3d eddy current analysis code using t-method," *IEEE Transactions on Magnetics*, vol. 26, no. 2, pp. 474–477, 1990.

[201] L. Barbato, N. Poulakis, A. Tamburrino, T. Theodoulidis, and S. Ventre, "Solution and extension of a new benchmark problem for eddy-current nondestructive testing," *IEEE Transactions on Magnetics*, vol. 51, no. 7, pp. 1–7, 2015.

[202] S. Mukherjee, L. Udpa, and Y. Deng, "Enhanced defect detection in nde using pixel level data fusion," in *2023 International Applied Computational Electromagnetics Society Symposium (ACES)*, pp. 1–2, IEEE, 2023.

[203] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions," *Communications on pure and applied mathematics*, vol. 44, no. 4, pp. 375–417, 1991.

[204] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, and Y. Zhang, "Multi-focus image fusion based on spatial frequency in discrete cosine transform domain," *IEEE signal processing letters*, vol. 22, no. 2, pp. 220–224, 2014.

[205] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 884–892, 2009.

[206] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

APPENDIX A

SUPPLEMENTARY MATERIALS OF CHAPTER 2

Below are the detailed derivation of (2.5) and Table 2.3 calculations.

By theorem 2 of (53) we know that for experiment i and case j, the weighted classification loss is minimized by the estimator:

$$\hat{\Theta}_w(i,j)[l] = 1\left\{\frac{(1-\pi)\phi(y_l|\Delta_0,\sigma_0)}{\pi\phi(y_l|\Delta_1,\sigma_1)} \leq \frac{1}{w_{ij}}\right\},$$

where, $\phi(y|\Delta,\sigma)$ denotes the normal density with mean Δ and standard deviation σ evaluated at y. Taking logarithm on both sides of the inequality above, $\hat{\Theta}_w(i,j)[l]$ reduces to:

$$1\left\{\log((1-\pi)w_{ij}/\pi) \leq \log(\phi(y_l|\Delta_1,\sigma_1)) - \log(\phi(y_l|\Delta_0,\sigma_0))\right\}.$$

The weighted Bayes classifier in the paper follows directly from above by noting that $\log\phi(y|\Delta,\sigma) = -(y-\sigma)^2/2/\sigma^2 - \log\sigma$. The FP and FN error rates of this estimator are given by:

$$\mathrm{FP}_{ij} = P_{Z\sim N(\Delta_0(i,j),\sigma_0^2(i,j))}\left\{\sigma_1^2(Z-\Delta_0)^2 - \sigma_0^2(Z-\Delta_1)^2 \geq 2\sigma_0^2\sigma_1^2\log\left(\frac{w_{ij}\sigma_1(1-\pi)}{\sigma_0\pi}\right)\right\}$$

$$\mathrm{FN}_{ij} = P_{Z\ N(\Delta_1(i,j),\sigma_1^2(i,j))}\{\sigma_1^2(Z-\Delta_0)^2 - \sigma_0^2(Z-\Delta_1)^2 < 2\sigma_0^2\sigma_1^2\log\left(\frac{w_{ij}\sigma_1(1-\pi)}{\sigma_0\pi}\right)\}.$$

These equations are calculated by Monte Carlo simulations and then the weights for each (i,j) pairs is optimized. The results are reported in Table 2.3 which also contains the FP and FN rates for unweighted classification loss when $w_{ij} = 1$.

In the chapter 2, figures 2.4, 2.5, 2.7 describe the different attributes of the data from the three cases of experiment I.

In figures A.1- A.7 the corresponding plots for the other experiments are provided. Figures A.1, A.2, A.3 supplements figure 2.4, figures A.4, A.5, A.6 supplements figure 2.5 and figure A.7 corresponds to figure 2.7 of chapter 2. These figures display reproducible signaling patterns that are described in chapter 2 across the concerned experiments.

