# Intelligent Machine Learning-based Leakage Detection and Localization in Vacuum-assisted Composite Manufacturing

by

Yussuf Reza Esmaeili Shahmiri

M.Sc., Sharif University of Technology, 2018
B.Sc., Sharif University of Technology, 2016

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in
The Department of Mechanical Engineering



We acknowledge and respect the Lək$^{w}$əŋən (Songhees and X$^{w}$sepsəm/Esquimalt) Peoples on whose territory the university stands, and the Lək$^{w}$əŋən and W̱SÁNEĆ Peoples whose historical relationships with the land continue to this day.

**Supervisory Committee**

---

Prof. Homayoun Najjaran, Supervisor
(Department of Mechanical Engineering)

---

Prof. Colin Bradley, Departmental Member
(Department of Mechanical Engineering)

---

Dr. Brandon Haworth, Non-departmental Member
(department of Electrical Engineering)

## ABSTRACT

Vacuum-assisted composite manufacturing methods, such as vacuum bag prepreg layup and vacuum-assisted resin transfer molding (VARTM), utilize atmospheric pressure as a uniform external force to consolidate and saturate fabric components. However, vacuum bag leakages can result in defects such as air bubbles, resin traps, voids, non-uniform surface finishes, and ultimately, inferior mechanical properties. Detecting and repairing these leakages before the autoclave curing stage is therefore essential.

The leakage localization method used in this study relies on volumetric flow rate measurements of air evacuation lines. Multiple air evacuation channels, known as vacuum ports, are strategically placed at different locations in the production layup. Each port is equipped with sensors capable of independently measuring the volumetric flow rates of air during the process. In the presence of a leakage, the measured flow rate values will not stabilize at zero because air continuously enters the vacuum bag through the leak. The flow rate values correlate with the location of the leak, the overall layup configuration, and the positions of the vacuum ports.

We introduced an intelligent machine learning-based framework for leakage detection and localization, designed to learn the complex relationships between flow rate values and leak locations. To generate sufficient training data, an electric circuit analogy was developed to simulate the vacuum process. This approach provides a fast and reliable alternative to complex analytical simulations and extensive physical experiments. The proposed method has been validated and compared across various experimental configurations, demonstrating its effectiveness.

Using the available and synthesized data, we employed various machine learning models, including regression models, a Grid neural network, a physics-informed Grid neural network, leakage classification models, and physical parameter training algorithms for leakage prediction. Our methods not only predict leakage locations with acceptable accuracy but also generalize well across different configurations. Additionally, we addressed challenges associated with complex, non-uniform layups featuring regions of varying permeability. For the first time, our framework also tackled scenarios involving multiple simultaneous leakages, successfully localizing all leaks on the layup.

Our results demonstrate significant advancements over state-of-the-art methods. These improvements go beyond higher prediction accuracy, focusing on enhanced gen-

eralizability across various layups, reduced data requirements for training, and the ability to tackle complex scenarios, such as non-uniform permeability and multiple leakages, which were previously unaddressed. Notably, the novel PI-GNN framework outperforms regression models in both generalizability and data efficiency. By integrating physical knowledge with data science, the PI-GNN framework establishes a robust foundation for addressing layups of varying sizes and geometries. Furthermore, our proposed physical parameter training algorithm effectively learns the permeability of different regions within the layup, enabling the development of a more accurate and robust simulation tool for model training.

Optimizing the placement of vacuum ports to improve leakage location prediction is another challenge addressed in our work. Each layup offers numerous possible configurations for positioning vacuum ports to enhance leakage localization. We tackled this optimization problem by maximizing flow rate variance among the vacuum ports. Given the problem's large state space, a hierarchical optimization approach was employed to identify the optimal configuration. Experimental validation confirmed that optimizing the port configuration significantly reduces leakage prediction errors.

**Keywords:** VARTM, Prepreg, Graph neural network, Leak detection, Machine learning, Leak localization, port placement, optimization.

# PREFACE

This research was conducted and written by Yussuf Reza Esmaeili under the supervision of Dr. Homayoun Najjaran at the Advanced Control and Intelligent Systems (ACIS) lab at the University of Victoria. The author led the experimental work, numerical simulations, and data analysis. The experiments have been conducted in Composite Research Network (CRN) at the University of British Columbia. This Ph.D. dissertation, along with related activities, has resulted in several publications and collaborations.

The following contributors also supported this work:

- Dr. Anoush Poursartip, Dr. Goran Fernlund and Malcolm Lane from Convergent Manufacturing Technologies: Providing the COHO leak detection and flow rate sensor system and experimental materials and space.
- Brett Cosco: Assisting in the experiment setup, design and data recording.

The following publications have resulted from the work of this PhD:

1. Portions of Chapter 3 including the experimental setup and data acquisition, simulation methodology and the first section of Chapter 4, simulation results were published in the "The International Journal of Advanced Manufacturing Technology".

   ***Esmaeili, Yussuf Reza**, Brett Cosco, and Homayoun Najjaran. "Vacuum bag leak detection for resin infusion: an electric current–based analogy." The International Journal of Advanced Manufacturing Technology 124, no. 5 (2023): 1775-1786.*

2. Section 3.10 of Chapter 3, Vacuum port placement optimization and the results in Section 4.5 were published in the "Production Engineering" journal.

   ***Esmaeili, Yussuf Reza**, and Homayoun Najjaran. "Optimal vacuum port placement in vacuum-assisted composite manufacturing for leakage detection and localization: a hierarchical optimization approach." Production Engineering (2024): 1-11.*

3. Sections 3.7 and 3.8 and their corresponding results presented in Chapter 4 on single leakage localization training and single and multiple leakage classification

methods have been submitted to *The International Journal of Advanced Manufacturing Technology* and are currently awaiting revision based on the reviewers' comments.

***Esmaeili, Yussuf Reza**, Brett Cosco, Yifan Pan, and Homayoun Najjaran. "Intelligent Vacuum Bagging Leakage Location Prediction" The International Journal of Advanced Manufacturing Technology (2025), https://doi.org/10.1007/s00170-024-14987-6*

# Contents

# List of Tables

# List of Figures

# ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Homayoun Najjaran, for his invaluable guidance, support, and mentorship throughout my research journey at the Advanced Control and Intelligent Systems (ACIS) lab at the University of Victoria. His insights and encouragement have significantly contributed to my academic and personal growth and have helped me shape this work professionally and with joy. Additionally, his cooperation with me to be able to pursue a part of my research in another city.

I am sincerely grateful to my Dr. Anoush Poursartip, from the Composites Research Network (CRN) and the University of British Columbia (UBC) Vancouver campus, for including and accepting me in his team. His valuable suggestions and assistance over the past 3 years, which clearly enhanced the quality of my research phases.

I would like to thank Dr. Goran Fernlund and Malcolm Lane from the Convergent Manufacturing Technology, for their assistance and contributions to my research and providing me with the COHO system and the Software. Their expertise has enriched this thesis and broadened my understanding of the field notably.

A special thanks goes to my colleagues and friends at CRN Vancouver and ACIS lab in Victoria, especially my dear friend Amir Mehdi Soufi Enayati, for their camaraderie, advice, and support, which made this journey a truly memorable experience.

I would like to express my heartfelt appreciation to my family, who have been a constant source of love and encouragement since the day I grew up, even when we were separated by distance. Their unwavering belief in me has been a pillar of strength.

Finally, my deepest gratitude goes to my beloved wife, Mahshid Mahbod, for her endless support and encouragement throughout my Ph.D. She has been a family member, a friend and a colleague in my PhD journey and has always believed in my potential, inspired me to take big steps, and reminded me to embrace every opportunity. This achievement would not have been possible without her love and unwavering faith in me.

# DEDICATION

*I would like to dedicate my thesis to my grand father, a man who I always have respected, and I am delighted to see one last time before he passed away...*

*And to my parents, for their encouragement and support, and their promise to be strong while I am away studying my PhD...*

*And to my beloved wife, who made this journey way easier and possible for me.*

*Dedicated from Yussuf*

# Chapter 1

# Introduction

Composite manufacturing methods such as Vacuum Assisted Prepreg Compression Molding or Vacuum Assisted Resin Transfer Molding (VARTM) are being widely used for producing high-quality large-scale components [1][2]. Composite manufacturing has applications in aerospace [3] [4], automotive [5] [6], marine [7] [8], wind energy [9] [10], construction [11], defense [12], sport and recreation industries [13] [14].

The growing use of fibrous composites can be attributed to their high specific strength and stiffness compared to conventional materials, as well as their ability to be shaped and tailored to achieve aerodynamically efficient structural configurations [15]. Recent advancements in composite materials for modern aircraft construction suggest that fiber-reinforced polymers, particularly carbon fiber-reinforced plastics (CFRPs), could contribute more than 50% of an aircraft's structural mass in the future. However, affordability remains critical for the competitiveness of aerospace manufacturing, both in civil and military applications [16]. Consequently, significant efforts should focus on the analysis and computational simulation of manufacturing and assembly processes, as well as on the performance simulation of the resulting structures, since these aspects are closely interconnected [17].

Both of the prepreg or VARTM approaches are often executed by arranging layers of a release film, peel plies, reinforcement fabrics and breather fabric on a mold, prior to being sealed in a vacuum bag. The principal role of the reinforcement is to provide strength, stiffness and other mechanical properties to the composite [18]. In either methods, the pressure difference between the ambient air pressure ($P_{atm}$) and the vacuum bag pressure ($P_{vac}$) helps with the consolidation of the component[19]. Figure 1.1 shows the assembly of the prepreg setup. After preparing the assembly, a vacuum pump evacuates the air from the bag. At this stage, it is essential to ensure

that there are no leakages in the vacuum bag. Afterwards, in the VARTM approach, resin will be transferred into the system, building the matrix around the reinforcement fiber [20]. Conversely, the prepreg method does not require resin injection as the resin already impregnated in the prepreg material [21]. Then the system will be cured in an autoclave for some time so that the resin properly adheres the fabric [22].

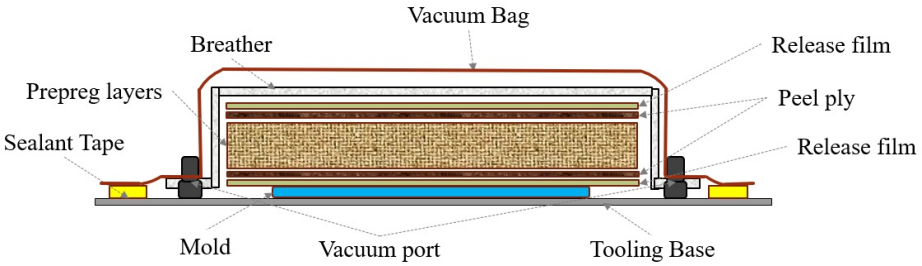


Figure 1.1: Schematic of the vacuum prepreg processing setup

The prepreg method is typically used in aerospace applications due to its capability of producing high quality components[23]. The environment's pressure and temperature can be controlled in the autoclave, and thus a large pressure gradient can be produced. The larger the pressure gradient between the interior and exterior of the vacuum bag, the larger the consolidation or clamping force provided by the vacuum. Higher force generally produces better quality composites which are more dense and free from voids[24].

Vacuum bagging is one of the most important steps before the curing process. The existence of a leakage in the vacuum bag will prove detrimental to the mechanical properties and surface smoothness. A leaking vacuum bag may require complex and expensive repairs or result in severe impairment of the product quality[25]. In the autoclave stage of the process, the pressure inside the autoclave is increased to seven atmospheres in order to maximize the consolidation force [26]. Under these extreme conditions, even the presence of a minuscule leakage jeopardizes the quality and characteristics of the final product. Fernlund et al. concluded that leakages in vacuum bags during the curing process can be a source of porosity and, consequently, poor mechanical properties in components [27]. A study by Haschenburger et al. [28] also demonstrated that leaks in vacuum bags and release films can cause porosity and result in inconsistent thickness in the final part. Therefore, it is imperative to locate and mend all leakages prior to the curing process [29]. The focus of this research is on leakage detection and localization in the vacuum bagging step of these composite manufacturing methods. In the next section, industrial leakage detection

and localization methods will be discussed.

## 1.1 Research Objective and Motivation

A perfectly sealed vacuum bag is essential for a quality product and to achieve that, potential leakages must be discovered and mended. An automated leakage localization method is mandatory because in many applications, for instance aerospace industry, the size of the product is so enormous that the time, effort and cost make manual leakage localization infeasible. Manufacturing processes must be able to produce economically large components with as few joints as possible. This presents challenges regarding the complexity of tooling and control of the process [30].

In the previous sections, state-of-the-art leakage detection and localization methods were discussed. The findings reveal no single method that consistently outperforms others across all criteria. Existing methods are often unreliable in autoclave environments or when multiple leakages occur, or they are costly, time-consuming, and lack the desired level of automation. Measuring flow rates on evacuation lines and applying machine learning-based methods shows greater potential due to their versatility and automation capabilities. However, these approaches are limited to handling symmetrical and simple geometries, making them inadequate for the complexities of industrial parts. Furthermore, data acquisition in these methods is entirely manual and requires the physical system. Few studies have explored this technique in real-world scenarios, but failing to consider and involve non-uniform permeability and multiple leakages. Moreover, while there are studies on optimizing resin inlet configurations, none have addressed the optimization of vacuum port locations on the layup.

The objective of this research is to provide a versatile framework for vacuum bag leakage localization in composite manufacturing using machine learning models. The proposed framework should effectively overcome the limitations of existing machine learning-based methods in four key areas:

1. Data acquisition/synthesis: Enhancing the ability to gather or generate high-quality data for training the model.

2. Generalizability and adaptability to new geometries: Ensuring the framework can effectively handle complex geometries with non-uniform permeabilities encountered in industrial parts.

3. Overall leakage prediction: Improving the prediction accuracy and reliability compared to other automated methods.

4. Single/multiple leakage location prediction: Enabling the framework can handle problems where there are multiple leakages in the system.

5. Vacuum port placement: Optimizing the configuration of vacuum ports on the layup for an improved leakage localization.

The novelties and further details of the proposed framework will be discussed in the next section.

## 1.2 Novelties

The following list highlights the novel aspects and potential impacts of the proposed method:

- In this study, we present a novel approach for simulating flowrate values in the vacuum process and demonstrate its validity through extensive validation. Our method draws an analogy between the flowrate in the vacuum bag and the current in an electrical circuit. By simulating the electrical circuit, we generate results that can be compared with the experimental outputs of the vacuum bag.

- Unlike existing frameworks that rely on manual data acquisition from the physical system, we present a novel approach utilizing a verified simulation tool for data synthesis. Our proposed tool draws upon an analogy between the physical system and an electrical circuit, enabling the rapid generation of training data. By doing so, we can bypass one of the most crucial and time-consuming steps in the process.

- Inspired from graph neural networks, we propose a new architecture called grid neural network that increases flexibility and generalizability of the trained agent. This new architecture is not limited to the leakage localization problem and could be used for any similar physical systems. In addition to its training phase, our proposed design allows for the integration of physical knowledge into the system, resulting in a remarkable reduction in the required amount of training data. Furthermore, the grid neural network facilitates the incorporation of

expert knowledge as priors, enabling Bayesian updates to refine the probabilities of leakage locations.

- For the first time in the literature, we have addressed multiple leakage scenarios. Using data generated from a validated simulation, we developed a classification algorithm capable of determining the number of leakages on the layup. Based on this classification, the system then selects the appropriate model for both single and double leakage localization.

- We have developed a framework that enables the model to learn its physical parameters using limited data. This approach is particularly effective for handling complex layups with constrained data availability. We applied physical parameter training to estimate the permeability of different regions within the layup. Enhancing the uncertain and unknown parameters in the simulation improves its accuracy and reduces error, making the synthesized data used for training more reliable.

- We have developed a framework for optimal vacuum port placement to enhance leak detection. To manage the large state space in the optimization problem, we employed a hierarchical optimization approach. The results demonstrate that using an optimal vacuum port placement can significantly improve leakage prediction accuracy.

## 1.3 Agenda

In Chapter 2, a literature review is presented, covering background information, composite manufacturing methods, and an overview of state-of-the-art leakage detection and localization techniques. Chapter 3 delves into the details of various methodologies employed and provides a comprehensive description of the proposed framework across different scenarios and case studies. Chapter 4 presents the detailed results in the same order as outlined in the methodology chapter. Chapter 5 focuses on drawing conclusions based on the research findings, discussing study limitations, and outlining future steps to achieve the study's objectives. The thesis concludes with the Bibliography section, which includes all references.

# Chapter 2

# Literature Review

## 2.1 Composite materials: an overview

Composite materials are engineered materials made from two or more constituent materials with significantly different physical or chemical properties [31]. When combined, these constituents form a material with properties distinct from the individual components. The key characteristics of composite materials include improved strength, stiffness, durability, or other properties compared to the constituent materials used alone [32]. The base material and the filler material are the components of a composite material.

The base material which is also called a matrix or a binder, holds the filler material together. It serves to transfer stress to the reinforcement material and protect it from environmental damage [33]. Common matrix materials include polymers (like epoxy, polyester, or thermoplastics), metals, and ceramics.

The filler material or the reinforcement is the dispersed phase within the composite, providing strength, stiffness, and other mechanical properties [34]. The reinforcement could take the form of sheets, particles, flakes or more importantly, fibers. Common reinforcement materials include glass fibers, carbon fibers, aramid fibers, ceramics, and natural fibers.