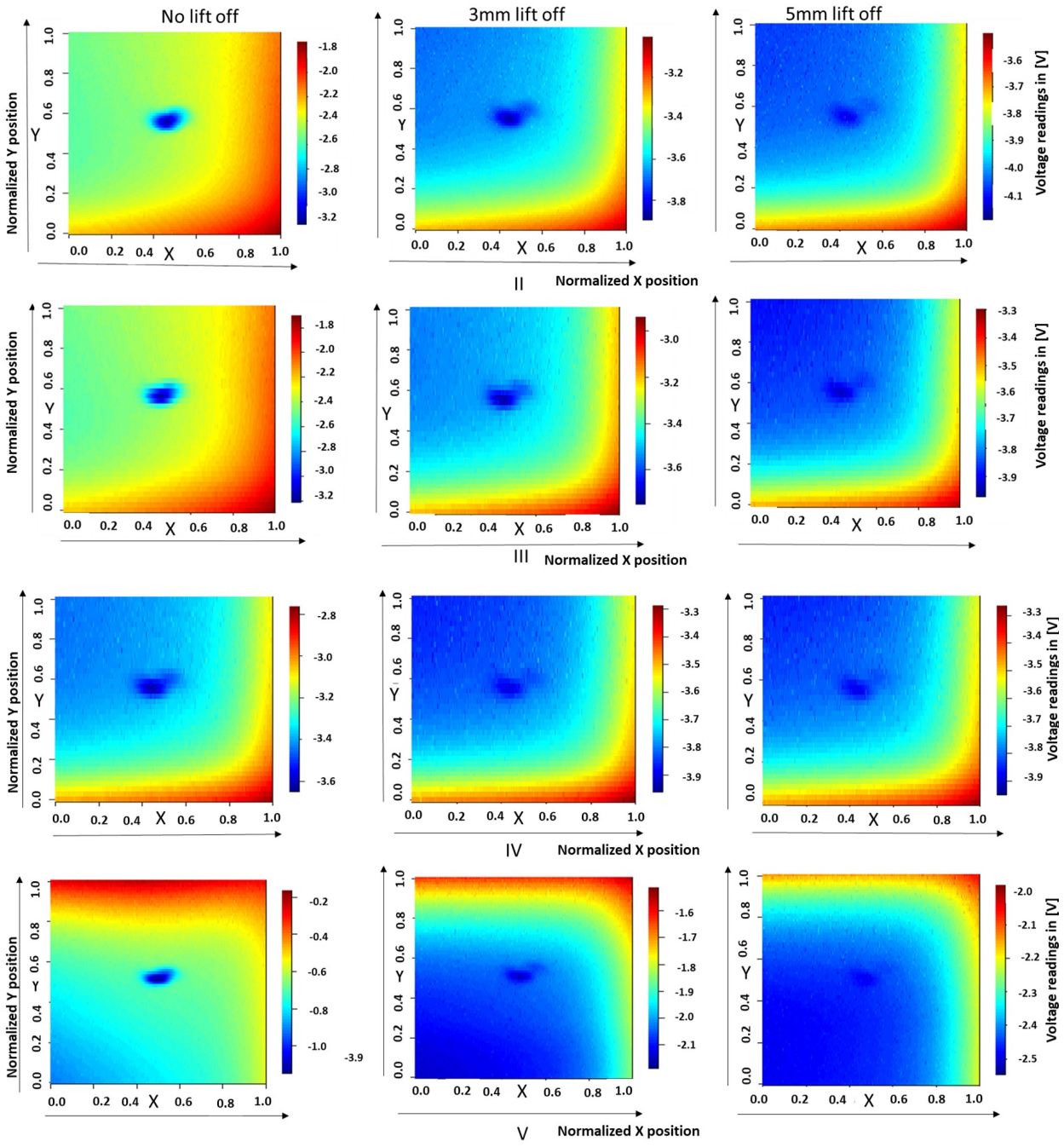

Figure A.1 Plot of voltage readings from three different scenarios (across columns from left to right, we have no lift-off, 3 mm lift-off, 5 mm lift-off) of experiments II to V (across rows) respectively.

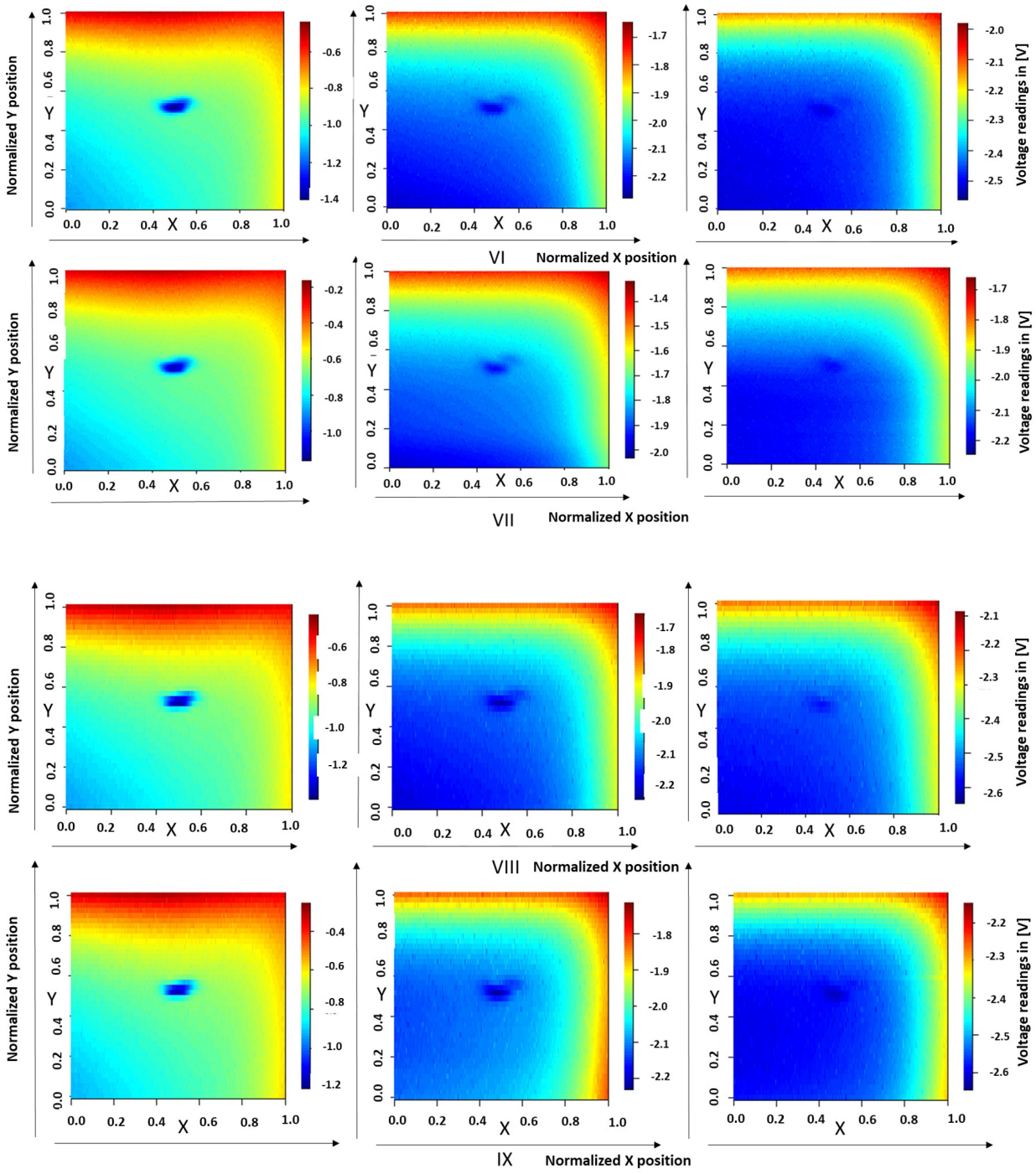

Figure A.2 Plot of voltage readings from three different scenarios (across columns from left to right, we have no lift-off, 3 mm lift-off, 5 mm lift-off) of experiments VI to IX (across rows) respectively.