Composites are classified into three main categories based on their structure [35]:

1. Fiber-reinforced composites

   Fiber-reinforced composites (FRCs) are materials made by embedding fibers into a matrix material (usually a resin or polymer). The fibers provide strength and stiffness, while the matrix holds the fibers together and transfers the load

between them. The fiber-reinforced composites are the most common and widely used type of composite materials. They dominate the composite materials market due to their excellent balance of mechanical properties, weight, and versatility, which make them suitable for a wide range of applications across various industries. FRCs are the main focus of this study and will be discussed more in Section 2.2.

2. Particle reinforced composites

   Particle-reinforced composites (PRCs) are composite materials where small particles are dispersed within a matrix material to enhance its mechanical properties. These particles, typically ranging in size from nanometers to millimeters, serve as the reinforcement phase, improving the performance characteristics of the matrix material. Although PRCs generally offer less significant strength and mechanical properties compared to fiber-reinforced composites, they are widely used in applications where enhanced hardness and stiffness are required [36]. A common example is in roadways and concrete structures, where cement acts as the binder, and aggregates like coarse rock or gravel serve as filler materials, providing increased hardness and structural integrity [37].

3. Sheet molded composites

   Sheet-molded composites (SMCs) are fabricated by bonding homogeneous layers of materials using a compression molding process to form non-homogeneous composite laminates. Sheet molded composites offer high strength-to-weight ratio, good corrosion resistance and good surface finish, but they have limited Recyclability and cannot match the strength of fiber-reinforced materials [38]. The main application of SMCs are in automotive industry [39].

## 2.2 Types of fiber-reinforced composites

In this section, various types of fiber-reinforced composites will be introduced. These composites can be classified in different ways, such as by their structure or by the source and origin of the fibers. Different materials impart unique properties to the fibers, influencing the overall performance of the composite.

### 2.2.1 Fiber classification by structure

Fiber-reinforced composites can be categorized based on their structure into continuous fibers, woven fibers, chopped fibers, and hybrid composites [40]. In continuous fiber composites, fibers can be arranged either unidirectionally or bidirectionally within the matrix, enabling efficient load transfer from the matrix to the fibers. Discontinuous fibers, however, must be of sufficient length to allow for effective load transfer and to prevent crack propagation, which is crucial in avoiding material failure, particularly in brittle matrices [41]. The properties and structural behavior of composite materials are largely determined by the arrangement and orientation of the fibers [42]. Continuous fibers usually have higher strength in one direction, while the chopped and discontinuous fibers offer more isotropic properties.

### 2.2.2 Fiber classification by origin

Fiber-reinforced composites can be categorized based on the type of fiber into synthetic and natural fibers. Synthetic fibers include materials such as glass, carbon, and aramid fibers, which are commonly used due to their high strength, stiffness, and durability. Natural fibers, such as flax, hemp, and jute, are increasingly used for their lower cost, availability and environmental benefits, including biodegradability and renewable sourcing [43], although they generally offer lower mechanical performance compared to synthetic fibers [44].

### 2.2.3 Different types of synthetic fibers

Different types of reinforcement fabric materials are used in the aerospace industry [45]. One of the most common composites is carbon fiber-reinforced polymer, an exceptionally strong [46] and lightweight material containing carbon fibers [47]. Fiberglass is another widely used fiber-reinforced polymer, consisting of a plastic matrix reinforced with fine glass fibers. It is a lightweight, strong, and robust material. Although its strength and stiffness are somewhat lower than that of carbon fiber [48], fiberglass is generally much less brittle, and the raw materials are significantly more affordable [49]. Aramid fiber, a class of heat-resistant and strong synthetic fibers, offers excellent resistance to chemicals and heat, a high tensile modulus, and low elongation at break. However, it has limited fatigue strength and poor compression strength [50]. Aramid fibers are used in aerospace and military applications, including

ballistic-rated body armor fabric and composites, bicycle tires, and as a substitute for asbestos [51]. Figure 2.1 illustrates the three most common types of fiber-reinforced fabrics used in the industry. Ongoing development of new and more advanced materials aims to improve the processing and production of existing materials, enhancing physical properties or enabling their application in new areas for future use [52].

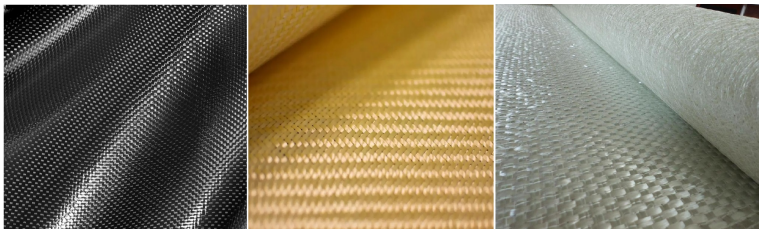

Figure 2.1: Different types of fiber reinforced fabrics; from left to right: carbon-fiber, aramid and glass-fiber reinforced polymers

## 2.3 Fiber-reinforced composites manufacturing techniques

Manufacturing of fiber reinforced polymers composite involves manufacturing of fiber preforms and then reinforcing these fibers with the matrix material by various techniques [53]. In this section some of the most common and related techniques will be introduced.

### 2.3.1 Hand layup

The hand layup method, also known as the wet layup method, is one of the most traditional techniques used in the industry. It is a straightforward process where each ply is manually placed and stacked layer by layer until the desired thickness is achieved [54]. A roller is typically used to press the resin into the fabric, ensuring better interaction between the successive layers of reinforcement and the matrix material. Figure 2.2 shows the schematic of hand layup manufacturing method [55]. Usually, curing is done in the room temperature [56]. While this method is reliable, it is labor-intensive and time-consuming compared to advanced manufacturing techniques. The quality of the final product is highly dependent on the skill of the worker [57]. Additionally, this method has limitations in terms of the complexity of the parts it

can produce, particularly in the aerospace industry. addition, it is difficult to achieve consistent volumetric fiber fractions and low void contents using this method.

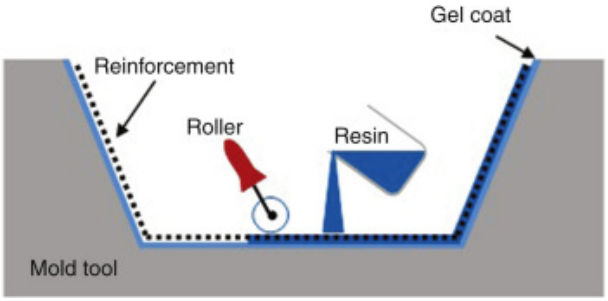


Figure 2.2: Schematic of hand layup manufacturing technique [43]

### 2.3.2 Spray layup

The spray layup process method is similar to the hand layup method, but it uses a hand gun to spray resin and chopped fibers on a mold. It uses a low cost mold, and it is a preferred method to use for big product finishing [58]. The stream of the hand gun is aimed at the mold and left until it is fully cured at room temperature. It also uses a roller and brushes to clear the air and wet the fiber. Figure 2.3 represents the hand layup process. The mechanical properties of the final product depend on the fiber orientation and the fiber constraints. However, the mechanical properties are usually weak due to short fibers. The sprayed mixture that consists of fiber and resin is rolled before the specimen is fully cured to avoid bubbles and voids. This method is generally faster than hand layup but has similar disadvantages such as low fiber volume fractions and high void contents [59].

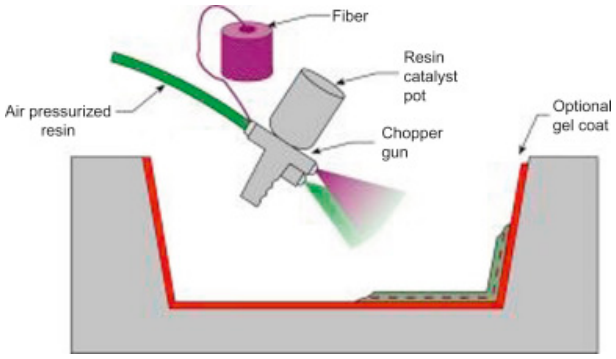


Figure 2.3: Schematic of spray layup manufacturing technique [44]

### 2.3.3 Resin transfer molding (RTM)

In this method, the fiber reinforcement is placed in the bottom half of the mold, and preheated resin is pumped at low pressure, typically less than 100 psi [60], through an injector until the mold is fully filled with resin [61]. Figure 2.4 illustrates the RTM technique. The RTM process can produce composite parts with complex shapes at a high production rate. With proper surface finishing, the quality of the final parts can be comparable to those made in an autoclave [62]. Additionally, this method achieves a good fiber volume fraction. However, a significant drawback is that the resin may not always fully penetrate the entire layup, making this method less suitable for large parts, particularly in the aerospace industry.

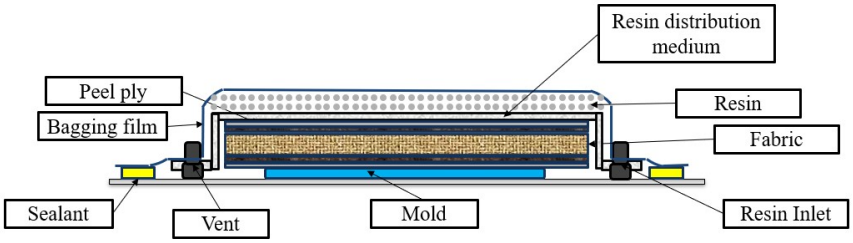


Figure 2.4: Schematic of resin transfer molding (RTM) technique

### 2.3.4 Vacuum-assisted method

The vacuum bagging technique is one of the most famous methods in fiber-reinforced composite manufacturing. This method is a more advanced version of the hand layup technique. Initially, the reinforcement fabrics and some other materials such as peel ply and release film agent are placed on a mold. A breather is also usually used so the air could be removed easily from the vacuum bag. In the vacuum-assisted method, a flexible film made of a material such as nylon polyethylene or polyvinyl alcohol (PVA) is used to enclose and seal the part from the ambient air, called the vacuum bag [63]. Afterwards, the air inside the bag will be drawn out using a vacuum pump. In the out-of-autoclave (OOA) or vacuum bag only (VBO) processes, the ambient air compresses the layup and helps with the consolidation of the laminate. It uses a technique where during the cure cycle, mechanical pressure is created [64].

There are different variations in this method such as vacuum bagging, vacuum assisted resin transfer moulding (VARTM), resin infusion under flexible tooling (RIFT), Seemann composites resin infusion molding process (SCRIMP), double vacuum bag-

ging and vacuum-assisted oven curing. In VARTM, dry fiber preforms are placed in a mold and covered with a vacuum bag. A vacuum is applied to compact the fibers and remove air. Resin is then drawn into the mold under the influence of the vacuum, saturating the fibers [20]. VARTM is suitable for large composite structures. Figure 2.6 shows the schematics of the VARTM process where the curing pressure is equal to the room pressure. RIFT is similar to VARTM but uses a flexible membrane instead of a rigid mold to create the cavity for resin infusion [65]. This variation is more suitable for curved parts. SCRIMP is a variant of VARTM where a flow media is placed over the dry fibers to facilitate faster and more uniform resin distribution [66]. SCRIMP is ideal for large and thick parts [67]. In double vacuum bagging, two vacuum bags are used: one directly over the laminate and a second one over the first bag. This method provides additional pressure on the laminate, helping to remove excess resin and improve fiber compaction [68]. Due to this higher compaction, double vacuum bagging is useful in high-performance application when the highest laminate quality is required. Studies in [69] and [70] show that the lowest porosity and void content compared to wet layup or single vacuum bagging occurs when double vacuum bagging technique was used. In applications where elevated curing temperature is required, vacuum-assisted oven curing is used. In this method, the vacuum-bagged laminate is placed in an oven for curing instead of being cured at room temperature [71].

The vacuum-assisted method uses vacuum to remove excess resin and trapped air from the mold, enhancing adhesion and improving the bond between layers. The consistent consolidation pressure provided by the vacuum results in a better fiber-to-resin ratio, reduced void content, and high-quality laminates in the final parts [72]. This method is also useful for prepreg molding. However, this method has some drawbacks, including the need for additional equipment like vacuum pumps and lines, its labor-intensive nature, and the sensitivity to potential leaks in the vacuum bag [73]. In addition, some of the components such as peel ply, sealant tape and resin tubes are not reusable.

### 2.3.5 Autoclave processing

This method is commonly used in aerospace applications. Although it is expensive, it produces high-performance components that are ideal for aircraft. By applying high heat and pressure, it creates composites that are void-free and denser. The autoclave typically operates at pressures between 3.5 and 7 bars. Studies of Chang et al.

shows that when the curing pressure is higher than 4 bars, the void content decreases to less than 1% [74]. In addition, mechanical properties such as interlaminar shear strength also increase with the curing pressure. In the autoclave manufacturing process, usually, prepregs are first laid up in the desired configuration in a mold and then vacuum-bagged before being placed in the autoclave for curing [24]. Figure 2.5 shows the layup settings inside the autoclave. Parts produced with this method exhibit excellent mechanical properties, high fiber volume fractions, and low void content. However, this method is costly and suitable only for high-performance applications, and it is extremely sensitive to potential leaks in the vacuum bag.

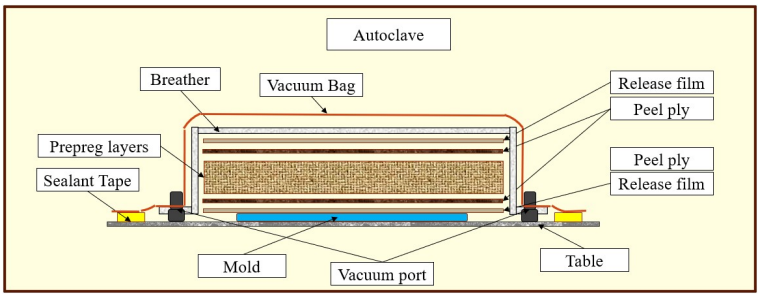


Figure 2.5: Schematic of the layup settings in the autoclave processing

### 2.3.6 Compression molding

In this manufacturing method, materials such as prepregs or fiber mats are placed into a heated mold, which is then closed and subjected to high pressure to shape and cure the composite. As the material is pressed at high temperatures, the increasing mold pressure forces out air from inside the mold, effectively eliminating voids in the composite [75]. Heat can be applied during the compression process if needed. There is also a variant known as cold compression molding [76], where no heat is applied during the process. This variant is performed at room temperature using a hydraulic or mechanical press. The final products often have sharp edges and require trimming after curing.

This process is well-suited for composites with a high fiber volume fraction due to its closed molding and high-pressure application. The final products are typically strong and can be produced in various sizes, depending on the mold. This method also offers a high production rate, with molding times often taking up to five minutes, depending on the size and thickness, before curing at room temperature. Additionally,

it yields products with good quality and a smooth surface finish. However, the method has some disadvantages, including high tooling costs, limitations to relatively simple shapes, and unsuitability for very large parts.

### 2.3.7 Other techniques

There are other manufacturing techniques such as pultrusion, filament winding, injection molding and additive manufacturing (3D printing).

In the pultrusion method, strands of continuous fibers are pulled through a resin bath, which are further consolidated in a heated die [77]. It is a continuous process, useful for fabrication of composites with a constant cross-section with a relatively longer length; it enables production with a high degree of automation and lower production cost [78].

The filament winding process is similar to the pultrusion process, where the fiber is pulled through the resin bath and then collected by a mandrel [79]. This process can be used to produce symmetrical shapes such as pipes [80].

In injection molding, the composites are produced by forcing molten material into the mold, where the material will be l left until cured [81]. It is suitable for thermoset or thermoplastic composites. The fibers, usually chopped fibers, and the pellets go through a funnel-shaped hopper with a rotating screw into a heated compression barrel [82]. Because of using chopped fibers, the mechanical properties of parts produced with this method are usually low.

In additive manufacturing method also called as 3D printing, continuous or short fibers are mixed with resin and printed layer by layer to build up a part. This method eliminates the requirement for mold and offers customization and complexity of the parts [83], but it is has limited material choices, slower production rates compared to traditional methods, and the need for post-processing [84].

### 2.3.8 Focus of this study

The objective of this paper is on any method involving vacuum bags in the process. This will include any manufacturing technique sensitive to leakages on the vacuum bag. This will include all different variations in the vacuum-assisted method is mentioned in Section 2.3.4. Additionally, autoclave processing introduced in Section 2.3.5, a primary method in the aerospace industry, which is extremely sensitive to leaks will be also a focus of this study. Leakages in vacuum bags can disrupt the vacuum or

curing pressure, preventing the system from achieving the necessary pressure differential required for proper consolidation. In the following section, a literature review will examine the effects of vacuum and curing pressure on the quality and properties of the final products.

## 2.4 Effect of vacuum and curing pressure on fiber composites

Using vacuum in the RTM process leads to an improvement in the terms of composite quality and mechanical properties [85]. Figure 2.6 shows the schematic of the VARTM process. The pressure inside of the vacuum bag could be reduced by a vacuum pump and it is called the vacuum pressure. On the other side, the curing pressure could be increased inside of an autoclave. The curing pressure could sometimes get up to 7 bars in an autoclave. In out-of-autoclave manufacturing processes, the curing pressure would be equal to the ambient or atmospheric pressure. The pressure difference between curing pressure and ambient pressure helps with the consolidation of the composite material.

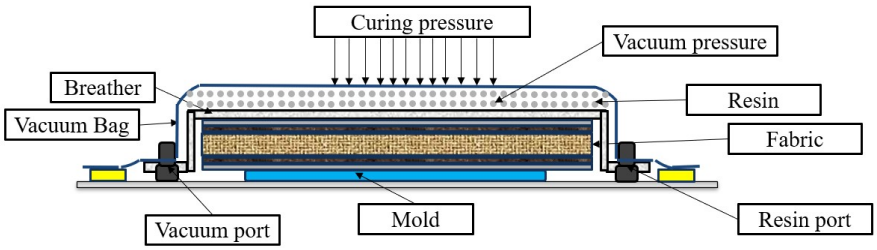


Figure 2.6: Schematic of the VARTM process; Pressure difference between the curing pressure and vacuum pressure helps with the consolidation of the fabric

Final mechanical properties of polymer composite materials highly depend on their curing parameters in the autoclave or oven, including curing pressure and vacuum pressure [86]. The void content and fiber volume fraction of a VARTM part can be affected by many parameters and is critical to the mechanical properties and the quality of the part. Stronger vacuum and higher mold temperature can better control and increase the fiber volume fraction. Verma et al. did a study on carbon fiber composited manufactured with and without vacuum [87]. Their results show more than 100% improvement in tensile strength and around 50% improvement in

hardness when parts were produced with vacuum. The effects of curing pressure on interlaminar shear strength (ILSS) has also been studied [74].

Voids are one of the most harmful defects in composite parts made by fiber-reinforced fabric [88]. Voids weaken mechanical properties such as shear and compressive strength mainly because the voids reduce the load-bearing cross-sectional area of composite laminates [89]. As a result, the stresses on selected fibers increase which, in turn, leads to the localized stress concentrations in the laminate and end up with poor mechanical properties and reducing fatigue life of the composites [90].

The void formation in a polymer composite is governed by many factors. Factors such as the vacuum pressure, the inlet pressure, curing pressure, curing time and the mold temperature are identified as the major factors for the void formation during the VARTM process [91]. Entrapped air in the laminate during layup [92] and dissolved moisture in the resin [93] are other factors causing void formation in composite materials. Boey et al. analyzed the effect of the bagging vacuum and the autoclave pressure on the void content, and subsequently on the bending strength and the modulus [94]. Their results indicated that increase in the vacuum bagging and achieving a lower pressure inside the vacuum bag decreases the void content exponentially. Their results show that the increasing the autoclave pressure to 12 bars could decrease the void percentage to more than 50 percent. Not achieving a proper vacuum could increase void content to four times compared with a proper vacuum as represented in Figure 2.7. In addition, better vacuum could decrease void sizes significantly [95].

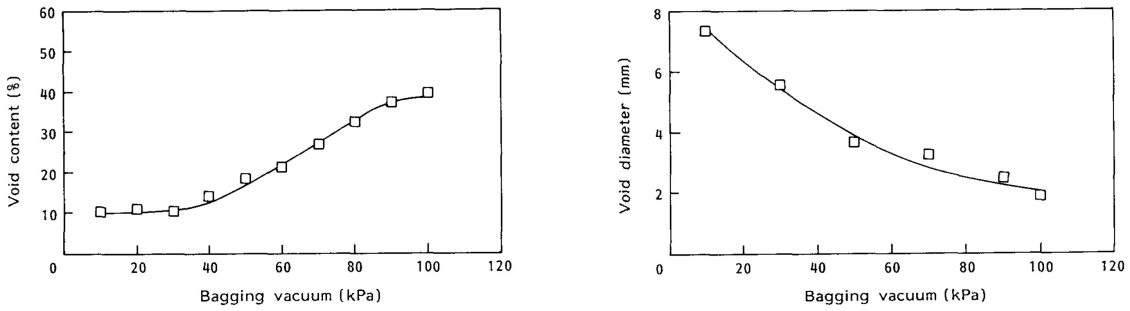


Figure 2.7: Effects of vacuum pressure on void content and diameter; lower bagging pressure means a vacuum with a higher quality

The experimental results of [96] indicate that a higher mold temperature combined with a higher vacuum at the vent (corresponding to a lower vent pressure) effectively increases the fiber volume fraction and enhances its consistency. This is

due to the reduction in resin viscosity and the increase in thickness-direction compression, respectively. While increasing the vacuum generally has a positive effect on minimizing voids, this is not always true for higher temperatures. These findings demonstrate that achieving a higher vacuum can lead to an improved fiber volume fraction and reduced void content.

The vacuum level in the bag significantly impacts surface porosity. Surface porosity also affect mechanical properties and studies in [97] showed that the presence of 25% porosity in a glass fiber reinforced polymer could cause a 50% reduction in the shear strength of the final part. Generally, trapped air inside the vacuum bag leads to surface porosity [98]. Furthermore, studies in [99] confirm that surface porosity is primarily associated with air trapped at the tool–prepreg interface during layup. Several factors can influence the magnitude and distribution of surface porosity, including vacuum hold time, freezer and out time, and modifications to materials and processes that affect air evacuation [100]. The results suggest that prepreg out time (and consequently tack) and vacuum quality are the primary contributors to surface porosity. Inadequate vacuum in the bag increases bag pressure, reducing the driving force for air evacuation and allowing more air to remain in the bag. Figure 2.8 shows trapped air on a laminate under vacuum condition. The visualization shows the reduction of voids in size as vacuum condition is present in the layup. Void contents is usually measured by water immersion. However, this method is not useful for measuring surface porosity or visible surface voids. Acid digestion methods according to the ASTM D3171 standard could be used to obtain void content and fiber volume fraction [101].

Vacuum pressure and the curing pressure also play a crucial role in the thickness of the final part. It is usually a challenge to manufacture a composite part with small dimensional tolerances [102]. Yenilmez et al. studied the effect of different factors on the final part thickness and concluded that duration of the vacuum and the curing pressure as well as shrinkage ratio of the resin have the most influential factors [103]. Experimental results represented in Figure 2.9 show how higher pressure difference could influence thickness and fiber volume fraction for ten different experiments [104]. Higher pressure results in more accurate and valid thickness and higher fiber volume fraction. Failing to achieve adequate vacuum and curing pressure result in high tolerance in the thickness and geometry of the final products.

Optimized vacuum infusion processing is a suitable method for producing thick composites up to 12 mm in thickness [105]. However, inadequate vacuum can result

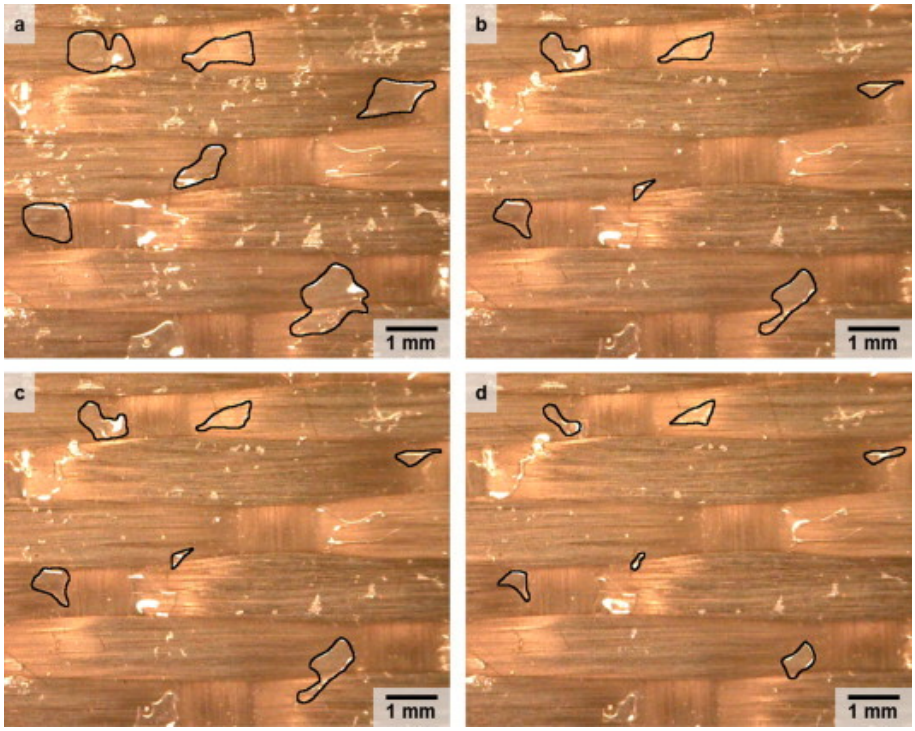


Figure 2.8: Images of glass plate–prepreg interface under vacuum at room temperature in different times; (a) immediately after the application of the vacuum, (b) 10 min, (c) 30 min and (d) one hour into the vacuum hold

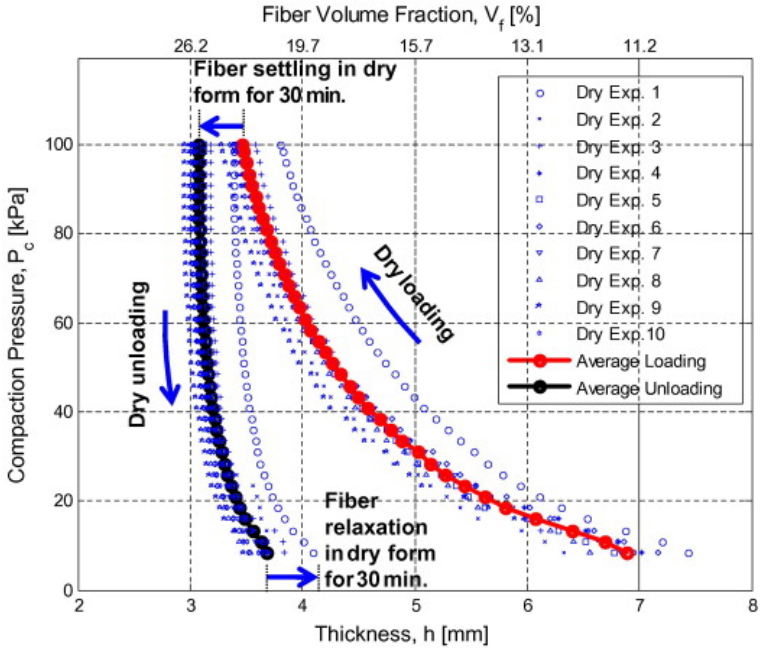


Figure 2.9: Diagram of part thickness and fiber volume fraction as a function of curing pressure for different experiments

in resin accumulation because the low pressure may not be sufficient to drive the high-viscosity resin through the porous fabric preform effectively. These infiltration heterogeneities can lead to variations in the physical and mechanical properties of the part, particularly in properties less dependent on fiber characteristics, such as short beam strength [106]. Short beam strength, also known as short beam shear strength, measures the interlaminar shear strength of laminated composites and is determined using a standardized short beam shear test based on ASTM standards [107]. This testing method is highly sensitive to defects within the laminates, making it an effective tool for assessing composite quality [108].

In this section, the literature behind vacuum and curing pressure effects in composite manufacturing was reviewed. Leakages in the layup can disrupt pressure differentials locally or throughout the entire part, leading to undesirable surface porosity, increased void content, lower-than-expected fiber volume fractions, and compromised mechanical properties. Therefore, it is crucial to achieve a leak-free, perfectly sealed vacuum bag and to locate and repair any leaks before and during the curing process. In the following sections, state-of-the-art methods for leakage detection and localization are presented. Active monitoring of leakages during the curing process is essential, as maintaining vacuum integrity throughout curing is as important as starting with a leak-free vacuum bag. Consequently, leakage detection methods that can be applied within the autoclave are particularly advantageous.

## 2.5 State-of-the-art Leakage Detection Methods

Numerous technologies are available to address the issue of leakage detection and localization. The initial task involves determining the presence of any leakages within the vacuum bag. Once confirmed, the subsequent step is to pinpoint their exact location.

The simplest method for leakage detection is called the pressure rise method [109], where the air inside the vacuum bag is evacuated and after disconnecting the bag from the pump, the pressure inside the vacuum bag is measured continuously for some time (5 to 15 minutes). A leakage exists in the system if the inside pressure rises more than an allowable threshold (5 to 15 $mbar/min$ based on [110]).

Helium has also been widely used for leakage detection [111]. This can be conducted in one of 3 ways: Detector-probe technique, Tracer-probe technique, or Hood technique [112]. The detector-probe variation is conducted by pressurising a compo-

nent with helium gas, and then scanning the component by the detector probe. This method is also known as the Ultra Sniffer Test (UST) [113]. This method is useful for positive pressure and is not useful in composite manufacturing because the vacuum bag cannot be filled with any gas including helium. Both trace probe and hood technique are conducted under vacuum or relative negative pressure and are applicable to composite manufacturing. In the tracer-probe method, the tracer is used to spray helium around the component under test. If there is a leak in a bag, the Helium will be drawn into the vacuum and the detector will sound an alarm [114]. The hood test method is similar but instead of using a probe, a good or envelope will be installed around the portion of the component under test and this envelope which is usually a plastic material or bag, will be filled with helium [115]. Similar to the tracer-probe test the helium will be drawn into the part due to the differential pressure and the mass spectrometer will be monitored to verify the presence of a leakage. Figure 2.10 shows the helium hood leak detection technique. Using helium leak detection method is highly sensitive, detecting leakages in the order of $10^{-11} mbar/s$.

Figure 2.10: Helium gas leakage detector hood technique used on a vacuum bag; a large area could be covered with this method [29]

Other gases such as Hydrogen, Methane, Carbon Dioxide, Sulfur Hexafluoride and other noble gases can also be used for leak detection. The most important requirements for a tracer gas are having environmental compatibility, chemical and physical inertness and having a low content in standard environments [116].

Another effective method for leak detection involves monitoring the flowrate values on vacuum lines [25] [28]. In large components, multiple vacuum ports in different locations of the layup are required to effectively evacuate the air from inside of the vacuum bag. The flowrate measurement sensors could be installed on the evacuation lines, act as sensors with multiple channels. These sensors could continuously measure air flow rate on evacuation lines. Under normal operating conditions, when the system reaches a steady-state state, the flowrate values on the vacuum lines should indicate zero. This indicates that there is no air remaining within the vacuum bag or entering the system that needs to be pumped out. In steady state and in presence of a leakage, air will continuously enter into the vacuum bag from the leakage and pumped out from the vacuum pumps. Therefore, the sensors will show non-zero flowrate values, indicating the existence of at least one leakage in the system. Direct flow measurement has many advantages. It is independent of part volume or configuration, and it's not sensitive to temperature changes [117]. Figure 2.11 shows an example of using flowrate measurements on evacuation lines [118]. This particular setup is consisted of six independent evacuation channels.

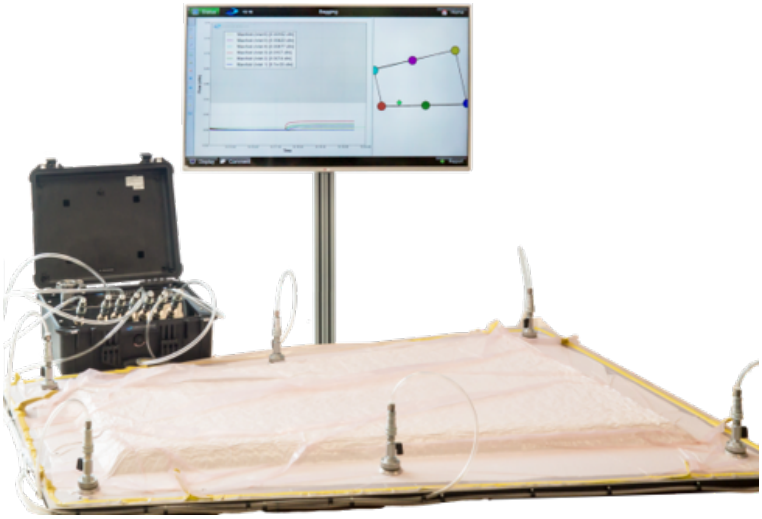

Figure 2.11: Flow rate measurement on evacuation lines for leakage detection; non-zero flowrate values in steady state indicate presence of a leakage [35]

## 2.6 State-of-the-art Leakage Localization methods

Once potential leakages are identified in the layup, their precise locations must be determined. Several state-of-the-art methods in the industry can be used to pinpoint

or localize these leakages. Although methods in Section 2.5 are simple and can detect leakages in the system, they are incapable of locating leakages.

When a vacuum bag is under low pressure, any small leak in the bag causes gas to enter into the layup. This entering gas produces high-frequency sound waves due to the turbulent flow of gas through the leak site. The high-quality ultrasonic leak detector operates using a microphone optimized for the high-frequency sound produced by gas flow, enabling the identification of leakages [119]. The accuracy of ultrasonic leak detection depends on several factors, including the operator's experience level, environmental background conditions, the ability to isolate unrelated background noise from the frequency of interest, and the type of leak (e.g., sealant tape leak or surface of bag leak). Although this technology is capable of detecting small leaks, it is highly sensitive and can be easily disrupted by external noise and vibrations, making it ineffective in autoclave environments. Additionally, there is a limitation concerning the distance between the sensors and the leakages, which renders the technology unreliable for large components. To improve localization accuracy, an array of ultrasonic microphones can be used in conjunction with an algorithm to finely pinpoint leakages [120]. Typical detection distance for leakages are between 30 to 50 cm, with a lower limit of 10 L/h for an advanced and industrial ultrasound leak detector [121] shown in Figure 2.12. Multiple leakages may also cause overlapping and signal interference in this technology.