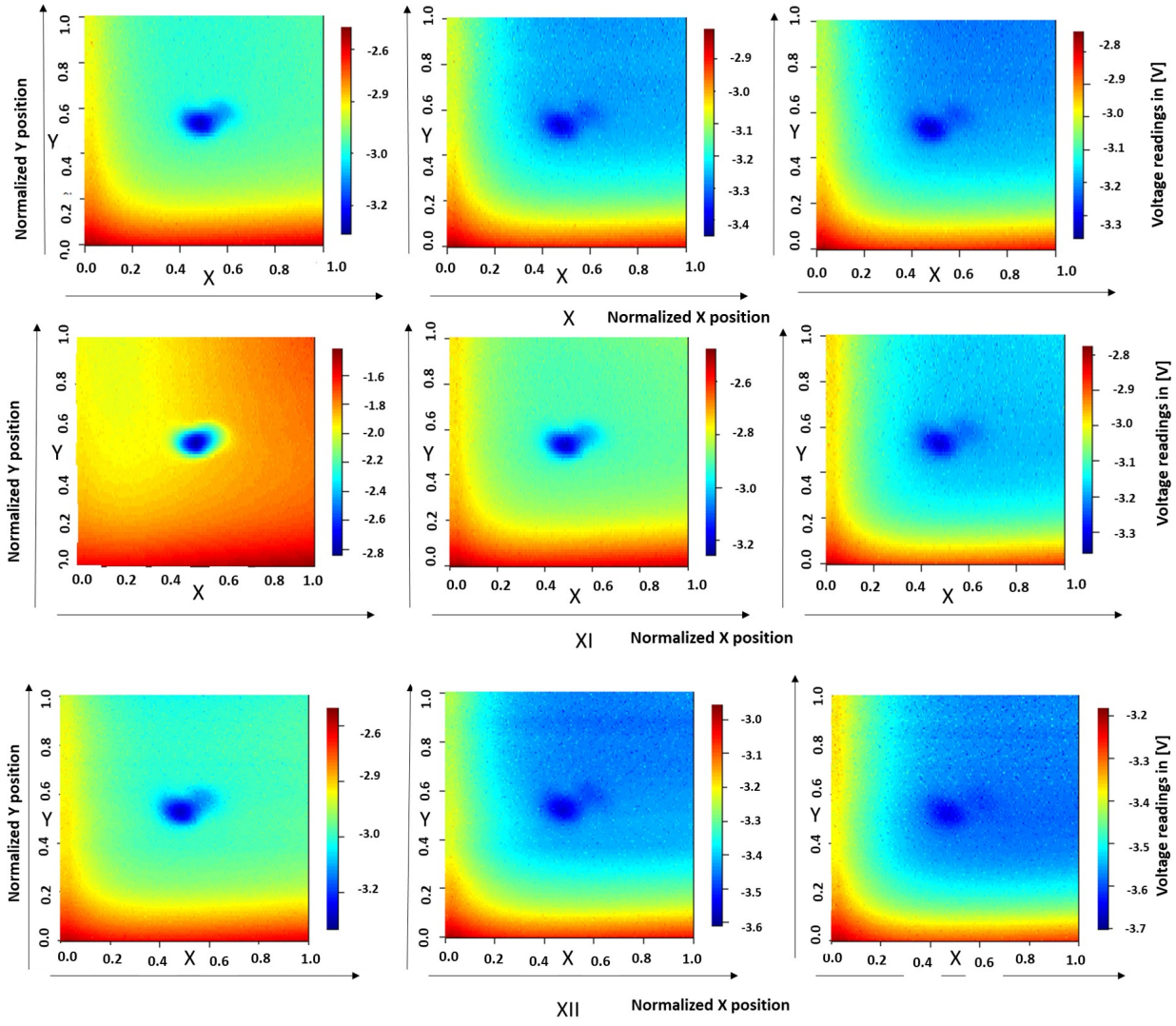

Figure A.3 Plot of voltage readings from three different scenarios (across columns from left to right, we have no lift-off, 3 mm lift-off, 5 mm lift-off) of experiments X to XII (across rows) respectively.

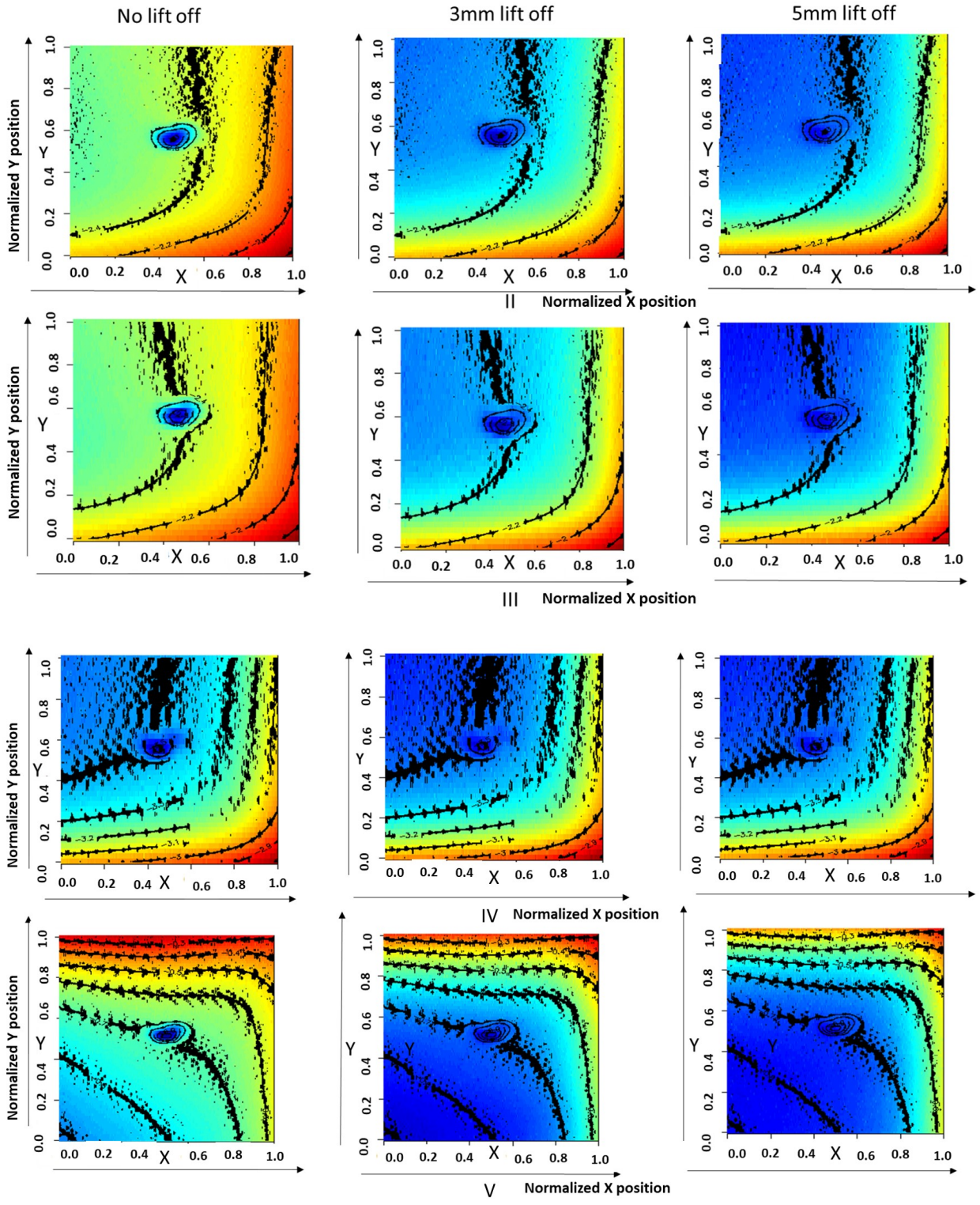

Figure A.4 The contour plot of the left most image is superimposed on each of later images for experiments II to V. The thresholds are selected based on the contour plots of the left most images in each experiment.

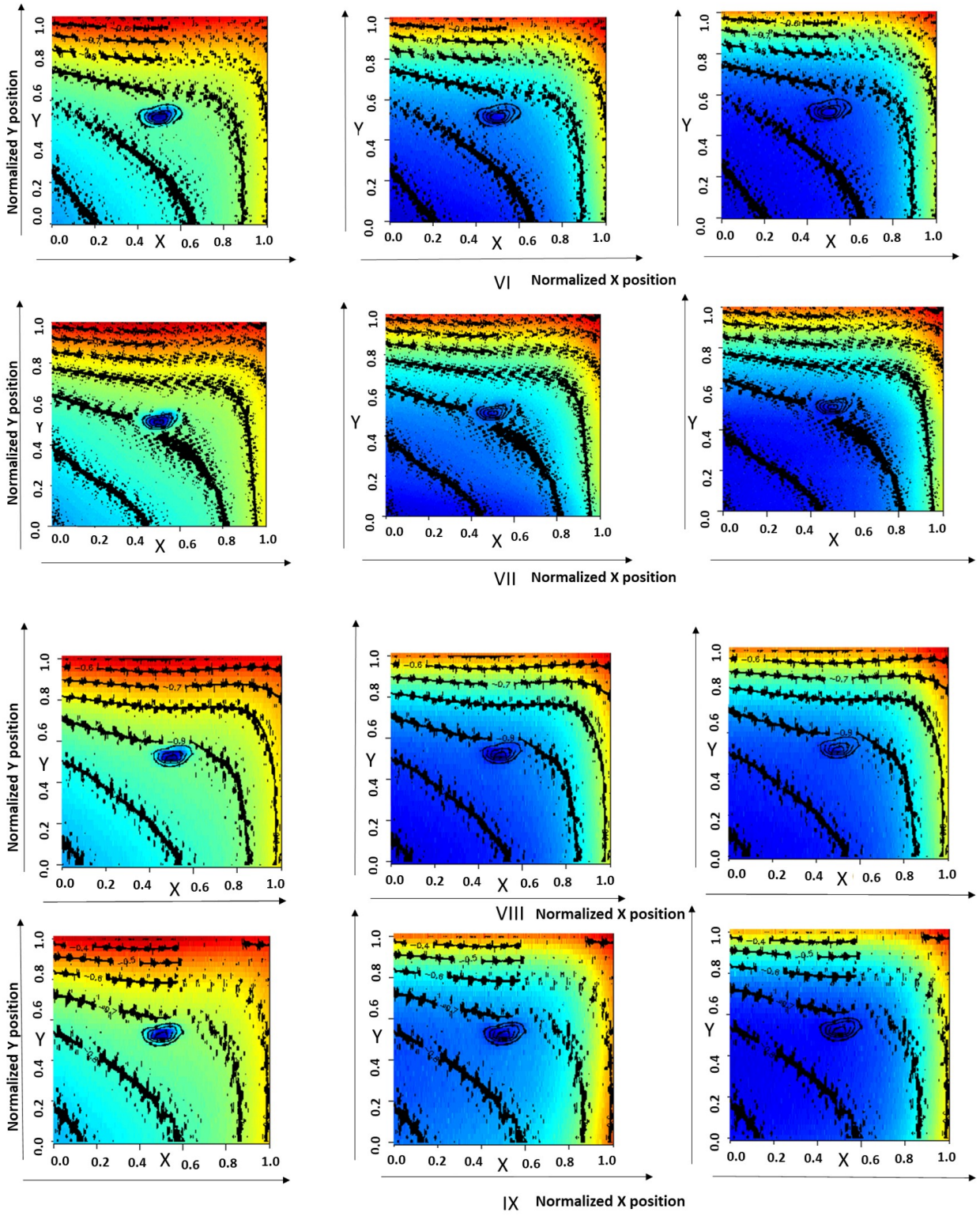

Figure A.5 The contour plot of the left most image is superimposed on each of later images for experiments VI to IX. The thresholds are selected based on the contour plots of the left most images in each experiment.