Figure 2.12: Ultra CE; an advanced ultrasound leak detector using 124 microphones, detecting leaks in quiet environments [37]

Piezoelectric sensors [122] could be used to measure strain variations and detect leaks [123] [124]. Typically, this sensors are made of certain materials such as quartz and ceramics, generating an electric charge in response to mechanical stress. When a leak occurs in a vacuum bag, the entering air or gas can create acoustic waves. Piezoelectric sensors can pick up these acoustic emissions, even at low levels, making them effective for early leak detection. These sensors, which have high sensitivity, high resonance frequency and good stability,[125] can be placed in different locations on the assembly and the leakage location can be calculated using triangulation[126]. Using piezoelectric sensors is highly effective against small leakages even in large assemblies. In addition, the method is non-intrusive and the sensors could be attached externally without significantly altering the manufacturing process. Similar to the ultrasound microphones, piezoelectric sensors are sensitive to various environmental factors, such as temperature changes and mechanical vibrations, which can potentially lead to false alarms and therefore not suitable for autoclave. In addition, the process is rather complex, as an extensive calibration and signal processing is required. Integrating piezoelectric sensors into composite manufacturing processes requires careful planning, as the sensors need to be placed in locations where they can effectively monitor for leaks. Therefore, generalizability of this method would be extremely difficult. Another drawback of this method is its lack of reliability and effectiveness when there are multiple leakages in the system.

Infrared thermography is a method used to detect leaks using an infrared camera. Due to the Joule-Thomson effect, regions in the vicinity of leakages appear as regions of lower temperature in the infrared spectrum[127] [128]. The air expands after falling into a lower pressure as it enters into the vacuum bag through the leaks. This expansion will cause cool down and will be visualized as cold spots in the thermal camera. This contact-less method is easy to use and it can be automated by computer vision. In addition, it could assist with identifying multiple leakages at once [129]. However, it is difficult to capture the whole assembly with a single stationary camera as thermal imaging camera scope is not able to capture the detail required for leak detection. Therefore, the camera should manually or automatically moved across the layup to cover the whole area. In addition, this method is not reliable as it may overlook small leakages due to their weak cool down effect. In addition, this method may struggle to pinpoint the leakage as cold spots could appear in a relatively large vicinity. Figure 2.13 shows the application of using a thermographic camera for leakage localization. Leaks will be visible as cold spots on the setup. Note that the

cold spots have been projected into the layup in this particular image.

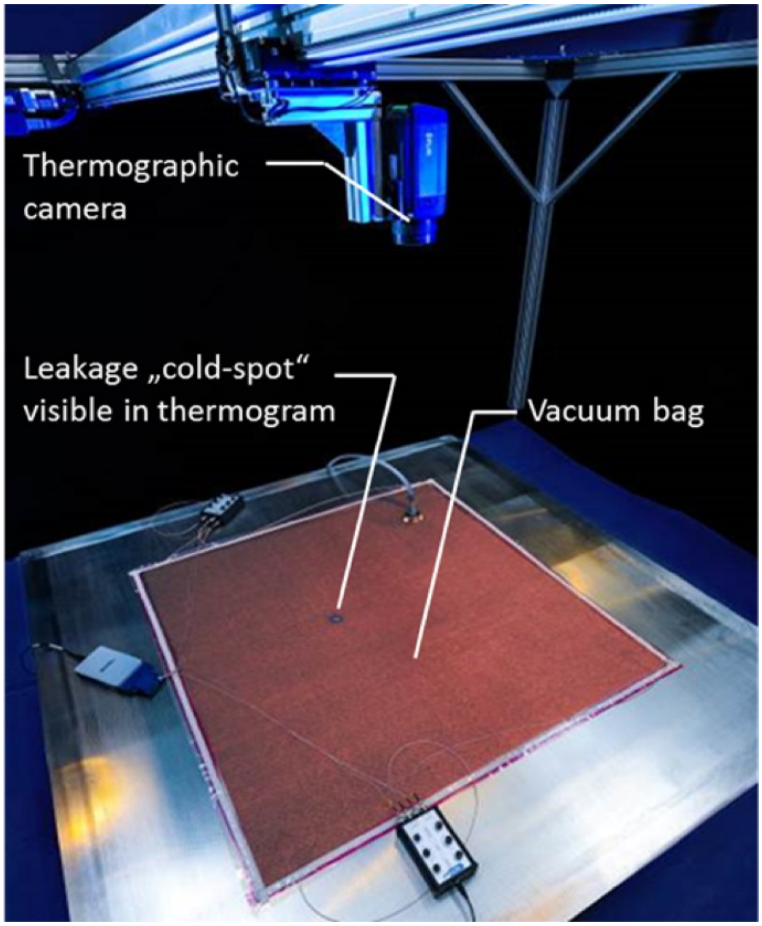


Figure 2.13: Leakage localization using infrared thermography. The thermal image acquired is projected onto the layup under test

Schlieren imaging has been explored for its potential to pinpoint leaks in vacuum bags [130]. This technique visualizes density variations in transparent media [131] and has been extensively used in aerospace applications to study supersonic motion and visualize airflow [132]. This method is non-invasive and can address multiple leaks at once, but it shares similar limitations with infrared thermography, such as a relatively small field of view, requiring the Schlieren imaging setup to be repositioned around the layup to cover the entire area. In addition, the method may struggle to identify small leaks with minimal airflow. Moreover, since we are dealing with a negative pressure on vacuum bags, meaning that the air will enter into the vacuum bag and would not exit, adding air density disruptive components such as gas and heat may be required for a proper visualization. Figure 2.14 presents the results of

using Schlieren imaging for leak detection. The image on the left shows heat from an extinguished candle being drawn into the leak, while the image on the right captures the candle's thermal plume being similarly pulled into the leak.

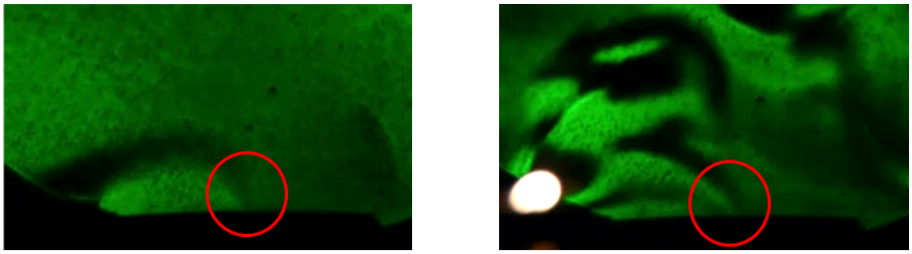


Figure 2.14: Application of Schlieren imaging in vacuum bagging leakage localization; (a) Heat from an extinguished candle being pulled into the leakage; (b) Candle's thermal plume being pulled into the leakage [47]

An Oxygen-sensitive material can be placed underneath the vacuum bag which will undergo a change of appearance when exposed to oxygen at the leakage location, allowing for visual identification of leakages [133]. The drawbacks of this method are two-fold. First, it requires the use of an additional layer in the layup. Second, it presents difficulties in visually locating small leaks. However, the advantages are simplicity and ability to address the problem when dealing with multiple leakages.

Another visual method similar to the oxygen-sensitive film is applying a fluorescent coloring agent or dye to the exterior surface of the vacuum bag. If a leak is present, the dye will be drawn into the vacuum bag. The leakage appear as dyed spots on the assembly and can be detected by UV lamps in a dark environment once the exterior dye is removed. This method is easy to implement yet highly effective, especially when there are multiple leakages. Unfortunately, this method is time-consuming and inefficient because it requires extra materials in the assembly and steps in the lay up process such as applying the agents and cleaning the dye [134]. In addition, the method is not applicable to the inside of the autoclave layups.

In 2019, Haschenburger et al. compared different technologies with criteria such as time, cost, handling, component size and several leakage detection [135]. The results of this analysis is shown in Figure 2.15. Most of these methods are time-consuming, costly and rely on the experience of the operator. Therefore, a more reliable, fast and automated method for locating leakages in vacuum bags is required. Additionally, the automation of leakage detection is an important step in improving the cost efficiency of the process [136]. Our proposed method is based on volumetric

flow rate measurement and and the application of intelligent methods to enhance its autonomy.

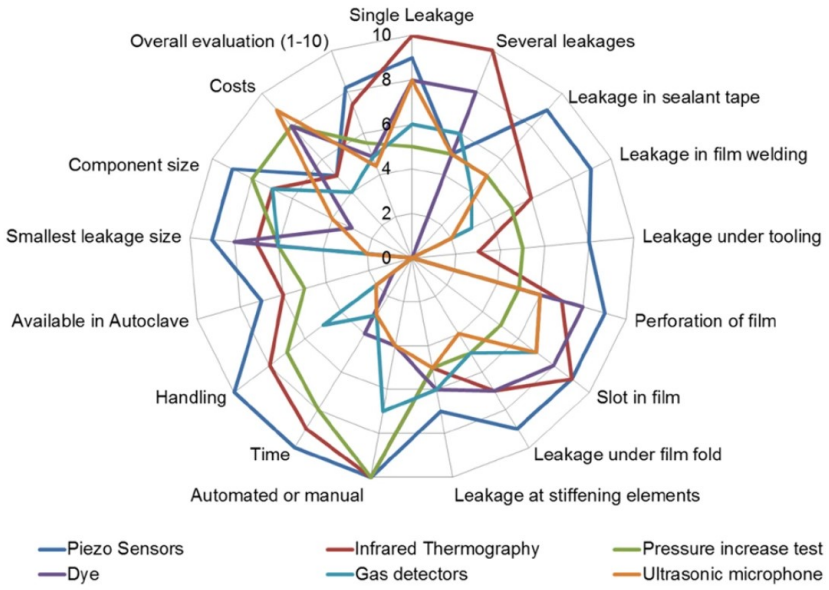


Figure 2.15: Comparison of six different leakage localization methods with different criteria on a scale of zero to ten done by [135]

Volumetric flow rate measurement is valuable for both leakage localization and detection. Menke utilized volumetric flow rates from vacuum ports and applied a trilateration method for leakage location prediction [25]. In 2022, Haschenburger explored four numerical methods on a symmetrical square-shaped tool [137]. These methods included potential flow modeling, numerical regression, volumetric flow rate matching, and machine learning models.

In the potential flow method, vacuum ports are modeled as sinks and the leakage as a source, allowing the complete velocity vector field to be determined. The most probable leakage area is identified by minimizing the difference between the measured flow rates and the average velocity ratios of corresponding areas. The numerical regression model establishes a relationship between the flow rate and the distance of the leakage from the vacuum port. This approach can be considered a simple form of a machine learning model. The volumetric flow rate matching method estimates the radial distance of the leakage from the vacuum ports. A grid is constructed, and the measured flow rates are compared with the calculated flow rates at each grid node. This approach resembles a complex multi-dimensional truth table, where each element of the table is a vector rather than a scalar.

One of the state-of-the-art automated methods for leakage localization is the application of machine learning and artificial intelligence to predict leakages based on measured flow rates from different vacuum ports [138]. Studies show that it is possible to locate leakages on vacuum bags by using experimental or generated data [139]. In 2022, Brauer et al. trained a neural network to predict leakage locations on a symmetrical square-shaped tool [140]. Most studies utilizing machine learning for leakage localization have focused on symmetrical and simple geometries, with data acquisition primarily performed manually. In some cases, additional data has been synthesized through rotation and symmetry.

However, the most novel AI-based approaches often lack generalizability to other geometries due to their reliance on manually acquired data. While these methods may involve complex setups, expensive sensors, and require medium to large datasets for training, they are highly versatile and capable of generalizing across various geometries and large components, provided sufficient data is available. Furthermore, they are highly automated, fast, and applicable even within an autoclave environment. With access to diverse datasets, these techniques could enable large-scale and complex layup leakage localization.

# Chapter 3

# Methodology

This chapter will outline the methodology used in the study. First, the experimental setup will be introduced, followed by a detailed explanation of the experimental procedures and data acquisition process. Next, the simulation process, including the methods for simulation and verification, will be discussed. Key challenges will also be addressed, such as the effects of leakage size, non-uniform permeability in the setup, and the presence of multiple leakages. Methods for addressing each of these challenges will be examined. Additionally, the placement of vacuum ports on the layup, which presents a gap in the leak localization problem, will be explored, along with strategies for optimizing their placement. Finally, the chapter will discuss various machine learning models, regression techniques, and methods for training them.

To effectively train any machine learning system, it is essential to have extensive and diverse datasets, regardless of the training algorithm used. In our case study, the datasets should encompass a wide range of layup sizes and geometries, vacuum port locations, and permeability variations. In fact, a well-known adage among artificial intelligence (AI) researchers highlights the importance of data over algorithmic superiority: “it’s not who has the best algorithm that wins, it’s who has the most data”. This aphorism from artificial intelligence (AI) researchers shows just how essential data is in machine learning [141].

Given the need for large datasets, conducting thousands of flowrate experiments becomes impractical. Therefore, it becomes necessary to generate artificial datasets. Keeping this in mind, we are in search of a method that is both reliable and efficient in simulating and synthesizing data, while also being applicable across various scenarios. Extensive research has already been conducted on modeling the resin flow mechanism and pattern in the mold [142, 143, 67, 144]. however, they are not helpful as there is

no model or prediction for flowrates in the presence of leakages.

In the next section, the experimental setup will be introduced.

## 3.1 Experimental Setup

The experimental setup is consisted of two main parts.

1. Layup stack: referred as the arrangement of mold, reinforcement layers and the vacuum bag shown in Figure 3.1.

2. Flowrate measurement system: vacuum pump, vacuum ports and connectors and the flowrate measurement sensors.

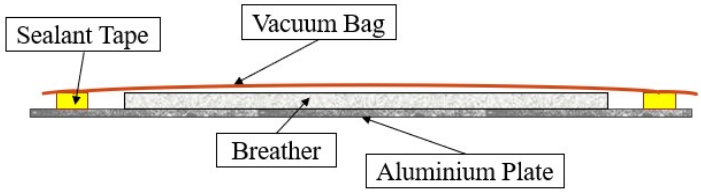


Figure 3.1: Schematic of the experimental setup

In the experimental setup, an aluminum plate has been used instead of the mold. With this setup, the complexities associated with the mold's 3D shape and the challenging behavior of air in the corners could be eliminated from the experiments and simulations. Two systems will be connected via vacuum connectors or ports. The vacuum connector which is shown in Figure 3.2 includes a valve that can isolate the two systems.

Figure 3.2: Vacuum connector or port which connects the layup to the vacuum system; the valve could be used to disconnect the system

During the experiments, the layup stack typically varies to accommodate different experimental conditions. These variations include changes in size, geometry, the number of material layers in the layup to adjust permeability, and the positioning of the vacuum ports. In contrast, the flow rate measurement system remains constant throughout the study, as its sole purpose is to evacuate air and measure flow rates.

### 3.1.1 Layup stack

The fiber-reinforced fabric is typically mounted on a mold, as shown in the schematic in Figure 3.1. To facilitate airflow beneath the vacuum bag, a porous medium, commonly referred to as a breather, is used. Without the breather, air would be unable to pass through the aluminum plate, the mold, or between the reinforcement layers, remaining trapped in the layup. Figure 3.3 illustrates our experimental setup, which follows a standard vacuum bagging configuration.

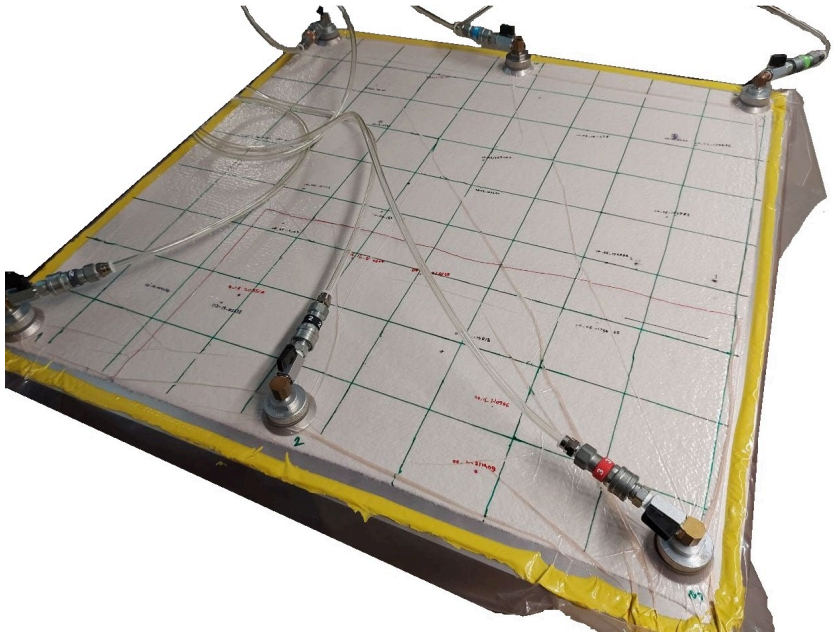

Figure 3.3: Experimental setup layup stack with the size of 90cm × 90cm

In this setup, an N10 breather is affixed to a square-shaped aluminum plate measuring one meter in width. The geometry of the breather which is the region that air could travel is a square with the size of 90cm. The vacuum bag is then securely attached to the aluminum plate using sealant tape. To ensure the integrity of the assembly, a vacuum test is performed to identify and rectify any potential leaks. The

vacuum test method is done by measuring flow rates on evacuation lines introduced in Section 2.5. The absence of leaks is crucial as the experiments focus on investigating leakages on the surface of the vacuum bag. Thus, starting from a leak-free state is essential to gather standardized leak data. Six vacuum connectors, which will connect vacuum pump are also placed in corners and the middle. The COHO flowrate measurement system will measure air flow in these locations.