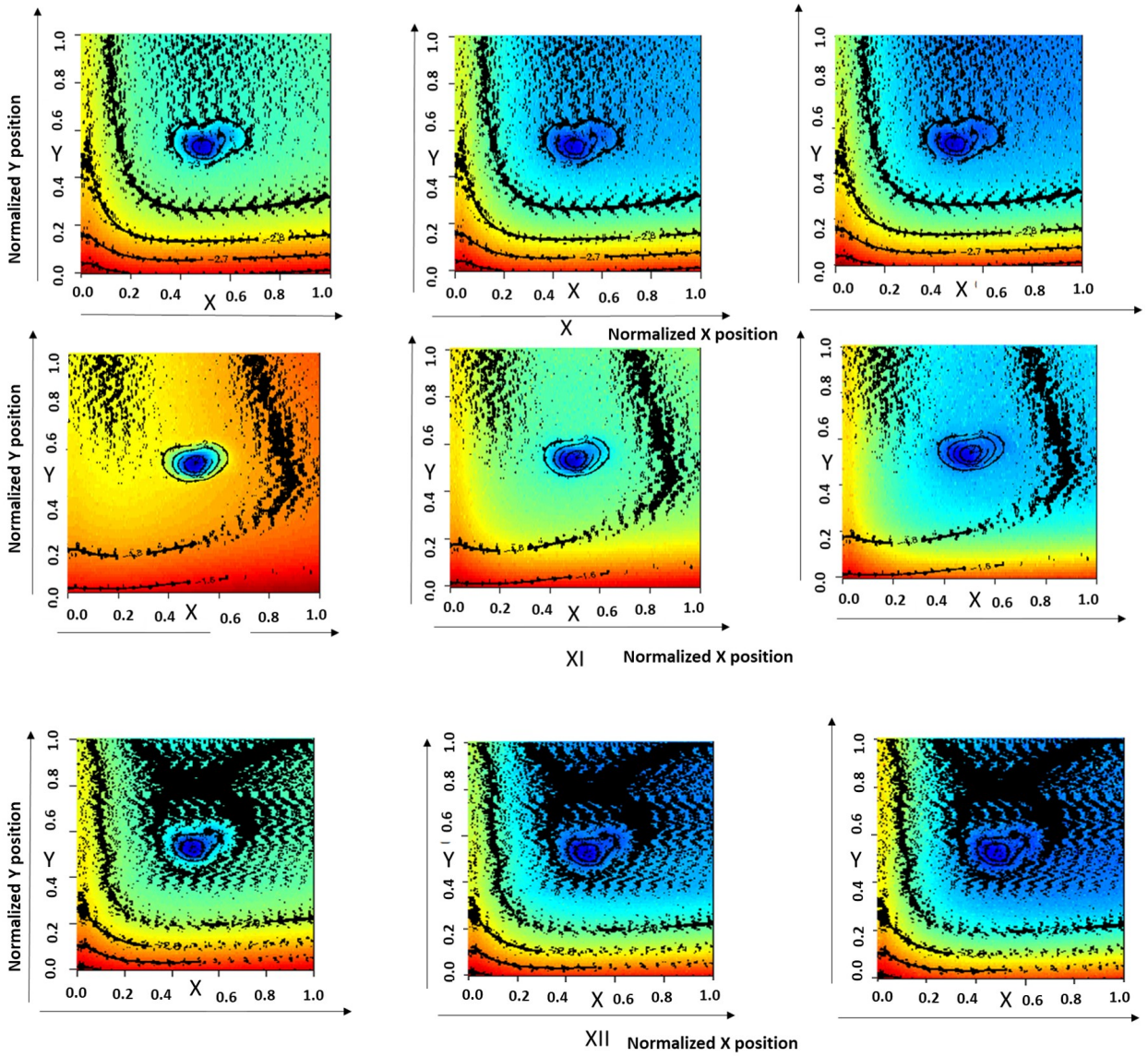

Figure A.6 The contour plot of the left most image is superimposed on each of later images for experiments X to XII. The thresholds are selected based on the contour plots of the left most images in each experiment.

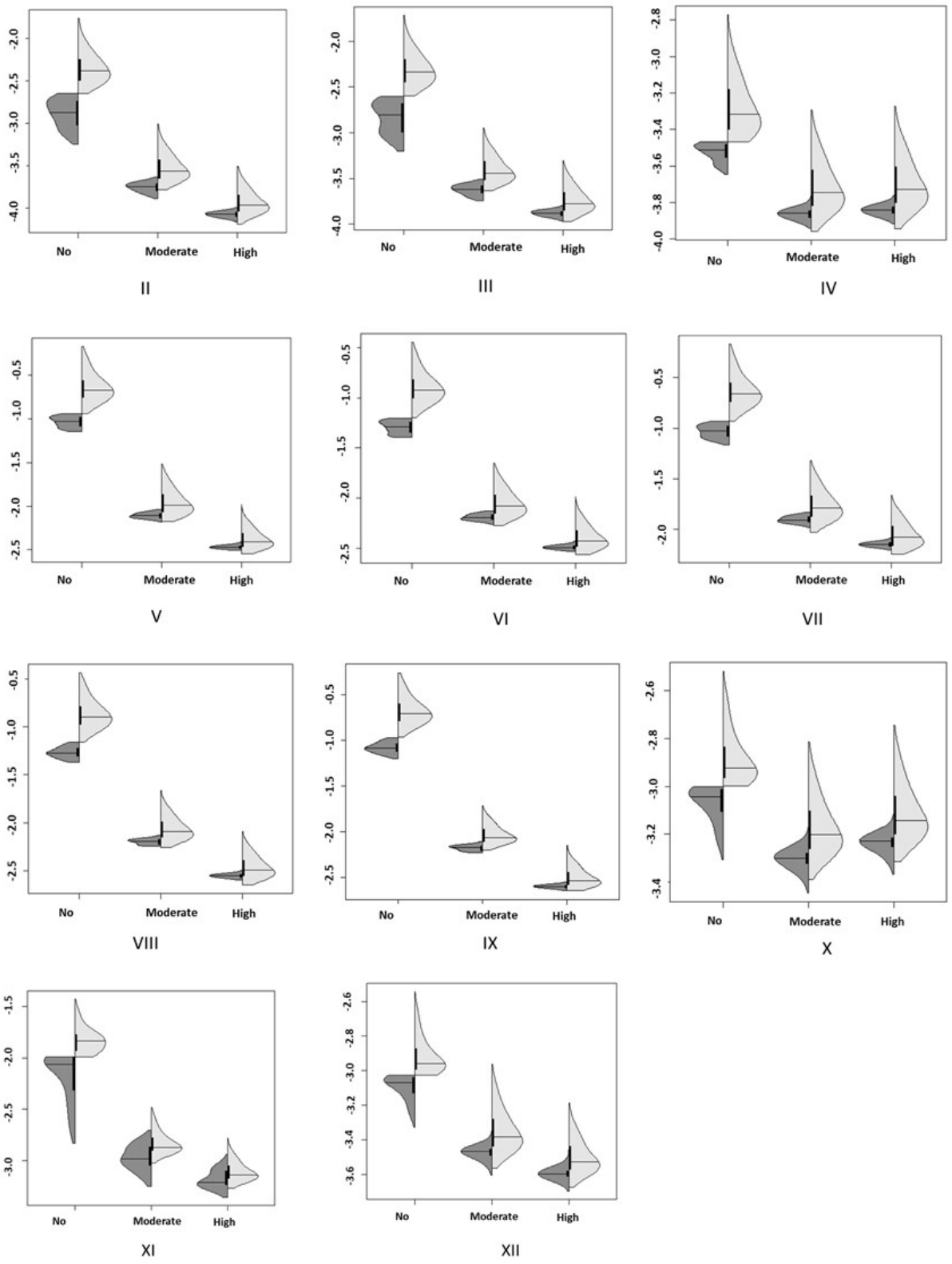

Figure A.7 Distribution of voltage readings in the form of violin plot for the three cases of experiments (II to XII) under no lift-off, 3mm (medium) lift-off and 5mm (high) lift-off respectively.