The configuration of the layup stack varies across different experiments, allowing for flexibility and the creation of varying permeabilities within the system. Certain sections of the breather may be selectively removed to introduce non-uniform permeabilities. On the other hand, more layers of breather could be added to the layup to increase the permeability in certain regions. Adding more layers of breather provides more porous media for the air to travel and therefore increases the permeability. Consequently, the vacuum bag is pressed against the aluminum plate, effectively restricting airflow within those specific areas. Additionally, other components, such as angle irons, can be introduced beneath the vacuum bag to intentionally create highly non-uniform permeabilities within the system. This approach enables the exploration of extreme permeability variations in the experimental setup.

### 3.1.2 Flowrate measurement system

To conduct the experiments, we employed the COHO 601 gas flow and vacuum leak detection system, as referenced in [145]. The COHO system is specifically designed for composite manufacturing applications and incorporates high-performance gas flow sensors. These sensors are seamlessly connected to the standard vacuum lines utilized in the process. Figure 3.4 shows the COHO 601 gas flow measurement system.

Figure 3.4: COHO 601 flowrate measurement setup

The COHO system features an onboard computer that instantaneously processes the sensor data, allowing for real-time analysis. This enables the operator to access precise gas flow measurement data and view flow-rate graphs at a frequency of 1 Hz. Notably, this setup offers the capability to measure air flow values independently on six evacuation lines, also known as vacuum ports. Such a configuration enhances the system's versatility and enables simultaneous monitoring of multiple evacuation lines during the experiments.

## 3.2 Experiment Procedure and Data Acquisition

After connecting the COHO gas flow measurement system to the layup through vacuum ports, it is crucial to ensure the absence of any leaks in the system. The leakage detection technique described in Section 2.5 should be employed for this purpose. Once the system integrity has been verified, data acquisition can proceed.

To gather data, intentional and controlled leakages must be created on the vacuum bag. To create these leakages on the vacuum bag, we utilized a needle with a diameter of 0.5 mm. These holes were carefully placed at designated locations on the bag and their respective flow rates were recorded. Throughout each experiment, both the hole positions and the corresponding flow rates were documented. This iterative process, illustrated in Figure 3.5, can be repeated to gather additional data for testing or validation purposes. It is essential to ensure the proper sealing of the vacuum bag following each data acquisition. This can be accomplished by confirming that the flow rate for each port remains below the predetermined threshold of 0.02 L/min in every experiment.rt is below the threshold of 0.02 L/min for each experiment.

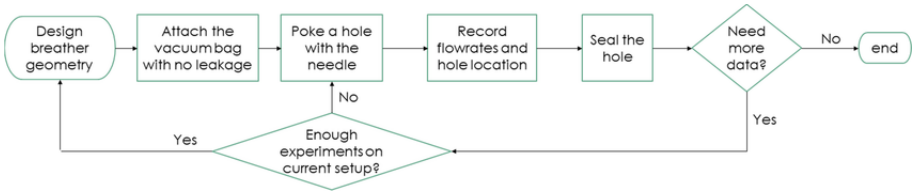


Figure 3.5: COHO 601 flowrate measurement setup

The COHO system measures the flowrates at a frequency of 1 Hz. The output of the sensor values for a specific example is illustrated in Figure 3.6. In this example, a leakage has been created in the system by poking a hole using the needle at the time of 45 seconds in the time a. Once the hole is poked and the needle is removed,

an overshoot appears in the graph and the values gradually settle. When the values have settled, the hole is covered, and the flowrates drop. There are non-linearities and noise in the system that may result in overshoots or oscillation in flowrates. To remove these effects, average flowrates must be extracted. Afterwards, the average flow rate values must be normalized; meaning that their percentage is represented instead of their absolute values. It is best to average the values after the overshoot and before covering the leakage. In our experiment, the vacuum pressure is 85,000 pa. This is the number introduced in Section 2.4 and plays an important role in the manufacturing process. Existence of leakages on the layup results in the reduction of this number. It is important to note that this value may vary depending on factors such as the vacuum pump and the breather type used [146]. Using a more powerful vacuum motor that can vacuum closer to zero pressure (about -100 KPa) will also decrease the oscillations in the flowrates.

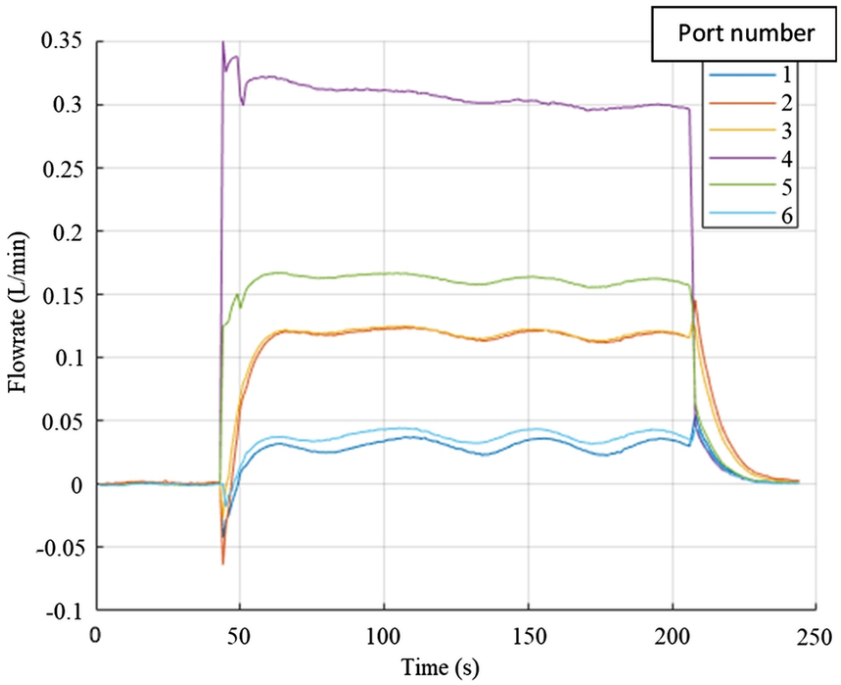


Figure 3.6: Flowrate outputs for six vacuum ports for a sample experiment; initially, there is no leakage in the vacuum bag and the flowrate values are zero in steady state. As a leakage is created at time t=45 (s), the flowrate values increase. Eventually, the leakage is covered at time 210 (s) and the flowrate values gradually tend to zero

In the experiment depicted in Figure 3.6, a steady-state behavior is observed between 80 and 150 seconds, with oscillations being disregarded. This indicates that all the air entering the vacuum bag through leakage is evacuated by the vacuum

pumps. The flow rate will not reach zero, as air continues to enter from the higher-pressure atmosphere into the lower-pressure vacuum bag unless the leakage is sealed. The sealing of the leakage happens just after 200 seconds and the flowrates eventually settle on zero. During the steady-state interval, vacuum port 4 exhibits the highest flow rate, followed by vacuum port 5, with ports 2 and 3 having lower flow rates. Vacuum ports 1 and 6 show the lowest flow rates, contributing minimally to the evacuation process.

The higher flow rate at port 4 suggests lower resistance for air reaching that port. Assuming uniform permeability in the system, it can be inferred that the leakage is closest to vacuum port 4, followed by port 5, and farthest from ports 1 and 6. This relationship could be modeled by training a machine learning algorithm with sufficient data. Figure 3.7 illustrates the actual location of the leakage (in red) and the vacuum port numbers. As shown, the leakage is closest to vacuum port 4, explaining why it generates the highest flow rate.

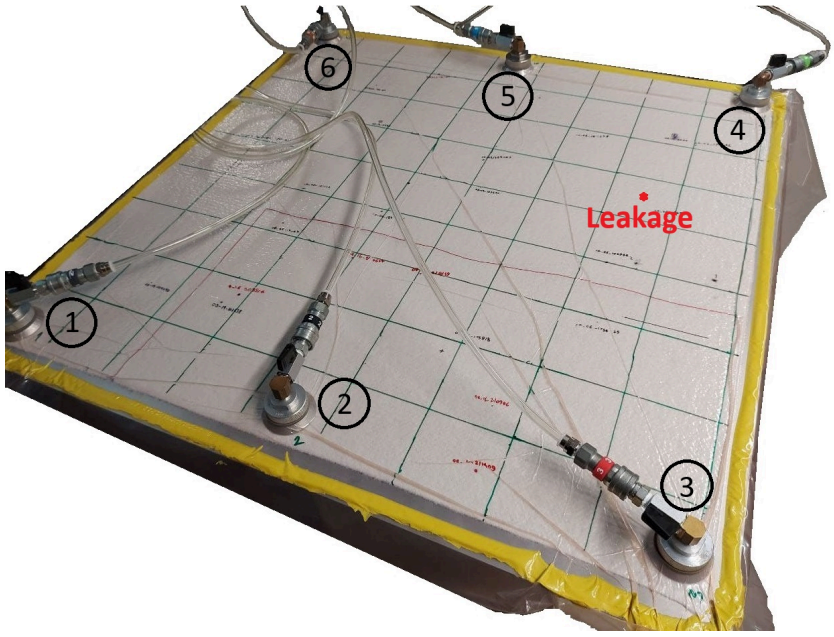


Figure 3.7: Experimental setup of the example with the vacuum port numbers and actual location of the leakage

Current state-of-the-art research primarily focuses on uniform, symmetrical, and lab-scale setups. However, in industrial applications, more complex and larger systems are common and must be addressed. As a starting point, a similar setup was used for experimentation, and the system's outputs were recorded. Additional config-

urations can be designed to further complement the dataset. One of the key challenges is predicting leakages when sections of the breather are removed, or when the permeability of the layup is non-uniform and asymmetrical. The breather geometry and the locations of the vacuum ports are additional variables that must be documented for each setup.

For each configuration, a grid was overlaid on the breather, as illustrated in Figure 3.8, with each square of the grid being sampled to produce a comprehensive dataset. Leakages were introduced at the center of each square using a standard 0.5 mm diameter needle. In the next section, other important experiments and layup configurations that are feasible within our lab-scale setup are considered, aimed at increasing the system's complexity.

## 3.3 Experimental Configurations and Setups

One of the key objectives and contributions of this research is the introduction of a leakage localization framework that is generalizable across different geometries and configurations. Consequently, it is vital to gather data from a diverse range of setups. Additionally, it is necessary to validate the simulation method on these specific configurations. To achieve this, various approaches such as altering the size and shape of the layup or introducing non-uniform permeable elements into the setup could be employed. In our experimentation, we utilized rectangle-shaped layups with different dimensions as well as an octagonal-shaped layup. Various configurations feasible with our lab-scale setup has been introduced in the following.

### 3.3.1 Changing the size and geometry of the layup

Initially, the experiments were conducted on a square-shaped layup measuring 90cm on each side. This layup was placed on an aluminum plate, serving as the mold, which measured one meter. The layup size could not exceed the plate's dimensions, setting a boundary for the experiments. Additionally, using a smaller layup would oversimplify the leakage localization problem. With these constraints in mind, we designed two different layups.

The first layup is rectangular, with a different aspect ratio than the square. We modified the 90cm by 90cm layup to a size of 81cm by 72cm. This change disrupted the symmetry in one direction, introducing more complexity and diversity into the

experiments. This setup is shown in Figure 3.8.

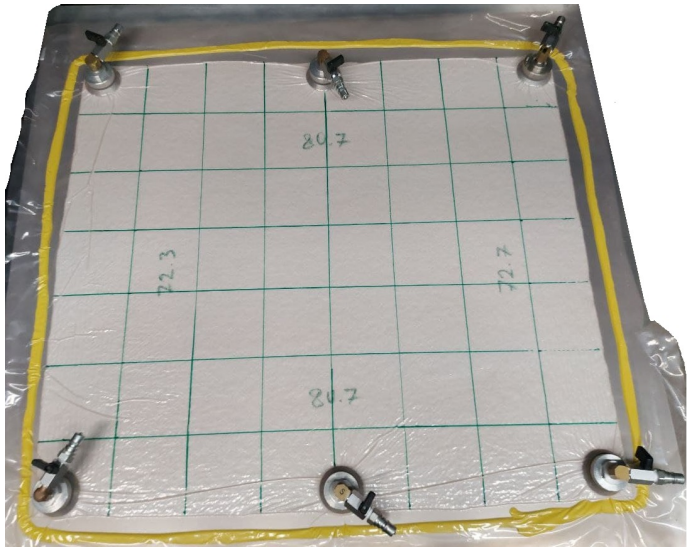


Figure 3.8: Rectangular layup used in the experiments with the size of 81cm × 72cm

Improper bonding between the sealant tape and the bag or tool can lead to boundary leaks. Typically, boundary areas are more prone to leakage compared to the central regions of the vacuum bag, where sealant or tacky tapes are not present [147]. To conduct more comprehensive experiments focused on boundary leaks, an octagonal-shaped breather, as shown in Figure 3.9, was created.

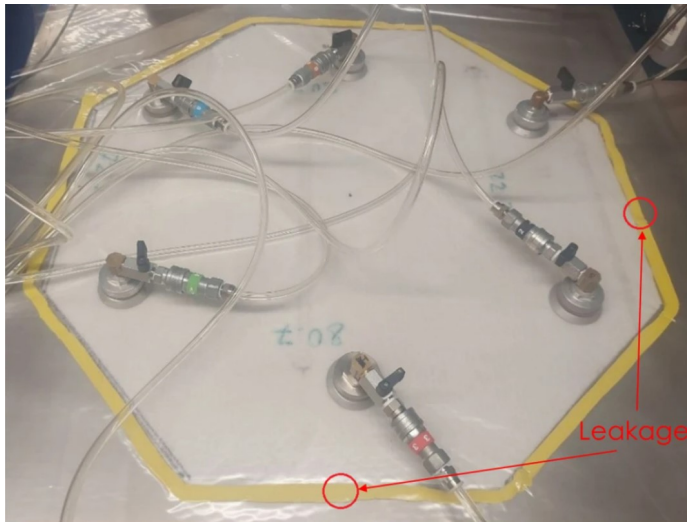


Figure 3.9: Octagonal-shaped breather for experiments on leaks on the boundary

Leakages are created on the boundaries and different sides of the octagonal-shaped

breather as presented in the figure. After performing experiments and generating sufficient data, the objective for the system is to correctly predict the location of the leaks on the sides of the octagon. However, before achieving that, a proper simulation must be designed and verified to solve and find the flow rates for this rather complex layup.

### 3.3.2 Changing the permeability

In all previous setups, the N10 breather was used as a porous media for air flow. When a setup is created with only one layer of breather under the vacuum bag, an area of uniform permeability is created. However, in industry it is rare to see a setup with uniform permeability. Instead, there are different layers with various permeabilities that all get sealed in the vacuum bag. For example, there can be different layers of reinforcement material, as well as complex shapes that contribute to a non-uniform media. The simulation used should be able to reliably predict leakage locations despite non-uniform permeability as well as other conditions.

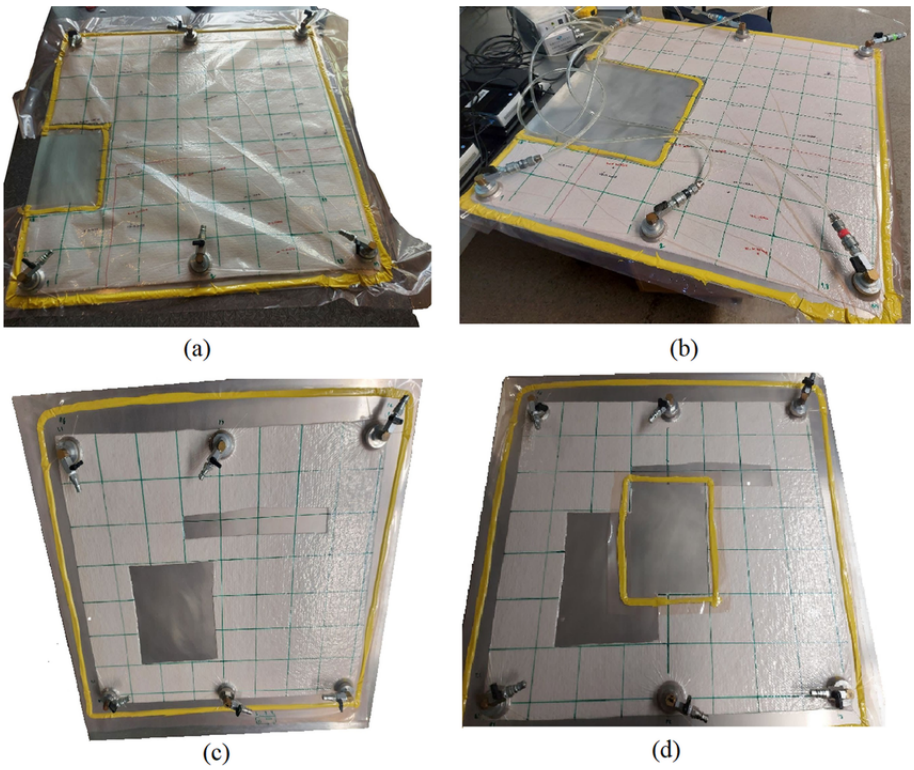


Figure 3.10: Designed layups with deducted breather to represent regions with low permeability

Another method of introducing new set of experiments involves modifying the breather permeability by selectively removing certain sections to prevent air from traversing through these specific regions, as depicted in Figure 3.10. In these deducted regions, the vacuum bag will attach to the mold, representing locations that very low permeability exists in the layup. The simulation process must also be able to compute flowrates in these scenarios. There are four different sets of experiments in this category.