APPENDIX B

SUPPLEMENTARY MATERIALS OF CHAPTER 6

Detailed Derivation of finding FDR for most efficient pointwise hypothesis test of power β is shown here. For the sake of simplicity, let us assume that the variance of each distribution is equal. Then most powerful (MP) test classifies a point to be defective if:

$$(\sigma_0)^{-2}(x-\Delta_0)^2 - (\sigma_1)^{-2}(x-\Delta_1)^2 \text{is large}$$

$$\iff 2(\Delta_1-\Delta_0)x - (\Delta_1+\Delta_0)(\Delta_1-\Delta_0)is\text{is large}$$

$$\iff x - \frac{(\Delta_1+\Delta_0)}{2} \text{is small as} \Delta_0 > \Delta_1$$

Now, let the rejection region for defect be $x - \frac{\Delta_1+\Delta_0}{2} < c$. Then the Power becomes:

$$P_{defect}\left(x - \frac{\Delta_1+\Delta_0}{2} < c\right) = P\left(\Delta_1 + \sigma Z - \frac{\Delta_1+\Delta_0}{2} < c\right)$$

$$= P\left(Z < \sigma^{(-1)}\left(c + \frac{\Delta_0-\Delta_1}{2}\right)\right) = \Phi\left(\sigma^{-1}(c + \frac{\Delta_0-\Delta_1}{2})\right) = \Phi\left(t+R\right)$$

where, Z is the standard normal variable. We substitute $c = t\sigma$ in the above equation and R is the SNR.

The false discovery rate (FDR) becomes:

$$\text{FDR} = P_{non\ defect}\left(x - \frac{\Delta_1+\Delta_0}{2} < c\right) = \Phi\left(\sigma^{-1}(c - \frac{\Delta_0-\Delta_1}{2})\right) = \Phi\left(t-R\right)$$

Now when Power = β. Then $\beta = \Phi\left(t+R\right) \Rightarrow t = \Phi^{-1}\beta - R$

Substituting the value of t to FDR,

$$FDR = \Phi\left(t-R\right) = \Phi\left(\Phi^{-1}\beta - 2R\right)$$

APPENDIX C

SUPPLEMENTARY MATERIALS OF CHAPTER 8

Here some extra registration aided results of the defect dimensions of Table 8.3 are presented. The fused results from the boundary conditions clearly shows the efficacy of registration aided fusion which is proposed in chapter8 of this thesis.

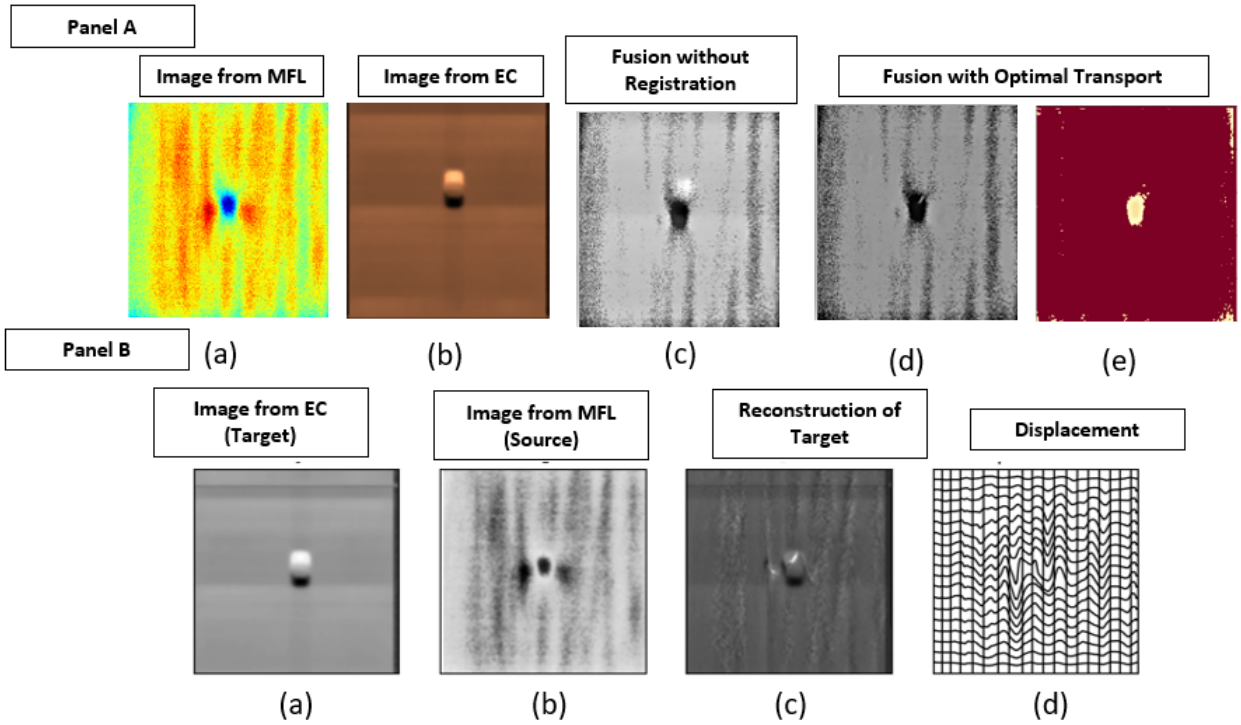

Figure C.1 Panel A: (a) Source MFL image, (b) Source EC image, (c) Result on applying fusion without registration showing misalignment, (d) Applying fusion based on OT, (e) Post processed results on (d), Panel B: (a) EC image which forms the target distribution, (b) MFL image as source distribution in OT, (e) Reconstruction of target after OT, (d) Displacement showing transfer of mass from source to target. Analysis is done based on 'Defect 8' of Table 8.3.