Previous experiments have aimed to reduce or eliminate permeability in specific regions. However, in industrial applications, certain areas may exhibit relatively higher permeability. To address this, technicians often use multiple layers of breather material to aid in the removal of excess air and resin. Additionally, variations in the thickness of the layup material can result in non-uniform permeability. To enhance permeability in certain regions, multiple layers of breather can be incorporated into the layup setup. As illustrated in Figure 3.11, the left-hand image shows an extra breather layer positioned diagonally within the system. In the right-hand image, approximately one-third of the layup is covered with an additional breather layer. As mentioned earlier, the breather serves as a porous medium through which air and resin can pass. By increasing the number of breathers, more porous media is introduced, thereby reducing viscous resistance and increasing permeability.

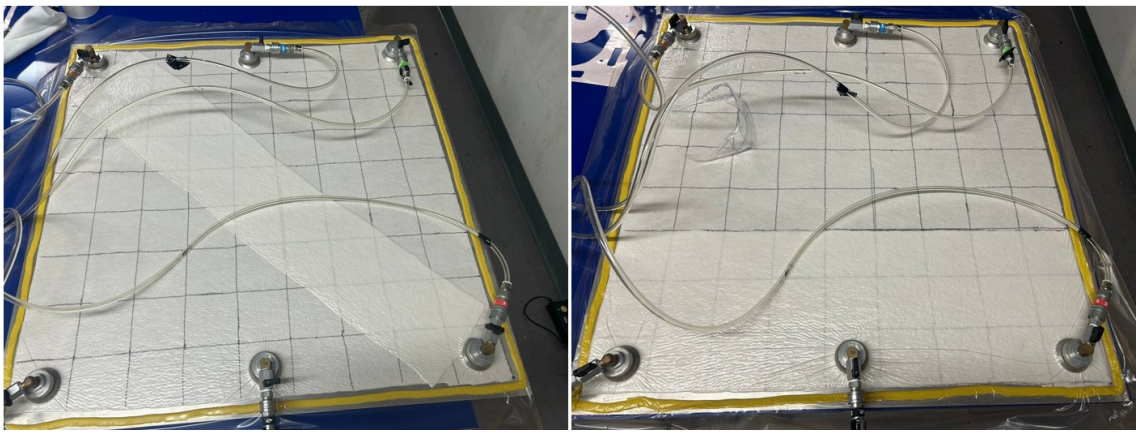

Figure 3.11: Designed layups with extra breather to represent regions with high permeability

In order to create areas with extremely high permeability, angle iron sections could be introduced inside the vacuum bag in order to create a channel for air flow. These channels offer a pathway with lower resistance compared to the original breather

setup. Figure 3.12 shows the 3/4" angle iron profiles used for creating these high-permeable regions.

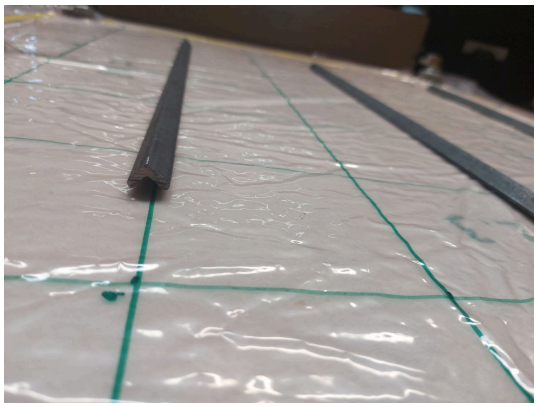

Figure 3.12: Angle irons used for creating non-uniform permeability in the setup

Using multiple layers of breather increases permeability in all directions. However, utilizing angle irons enhances permeability primarily in one direction, making them suitable for specific layups. For example, the configuration shown in Figure 3.13 facilitates airflow more efficiently between the right and left sides of the bag. Three angle irons have been placed horizontally beneath the vacuum bag.

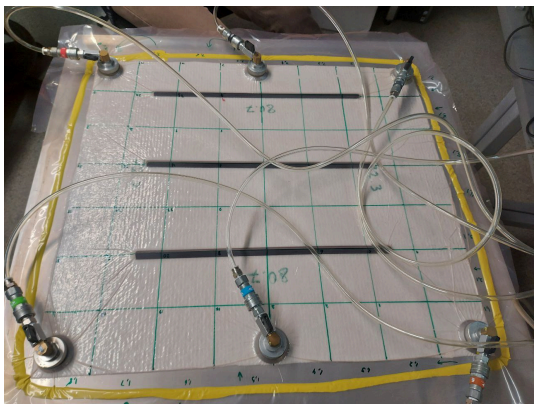

Figure 3.13: setup used for creating horizontal non-uniform permeability in the vacuum bag

Additional setups incorporating angle irons were designed and tested. In the configurations shown in Figure 3.14, angle irons were placed vertically along the y-axis. The air channels created by the angle irons reduce flow resistance in their direction, allowing for airflow between the top and bottom ports. In the setup on the

left-hand side, five angle irons were used to create multiple channels, facilitating easy airflow in the vertical direction.

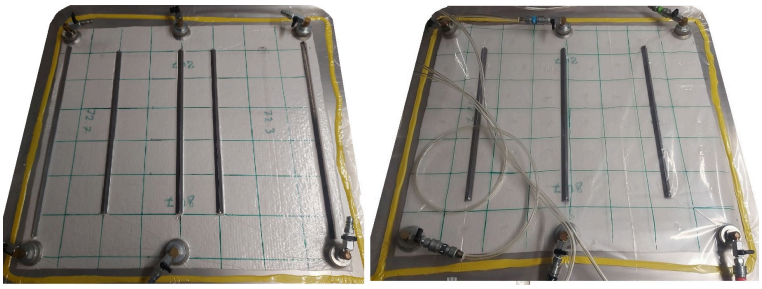

Figure 3.14: Setup used for creating vertical non-uniform permeability in the vacuum bag

This section briefly outlines various methods and procedures for introducing non-uniform permeability. In each setup, a grid was drawn on the layup, dividing it into small squares of approximately 10 cm each. During the experimental process, samples were taken from the center of each square. Leakages were introduced using a standard needle with a diameter of 0.5 mm, and flow rate values were recorded. After each measurement, the leakages were sealed, ensuring no other leaks were present in the vacuum bag before the next sampling. In all these experiments, the size of the leakage was controlled and kept constant. However, in real-world scenarios, the size of the leakages cannot be controlled and is considered an uncontrollable factor. The next section will discuss methodologies for studying leakage size variability.

## 3.4 Effect of the Leakage Size

In industry, the size of pre-existing leakages is an uncontrollable factor. The leaks can range in size from as small as 0.1 mm in diameter to big holes and leaks in the sealant tape on the boundaries. However, the leakage localization method must be robust against variations in leak size. Generally, larger leaks will increase the volume of air entering into the vacuum bag and subsequently the outbound flowrates at steady-state. In addition to the leakage size, the amount of penetration of the needle into the vacuum bag into the subsequent layers such as breather also affects the flowrate absolute values. Since leak size is an uncontrollable factor, a prediction method which relies on the normalized flowrate values is ideal. Using the normalized flowrate values means that any sized leak at a specific location should yield the same

location prediction.

We also refer to normalized flow rate values as port contribution. Previously in [138], we proposed considering the contribution of each vacuum port instead of the absolute values in order to suppress the effects of leakage size variance. Contribution of each port is defined by Eq. 3.1.

$$f_i = F_i/(\sum_{i=1}^{n} F_i) \tag{3.1}$$

In which, $F_i$ is the flowrate of port $i$ that has been calculated by averaging the flowrates in the steady-state time. The number of vacuum ports is shown with $n$ and in our study case it is equal to six. As an example, absolute and normalized flowrates for the sample experiment shown in Figure 3.6 is presented in Table 3.1. The flow rates are averaged in between 100 and 200 seconds time interval.

Table 3.1: Absolute and normalized flow rate values of the example in Figure 3.6

| Vacuum port number (i) | 1 | 2 | 3 | 4 | 5 | 6 | Total |
|---|---|---|---|---|---|---|---|
| Absolute flow rate value in liter per minute: $F_i$ | 0.037 | 0.119 | 0.122 | 0.310 | 0.165 | 0.042 | 0.795 |
| Normalized flow rate: $x_i$ | 0.047 | 0.150 | 0.153 | 0.390 | 0.208 | 0.053 | 1 |

As previously discussed, port number four is the closest to the leakage and since the setup is symmetrical with uniform permeability, the port has the maximum flow rate compared to other vacuum ports. The absolute flow rate value of this port is 0.310 $L/min$ but when normalized, it is contributing to 39 percent of the total flow in the vacuum line.

Absolute flow rate values vary with leakage size, and their time series often exhibit oscillations or transient behavior, making them unreliable inputs for training machine learning models. This raises the question: Are normalized flow rate values reliable, and do they remain consistent despite variations in leakage size? The debate over using port contributions or normalized flow rates can be informed by the nonlinear effects of leakage size on flow rate values. What if changes in leakage size disrupt the contribution of each port? To address this, well-designed experiments are necessary. As a starting point, four needles with diameters of 0.5, 0.6, 0.7, and 0.8 mm have been selected to generate leakages of varying sizes. Figure 3.15 illustrates the flow rates for a specific leakage, created using different needle sizes. It is evident that the total volumetric flow through each port increases as the needle size increases. There

is a small decrease in flow rates before the jump in their value because of the bigger needle. That is the moment when the new and bigger needle was used to increase the leakage size and for a moment blocked incoming air into the vacuum bag and resulted in a short decrease in the flow rate outputs measured in vacuum ports.

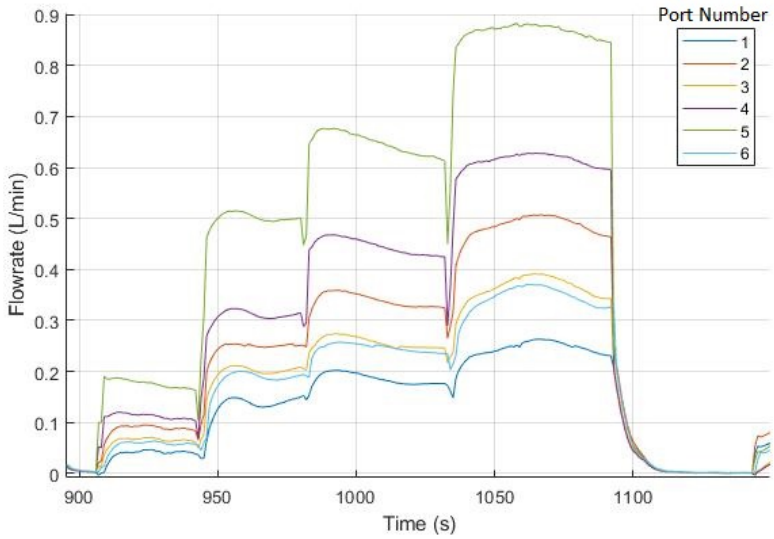


Figure 3.15: Measured flowrates when using four different needle sizes

The question we are looking to answer is that does the normalized flow rate significantly change when having different leakage sized. To properly answer this question, the experiments have been repeated with the four sizes of needles at sixteen distinct locations on the vacuum bag, constituting a total of 64 data points. The objective of this experiment is to determine the effect that different leakage sizes have on flow rate contributions. This question was addressed through a two-way ANOVA (analysis of variance) analysis. The two factors in this analysis are:

1. Variance in leakage size (The different needle sizes)

2. the remaining factors that have been blocked together (Leakage location, permeability, vacuum port locations, geometry)

Before starting the analysis, it is clear that the second factor, block of remaining factors (such as port number and leakage location), have significant effect on the flowrates. However, in this ANOVA, these effects have been grouped together as a block in order to focus solely on the significance of different leakage sizes. The null-hypothesis for this analysis is that the source of variation (needle size or the block)

has no significant effect on the port contributions. Table 3.2 shows the t-test results for the two-way ANOVA analysis.

Table 3.2: Two-factor ANOVA for the leakage size experiment

| Source of Variation | $SS$ | $df$ | $MS$ | $F_{stat}$ | $P_{value}$ | $F_{cr}$ |
|---|---|---|---|---|---|---|
| Leakage size | 2.2 E-15 | 3 | 7.4 E-16 | 2.2 E-11 | 1 | 2.64 |
| Block of remaining factors | 3.2463 | 95 | 0.034 | 1030.95 | 0 | 1.30 |
| Error | 0.009 | 285 | 3.3 E-05 | | | |
| Total | 3.26 | 383 | | | | |

As presented in Table 3.2, for the leakage size factor, the $F_{cr}$ is orders of magnitude larger than the $F_{stat}$ and the P-value (probability) of the null-hypothesis to be true is equal to 1. Therefore, it can be concluded that, with a confidence of 100 percent, different leakage sizes have no significant effect on the flowrates. Om the other hand, the P-value for the blocks (different ports and locations) is equal to zero which means the null-hypothesis is rejected and the blocks have a significant effect on the flowrates. In conclusion, this study shows that using normalized flow rate values are reliable parameters for training a machine learning model. Any penetration tool with any size could be used for leakage creation as long as it is not excessively big to create turbulent flows or flows that exceed our flow rate sensors upper limit which is 1 L/min.

## 3.5 Simulation Tool for Data Synthesis

As previously discussed, a reliable, validated, generalizable and rapid simulation tool is required in this process to be used for data generation. Even though a valid lab-scale experimental setup is available and the methodology for experiments and cleaning the data independent of the leakage size has been designed, it is still not feasible to generate sufficient data using experiments. In addition, the challenge of addressing the problem when multiple leakages are present requires an enormous amount of data and the problem gets more complicated.

Before proposing any simulation method, it is important to clarify the expected outcomes. To do so, we must revisit the experimental procedure outlined in Section 3.2. The simulation should adhere to the constraints, properties, and results of the experiments. Starting with simple configurations, the simulation must eventually handle varying configurations, including different sizes and geometries, non-uniform permeability introduced in Section 3.3, and variations in vacuum port locations.

Additionally, the simulation method must be relatively fast, ideally faster than conducting the experiments and processing the data. In the experiments, the transient response during the creation or repair of a leak, or the flow rate behavior before reaching steady-state, holds minimal significance. This is because leaks generally exist in the layup before the vacuum is applied, and once the leak location is identified, the issue is resolved. Consequently, there is no need to simulate every detail—especially the transient response—as this would unnecessarily increase complexity, computation time, and cost. It is more efficient for the simulation to directly calculate the flow rates during the steady-state interval. In summary, a proper simulation must be:

1. Reliable, valid and verified with experimental data at least in the steady-state phase.

2. Rapid and be able to generate data faster than the experiments.

3. Able to simulate cases with non-uniform permeability and various sizes.

4. Able to simulate different location of leakages, such as leakages on the vacuum bag and on the tacky tape on the boundary.

5. Able to address having multiple leakages on the layup.

Simulation of the process enables the determination of flow rate values based on leakage locations. By comparing and verifying experimental results with simulations, valuable insights can be gained.

### 3.5.1 Simulation using finite element method

The initial and basic method that comes to mind is the Finite Element Method (FEM). Components of the system, such as the vacuum bag, air inside and outside the bag, and breather material, can be discretized into meshes for simulation. By assuming a valid simulation, prerequisites such as various sizes and leakage locations can be satisfied. However, traditional numerical methods, like those implemented in software such as Flow3D, COMSOL, and Ansys, are often impractical due to their extensive computational demands, requiring hours or even days for a single configuration on regular computers.

Moreover, these methods introduce unnecessary complexities and non-linearity irrelevant to the specific analysis. Transient response often cannot be neglected,

as the equilibrium volume and pressure of the air inside the flexible vacuum bag are unknown, necessitating a full transient solution from initial to equilibrium conditions. Additionally, the vacuum bag's flexibility introduces further complexity, as wrinkles, contractions, and expansions may occur. Figure 3.16 shows the setup modeled in Flow 3D. The simplified setup is 50×50×20 $cm^3$ in dimensions, with a single vacuum port of 2 cm diameter and no leakages to illustrate the results of a leakage-free simulation.