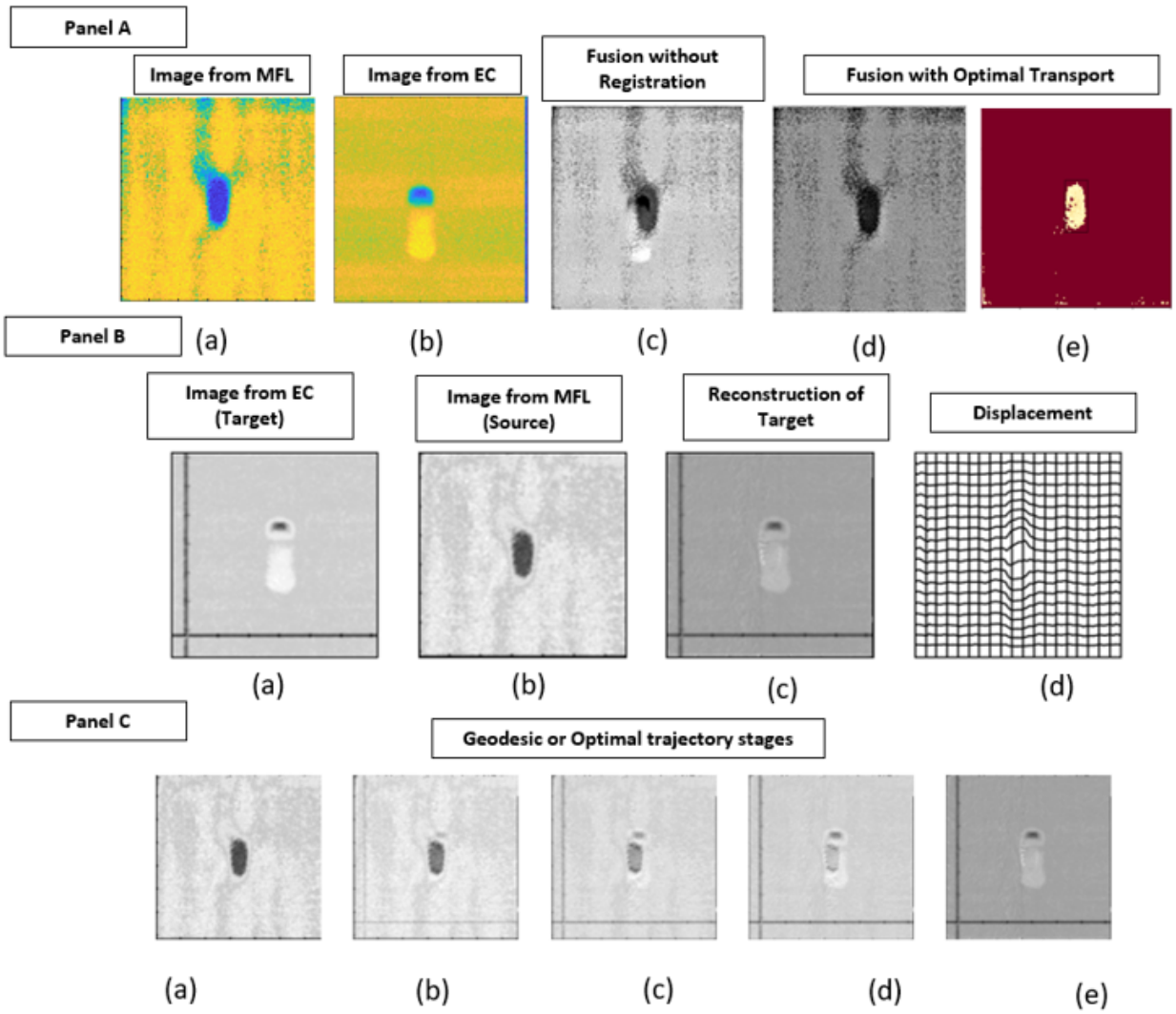

Figure C.2 Panel A: (a) Source MFL image, (b) Source EC image, (c) Result on applying fusion without registration showing misalignment, (d) Applying fusion based on OT, (e) Post processed results on (d), Panel B: (a) EC image which forms the target distribution, (b) MFL image as source distribution in OT, (e) Reconstruction of target after OT, (d) Displacement showing transfer of mass from source to target. Analysis is done based on 'Defect 11' of Table 8.3.

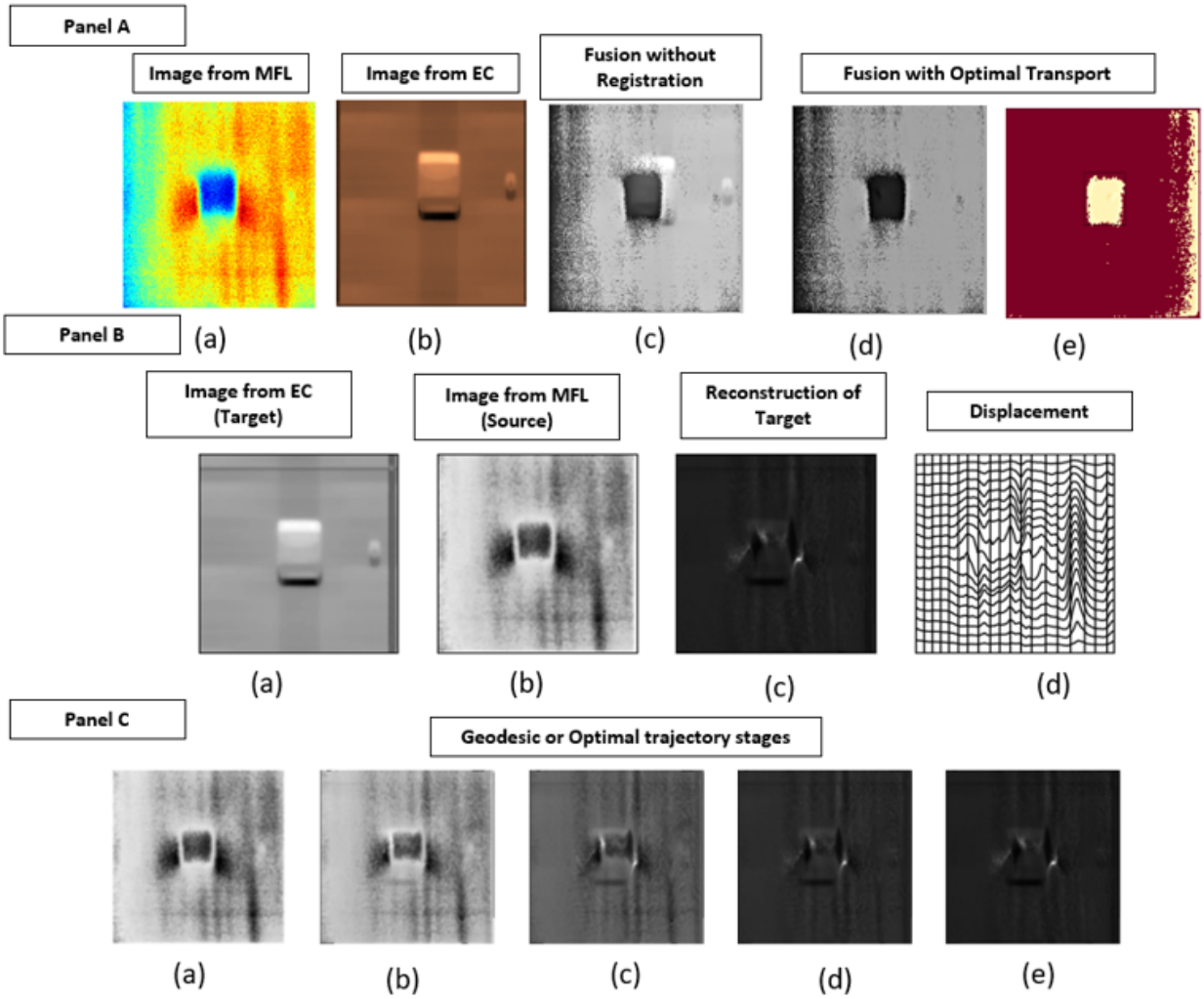

Figure C.3 Panel A: (a) Source MFL image, (b) Source EC image, (c) Result on applying fusion without registration showing misalignment, (d) Fusion based on OT, (e) Post processed results on (d), Panel B: (a) EC image which forms the target distribution, (b) MFL image as source distribution in OT, (e) Reconstruction of target after OT, (d) Displacement showing transfer of mass from source to target. Analysis is done based on 'Defect 21' of Table 8.3.

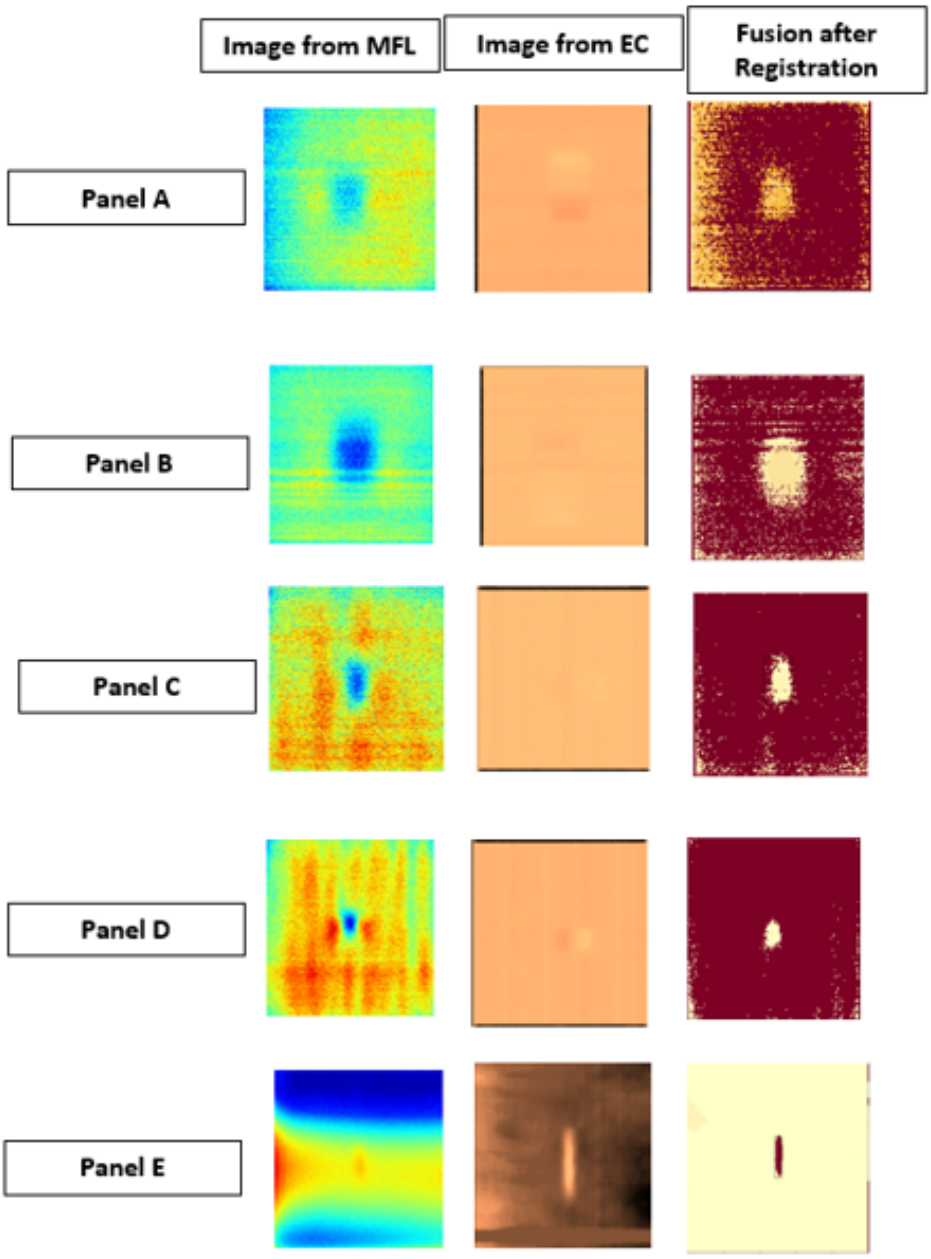

Figure C.4 Shows the efficacy of registration based fusion in boundary conditions due to lift offs, noisy conditions and limiting dimensions of defects. Panel A represents Defect 21, Panel B 17, Panel C as Defect 13, Panel D as Defect 8 and Panel E as Defect 3 respectively.