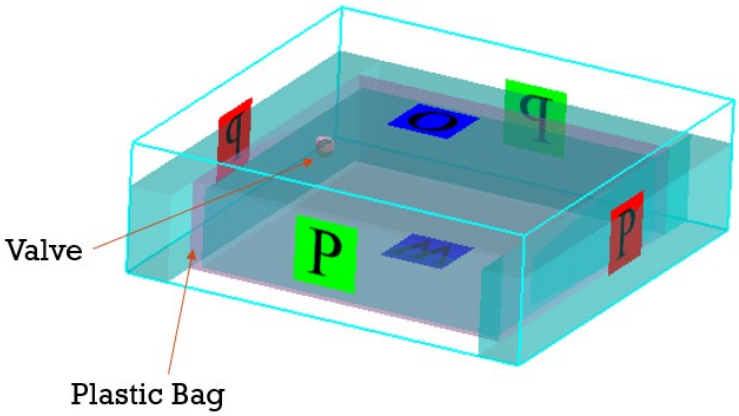


Figure 3.16: Simple simulation setup using Flow 3D Software

In this simulation, the ambient air, with atmospheric boundary conditions, was modeled surrounding a vacuum bag. A low pressure of 1000 Pa was applied at the valve. Despite extensive computational efforts, the results for this simplified setup could not be validated against the experimental data. It is important to note that the actual experimental setup is far more complex, with larger dimensions, six vacuum ports, and leakages that must be accounted for in the simulation.

Additionally, simulations using other FEM-based software were conducted. A model with dimensions of 10×10×2 $mm^3$ incorporating a vacuum valve, was developed using COMSOL Multiphysics. A low-pressure boundary condition was applied to the vacuum bag plane, with ambient pressure conditions elsewhere. Figures 3.17 and 3.18 depict the pressure contours and cross-sectional air velocities, respectively. The cylindrical region represents the vacuum, where low pressure was applied as a boundary condition. While the pressure contours appeared reasonable, the velocity values in the cross-sectional area of the vacuum port were excessively high, resulting in flow rate values that were orders of magnitude larger than the experimental flow rate, which was below 1 L/min.

All these basic and simplified FEM simulations demonstrate how using this method

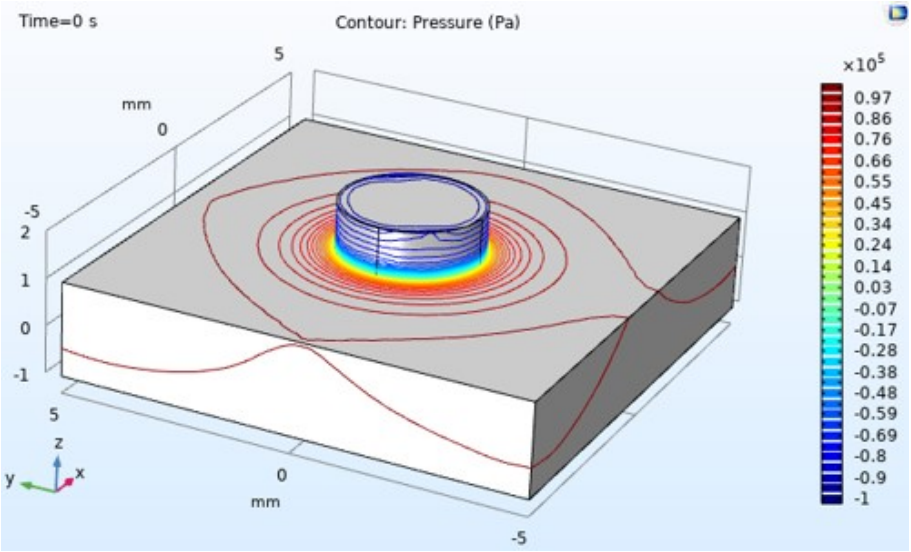


Figure 3.17: Results for pressure contours for a small-scale setup using COMSOL Multi physics Software

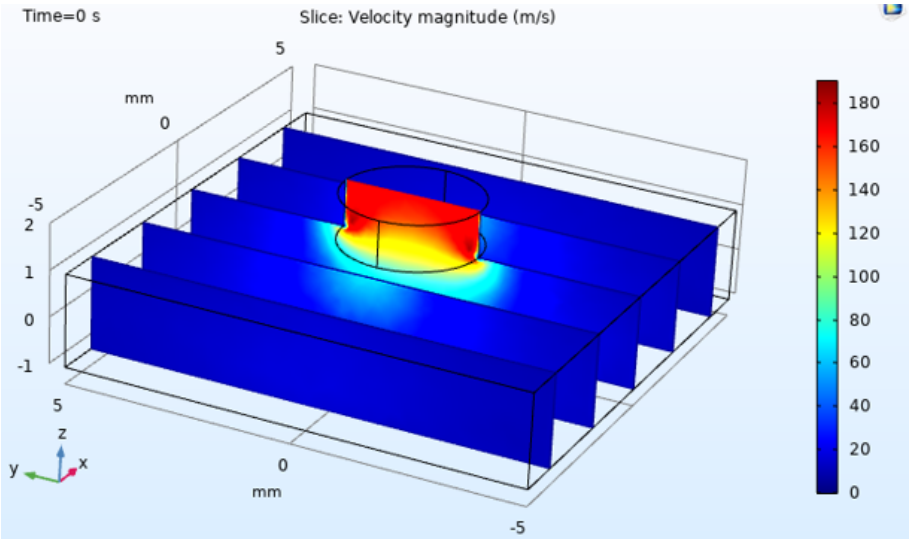


Figure 3.18: Results for velocity in various sections for a small-scale setup using COMSOL Multi physics Software

can be both complex and time-consuming. Additionally, the generalizability of these methods is low, meaning that any changes to the setup or leakage locations require redoing the entire simulation from scratch. Moreover, these methods often take more time to execute than conducting experiments, even in this simplified form. Therefore, alternative methods have been proposed and utilized in this work to better meet the conditions outlined for proper simulation in Section 3.5.

### 3.5.2 Simulation using analogy

To address the aforementioned challenges, a novel approach inspired by control theory suggests establishing an analogy between the vacuum bagging experiment and an electrical circuit. Electrical and mechanical components often exhibit similar effects on systems. For example, a second-order mass-spring-damper mechanical vibration system can be analogized to a resistor-inductor-capacitor (RLC) circuit shown in Figure 3.19.

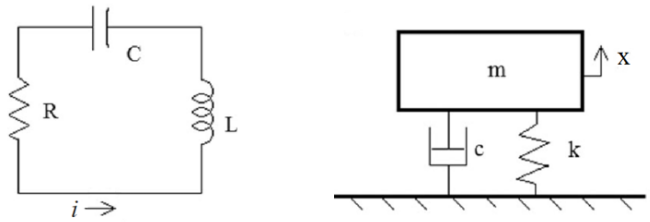


Figure 3.19: Well-known analogy in mechanical and electrical systems

Differential equations 3.2 and 3.3 describe the behavior of displacement ($x$) in the mechanical system and current ($i$) in the electrical systems, respectively. The responses of both systems are analogous, and either system could be used interchangeably if properly normalized. Leveraging this similarity, the flow rates of vacuum ports can be modeled using the current in an electrical circuit.

$$m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = 0 \tag{3.2}$$

$$L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{1}{C}i = 0 \tag{3.3}$$

Darcy's law is the governing formula in the air flow underneath the vacuum bag shown in Eq. 3.4. Based on Darcy's law, the flux ($q$) decreases as the distance between and inlet and outlet ($L$) increases [148].

$$q = -\frac{k}{\mu L}\Delta p \tag{3.4}$$

The flux in this formula, when multiplied by the area, corresponds to the flow rate values measured by COHO sensors. Since we are using vacuum connectors with equal cross-sectional areas, as shown in Figure 3.2, the area remains consistent. Additionally, $k$ represents the permeability of the breather material, and $\mu$ is the dynamic viscosity of air at a given temperature. We assume constant values for $k$ and $\mu$, as we are dealing with a uniform setup in steady state.

Darcy's law in this form is applicable for one-dimensional problems and would hold if there were a direct pipe connecting the leakage to each vacuum port. The pressure difference across the vacuum ports will be uniform, as the pressure at the leakage point is atmospheric, while the vacuum ports are near vacuum pressure due to their connection to the same pump.

However, in our setup, the breather introduces a two-dimensional resistance, and instead of a pipe, the breather area connects the leakage to each vacuum port. This means that, in addition to the direct path between the leakage and the vacuum ports, air can take parallel paths to reach the ports. Using Darcy's law in its 1D form essentially calculates flow rates based on the inverse of the leakage-to-port distance, which provides a reasonable approximation. However, this method is not entirely accurate due to the presence of parallel flow paths. This is analogous to electrical circuits, where calculating the equivalent resistance between two nodes requires accounting for all parallel resistances.

To begin with, it is advisable to employ the simplest model possible, which is the zero-order model comprising solely resistors. A corresponding component mapping between the two systems is provided in Table 3.3. In the vacuum bagging experiment, a pressure difference between two locations prompts the flow from the higher-pressure area to the lower-pressure area. Similarly, a voltage difference between two nodes in an electrical circuit induces a current across the nodes. In both systems, the resistance governs the magnitude of the flow or current, which can be quantified using sensors.

Table 3.3: Analogous terms between physical system and the simulated model

| Physical system | Unit | Simulation model | Unit |
|---|---|---|---|
| Flowrate (COHO sensors) | $L/min$ | Current (Ampere meter) | A |
| Pressure difference | $Pa$ | Voltage difference | V |
| Breather viscous resistance | $N.s/m^2$ | Electrical Resistance | $\Omega$ |

### 3.5.3 Modelling the breather

To simulate the vacuum bagging experiment, an electrical resistor network is devised to match the behavior of the breather. The parallel connection of two resistors in the network results in a decrease in the overall resistance, analogous to the presence of parallel airflow channels. By increasing the number of resistors in the network, the simulation's resolution and precision can be enhanced, akin to using finer meshes in other numerical simulation methods.

Within the network, a specific resistance is assigned to each pair of connected points, and this resistance increases as the distance between the nodes grows. This mirrors the scenario within the vacuum bag, where ports closer to the hole exhibit higher flow rates. Initially, each resistor is arbitrarily set to a value of 100Ω, as the network resistance is subsequently normalized to become dimensionless.

Figure 3.20 shows the network of electrical resistors to model the viscous resistance of the breather used in the experimental setup previously shows in Figure 3.3. There is a 9 by 9 grid on the setup and similar configuration has been designed on the model. Since the experimental setup is uniform, the electrical resistor values are constant in the network.

### 3.5.4 Modelling the vacuum ports

Next step is modelling the vacuum ports in the setup. The vacuum ports in the experiment operate at a significantly lower pressure than atmospheric pressure which is around 20 KPa absolute pressure within our system. This lowest pressure level in the entire system can be regarded as equivalent to the ground reference in an electrical network. Accordingly, in our electrical network model, we establish grounding at six locations to replicate the configuration of the COHO system's six ports.

To enable the measurement and recording of current passing through each grounding location, virtual ammeters are introduced before each grounding point. These virtual ammeters allow us to extract and monitor the current flow at each specific location within the network. In the physical system, the COHO sensors measure flow rates on the evacuation lines. Sensors and pipes will have a resistance for the out bounding air, but have been neglected compared to the breather viscous resistance. This internal resistance could be modelled by assigning an internal resistance to the ammeters, but have been neglected similarly.

Figure 3.21 shows the modelling of vacuum ports and COHO sensors using elec-

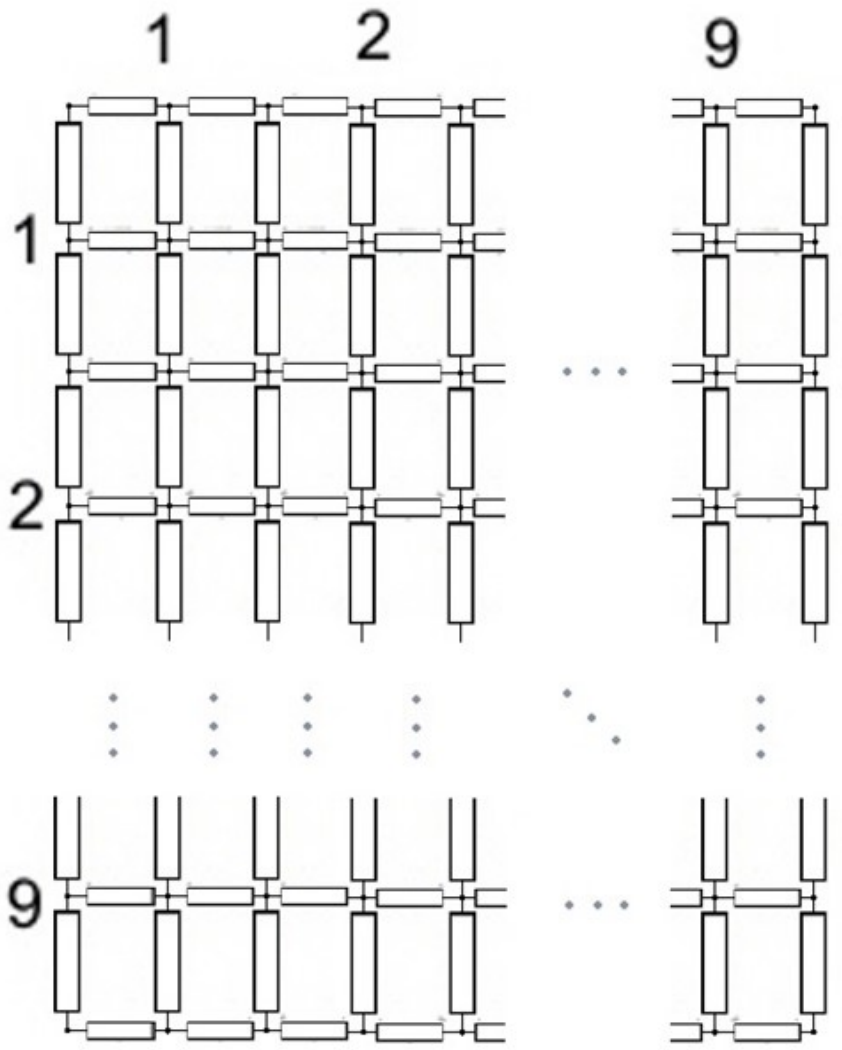


Figure 3.20: Network of electrical resistors used for modelling the breather in the experimental layup

trical grounds and ammeters on the current's pathway. The locations of electrical grounds will be on corresponding location of the vacuum ports. therefore, the design has the flexibility to change the number and location of the vacuum ports, promptly.

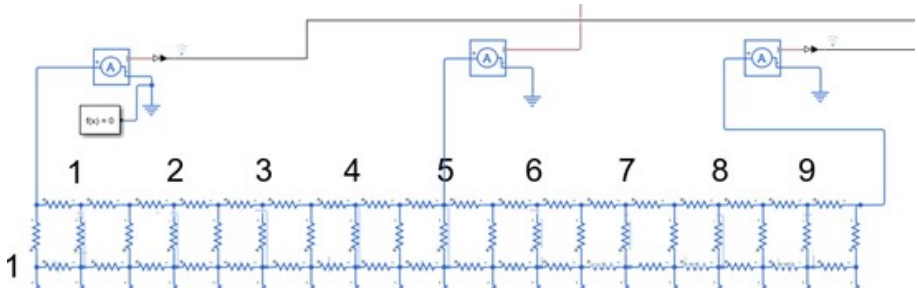


Figure 3.21: modelling of vacuum ports and COHO sensors in the electrical circuit; electrical ground with low voltage acts as the location with lower pressure and ammeters measure current, the equivalent of measuring flow rates by sensors

### 3.5.5 Modelling the leakages

In the mechanical system, the presence of a leakage allows atmospheric air, at a higher pressure, to enter the vacuum bag. This air flows from the leakage points towards the vacuum ports, which maintain a lower pressure. Analogously, in an electrical system, the hole in the vacuum bag can be represented as a node with a positive voltage source. In this analogy, current flows from the positively charged node (representing the hole) towards the six ground nodes (representing the six vacuum ports).

To simulate this behavior, each leakage point in the mechanical system is mapped to a corresponding node in the electrical system, which is assigned a positive voltage. This allows for the calculation of the electrical currents flowing through the six distinct ground channels. MATLAB's Simulink has been used to simulate the electrical network.

With Simulink, the electrical circuit's behavior can be modeled, enabling the analysis and calculation of current flow through the system's components and channels.

### 3.5.6 Comparison method

Directly comparing the outputs of the experiment and simulation is not feasible due to dimensional incompatibility. To objectively compare the two systems, it is necessary to first convert the results into dimensionless quantities. Utilizing dimensionless

outputs enables not only the comparison between the systems but also eliminates the need to specify parameters of the electrical network, such as resistances and voltages. Moreover, by employing dimensionless analysis, the flow rate evaluation becomes independent of the system's dimensions and the size of the hole.

In the experiment, the crucial value lies in the normalized or relative flow rates. The absolute flow rate value of a port provides limited information about the distance between the leak and the port. Therefore, the focus should be on the flow rate of each port relative to the others, as it delivers valuable insights into the comparison between different ports and their respective leakage distances. Therefore, we will compare normalized flowrate values of the experiments with the normalized values of currents acquired from the simulated model. The normalized flow rate values ($f_i$) could be calculated using Eq. 3.1. Normalized current could be calculated similarly using Eq. 3.5.

$$\bar{I}_i = I_i / (\sum_{I=1}^{n} I_i) \tag{3.5}$$

The Root-Mean-Square errors will be calculated to discover model error on each sample using Eq. 3.6.

$$error(node) = RMS_1^n(f_i - \bar{I}_i) = \sqrt{\frac{\sum_1^n (f_i - \bar{I}_i)^2}{n}} \tag{3.6}$$

To assess the efficacy of this analogy, tests could be conducted using different setups and compared the simulated data with corresponding experimental data. With a reliable a proper simulation tool, artificially synthesized data can be employed as training sets for machine learning systems. Furthermore, experimental data can be used for cross-validation or testing purposes once the machine learning system has been trained. In the next section, the problem will be addressed for multiple leakage scenarios.

## 3.6 Multiple leakage scenarios

Leakage detection becomes exceedingly complex when an unknown number of leaks with unspecified locations exist in the system. This challenge must be considered when training the system. Therefore, it is crucial that the proposed simulation performs accurately under conditions of simultaneous leakages. Predicting the locations

of multiple leaks is far more complicated than single-leak scenarios. Currently, no state-of-the-art methods based on flow rate measurements effectively address this issue, with most deferring it to future research.

Before locating multiple leakages on the vacuum bag, the first step is to determine how many leaks exist. The first step in addressing multiple leakage scenarios is determining the number of leaks present in the system. This can be framed as a classification problem, where the task is to classify examples into one of two categories: systems with either one or two leaks. Our objective in this research stops at having one or two leakages in the setup and cases with three or more leakages are not a part of our studies due to further complexities.

### 3.6.1 Feature recognition in multiple leakage scenarios

To be able to successfully classify the examples, is to understand feature differences and variances between two classes. To understand if flow rates measured from having one leakage in the setup have a difference with flow rates measured when having two leakages on the layup. This is where training a machine learning model would be possible. To start, two different scenarios shown in Figure 3.22 have benn considered.

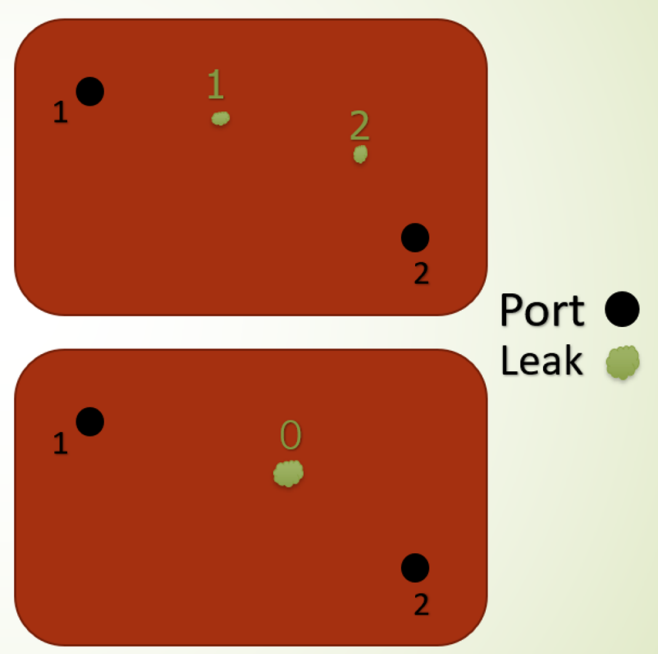


Figure 3.22: Setup with having two (top) and one (bottom) leakage for flow rate comparison case study; all elements including setup size and vacuum ports remain the same in both cases

In both of the case studies, the layup and vacuum port locations are identical. However, in the first case shown in top, two leakages with the same size exist in the setup. In the second scenario shown in bottom, one leakage exists in the setup. The objective is to estimate and compare normalized flow rates generated by these two case studies. For simplification purposes, only two vacuum ports exist in the setup and their distance is a constant value of $\vec{R}$. Note that the leakage size of both leakages in case (1) is similar and we are studying multiple leakages with same leakage size. However, there is no constraint on the leakage size of the leakage in the case (2). Each leakage, generates a flow on each vacuum port and this flow increases with leakage size and decreases with the viscous resistance between the leakage and the vacuum port. The volumetric flow rate or absolute flow rate in each vacuum port could be calculated with the following equation.

$$F_j = \sum_i \frac{c_i}{g(r_{ij})} \tag{3.7}$$

In the equation 3.7, the term $F_j$ is the volumetric flowrate flowrate of vacuum port $j$ which could be calculated by adding the effects of all leakages that is shown by $i$. $i$ is either 1 or 2 in our case study. $c_i$ is the leakage size factor and a constant that depends on the size of the leakage and vacuum pump's efficiency and power. $r_{ij}$ is the distance between leakage number $i$ to vacuum port number $j$. The term $g(r_{ij})$ represents the viscous resistance existing between the leakage and the vacuum port. This nonlinear function depends on the distance, the permeability structure of the layup, the location of the vacuum ports and surrounding area and the amount of region and paths the air could travel from the leakage with the high pressure to the port with the low pressure. This term, however nonlinear, could be estimated and written with its Taylor Series as a polynomial function shown in Eq. 3.8.

$$g(r) = a_0 + a_1 r + a_2 r^2 + ... \tag{3.8}$$

The term $a_0$ depends on vacuum port and connecting pipes' internal resistance and other terms depend on the permeability and settings. By using these equations, volumetric flow rates of vacuum ports for the case 1 could be written as:

Case 1:

$$F_1^{(1)} = \frac{c}{g(r_{11})} + \frac{c}{g(r_{21})} \tag{3.9}$$

$$F_2^{(1)} = \frac{c}{g(r_{12})} + \frac{c}{g(r_{22})} \tag{3.10}$$

$R$ is the distance between vacuum ports and therefore Eq. 3.10 could be rewritten as:

$$F_2^{(1)} = \frac{c}{g(R - r_{11})} + \frac{c}{g(R - r_{21})} \tag{3.11}$$

There is only one leakage existing in case 2, and a different size factor exist for this scenario, therefore the flow rates could be written as:

$$F_1^{(2)} = \frac{d}{g(r_{01})} \tag{3.12}$$

$$F_2^{(2)} = \frac{d}{g(r_{02})} = \frac{d}{g(R - r_{01})} \tag{3.13}$$

Since there are two vacuum ports in this case study, normalized flow rates could be calculated as following:

$$f_1^{(1)} = \frac{F_1^{(1)}}{F_1^{(1)} + F_2^{(1)}} = \frac{\frac{1}{g(r_{11})} + \frac{1}{g(r_{21})}}{\frac{1}{g(r_{11})} + \frac{1}{g(r_{21})} + \frac{1}{g(r_{12})} + \frac{1}{g(r_{22})}} \tag{3.14}$$

For case 2, the normalized flowrate is equal to:

$$f_1^{(2)} = \frac{F_1^{(2)}}{F_1^{(2)} + F_2^{(2)}} = \frac{\frac{1}{g(r_{01})}}{\frac{1}{g(r_{01})} + \frac{1}{g(r_{02})}} = \frac{1}{1 + \frac{g(r_{01})}{g(r_{02})}} \tag{3.15}$$

We are looking for the leakage locations when both scenarios lead to a same result in the terms of normalized flow rates. The reason is to study the cases were both cases could generate equal flow rates and if they exist, it will be more difficult to train a machine learning model on that base.

$$f_1^{(1)} = f_1^{(2)} \tag{3.16}$$

$$f_2^{(1)} = f_2^{(2)} \tag{3.17}$$

By substituting Equations 3.14 and 3.15 into the Eq. 3.16, for vacuum port number one, results in:

$$\frac{\frac{1}{g(r_{11})}+\frac{1}{g(r_{21})}}{\frac{1}{g(r_{11})}+\frac{1}{g(r_{21})}+\frac{1}{g(r_{12})}+\frac{1}{g(r_{22})}}=\frac{1}{1+\frac{g(r_{01})}{g(r_{02})}} \tag{3.18}$$

Eq. 3.18 could be written as:

$$\frac{1}{1+\frac{\frac{1}{g(r_{12})}+\frac{1}{g(r_{22})}}{\frac{1}{g(r_{11})}+\frac{1}{g(r_{21})}}}=\frac{1}{1+\frac{g(r_{01})}{g(r_{02})}} \tag{3.19}$$

Therefore:

$$\frac{\frac{1}{g(r_{12})}+\frac{1}{g(r_{22})}}{\frac{1}{g(r_{11})}+\frac{1}{g(r_{21})}}=\frac{g(r_{01})}{g(r_{02})} \tag{3.20}$$

By substituting $R$, it succumbs:

$$\frac{\frac{1}{g(|\vec{R}-\vec{r_{11}}|)}+\frac{1}{g(|\vec{R}-\vec{r_{21}}|)}}{\frac{1}{g(r_{11})}+\frac{1}{g(r_{21})}}=\frac{g(r_{01})}{g(|\vec{R}-\vec{r_{01}}|)} \tag{3.21}$$

The term $g(r_{11})$ could never be equal to $g(|\vec{R}-\vec{r_{11}}|)$, based on the form showed in Eq. 3.8, unless the distance of the leaks is equal to both vacuum ports, which is not valid in general. Therefore, the only condition that Eq. 3.21 would be valid, is satisfying both of the following equations:

$$g(r_{11})=g(r_{21}) \tag{3.22}$$

$$g(|\vec{R}-\vec{r_{11}}|)=g(|\vec{R}-\vec{r_{21}}|) \tag{3.23}$$

Due to the polynomial form of $g(r)$ shown in Eq. 3.8, Equations 3.22 and 3.23 are only valid if and only if:

$$\vec{r_{11}}=\vec{r_{21}}=\vec{r_{01}} \tag{3.24}$$

This means that the vector connecting leakage number one to the vacuum port number one, should be equal to the vector connecting leakage number two to the vacuum port number one. By substituting $R$, same expression is valid for vacuum port number two as well:

$$\vec{r_{12}}=\vec{r_{22}}=\vec{r_{02}} \tag{3.25}$$

The results mean that the leakage number one should be in the exact location of leakage number two so that both cases shown in Figure 3.22 result in same normalized flow rate values. The findings could be applied to any number of vacuum ports and therefore applicable to our case study.

There are two key findings from these calculations that must be considered in the classification process.

1. In a given setup, the normalized flow rates generated by two simultaneous leakages differ from those generated by a single leakage. This distinction makes it feasible to train a classification model to identify the number of leakages.

2. As the distance between two leakages decreases, the flow rates become more similar, resembling the pattern of a single leakage between them. This can cause the classification model to suffer from overfitting and high variance errors if the training set contains leakages that are too close together. To ensure effective training, the training and test datasets should enforce a minimum distance between leakages. If leakages are too close, the algorithm may ultimately fail to differentiate between them and classify them as a single leakage.

In addition, there are specific assumption that the leakages in the multiple leakage scenario should have same size and we are studying the cases when we have one or two leakages in the setup.

### 3.6.2 Simulation method for multiple leakage scenarios

Data is arguably the most important and costly element of any machine learning model, and having a proper simulation that enables data synthesis is crucial. In Section 3.5.2, we presented a simulation based on the analogy between electrical circuits and our physical model. Although this approach was initially used for single-leakage scenarios, it shows promise for application in multiple leakage cases as well. Generalization from single to multiple leakages is feasible if both the experiments and the simulation adhere to the superposition rule. This rule states that the effects of two different inputs can be summed without interference.

From a simulation perspective, the superposition rule holds in linear electrical circuits [149]. Our proposed simulation model consists only of linear components such as resistors, capacitors, and inductors. As long as nonlinear components like

diodes and transistors are avoided, we can add the effects of two distinct high voltages (representing leaks), compute their respective currents, and sum them accordingly.

Experimentally, however, the flows generated by two simultaneous leakages may interfere with each other, potentially disrupting linearity and violating the principle of superposition. The general formula for estimating volumetric flow rate, as shown in Eq. 3.7, assumes non-interfering flows from multiple leakages. The flow behavior—whether laminar or turbulent—can be characterized by a constant known as the leak rate, which can be defined as in Eq. 3.26 [116].

$$q_{pV} = F\frac{\Delta p}{\Delta t} = const. \tag{3.26}$$

in Which $F$ is the volumetric flow rate, $\Delta p$ change of pressure during a certain time interval. larger leak rates bigger than $10^{-2}$pa $l/s$ [150] [151] indicate practically complete laminar flow whereas small leak rates lower than $10^{-5}$pa $l/s$ suggest entirely molecular flow and these conditions get more complicated in the transition region. In all the experiments conducted in this research, the volumetric flow rates range from 0.02 to 1 $l/s$. Considering a pressure drop of 85 $kPa$, the corresponding leak rates are estimated to be between 28 to 1400 pa $l/s$, remaining within the laminar flow region. As a result, it is expected that multiple leakages will not induce turbulence in the flows, and the superposition rule should hold.

To put this theory on test, experiments and simulations have been done and compared with each other. The 81cm×72cm setup shown in Figure 3.8 has been used for the experiments. In this case, two simultaneous leakages are created using a particular needle and flow rate values are recorded. On the simulation side, high voltage will be applied to the two corresponding locations on the electrical circuit network. Currents in six ammeters will be recorded and normalized and two systems will be compared with each other.

## 3.7 Machine learning model for single leakage localization

Once the proposed analogy has been verified across various configurations, it can serve as a valuable tool for data generation. After synthesizing a sufficient amount of unbiased data, a regression model can be trained. The primary objective of this trained model is to accurately predict leak locations based on the normalized flow

rates read by the COHO system.

To prevent overfitting during training, it is crucial to split the dataset into two parts: one for training and one for validation. Typically, 80% of the data is used for training, while the remaining 20% is reserved for 5-fold cross-validation. This cross-validation method helps mitigate overfitting. Additionally, to assess the model's generalizability, a separate test dataset consisting of new samples the model has never seen before can be used. In the 5-fold cross-validation process, the data is divided into five subgroups. One group is used for hyper-parameter tuning, while the remaining four are combined to form the training set. The model is trained using all five possible validation subgroups, and the model with the highest accuracy is selected as the final output. Figure 3.23 shows the training procedure for the single leakage localization problem.

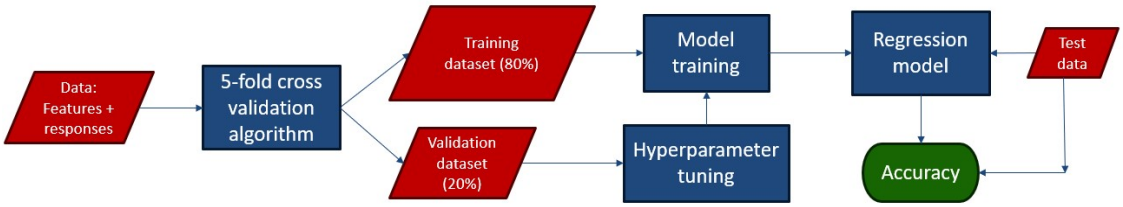


Figure 3.23: Training procedure for single leakage localization problem

Each data entry should include the flow rates of the six vacuum ports (features) as well as the coordinates of the leaks (labels) relative to a designated origin. Different models of linear, regression tree, SVM, GPR and neural network can be trained in order to accurately model the system. In the next section, the data structure will be discussed.

### 3.7.1 Data structure

In any machine learning model, the data is represented by two components: (1) inputs or features, and (2) classes or responses. Since we are working with a regression model, we refer to these components as features and responses. The features are the inputs the machine learning model uses for training, which in our case study are the normalized flow rate values. As discussed in Section 3.4, while the absolute flow rate changes with leakage size, the normalized flow rate values remain consistent. Therefore, using absolute flow rate values would be redundant for training, and normalized flow rates should be employed. This is particularly important because normalized

flow rates are independent of leakage size, making them generalizable to any leakage. In our study, the features consist of six distinct normalized flow rate values.

The responses, on the other hand, are the labels the machine learning model aims to predict. In our case, these are the locations of the leakages. Given a predefined layup, the location can be described using two elements: either in Cartesian coordinates (X-Y) or in polar coordinates (r-$\phi$) with a selected origin. Since the location is described by two components, two separate models must be trained independently. However, the features remain the same for both models. Figure 3.24 shows a portion of the data structure used for training the single leakage localization model. The first six columns represent features that are normalized flow rates and the columns "Row" and "Column" represent the y and x location of the leakages that are responses in the machine learning model. The model uses flow rates from all different possible leakage locations in the structure. This is the data gathered from the 90cm×90cm layup previously shown in Figure 3.3.

| F1 | F2 | F3 | F4 | F5 | F6 | Row | Column |
|---|---|---|---|---|---|---|---|
| 0.048353 | 0.060913 | 0.028002 | 0.036288 | 0.15739 | 0.669052 | 1 | 1 |
| 0.071843 | 0.090343 | 0.045972 | 0.063108 | 0.242917 | 0.485817 | 1 | 2 |
| 0.071734 | 0.099463 | 0.051961 | 0.083368 | 0.372437 | 0.321037 | 1 | 3 |
| 0.06253 | 0.091125 | 0.05268 | 0.098389 | 0.487504 | 0.207772 | 1 | 4 |
| 0.044487 | 0.085046 | 0.062927 | 0.199407 | 0.519334 | 0.088798 | 1 | 6 |
| 0.048104 | 0.093697 | 0.077792 | 0.347981 | 0.352795 | 0.079631 | 1 | 7 |
| 0.038609 | 0.083875 | 0.073043 | 0.507773 | 0.238676 | 0.058025 | 1 | 8 |
| 0.029931 | 0.067901 | 0.06239 | 0.640752 | 0.157011 | 0.042014 | 1 | 9 |
| 0.088273 | 0.10676 | 0.047096 | 0.056667 | 0.217906 | 0.483298 | 2 | 1 |
| 0.097259 | 0.118571 | 0.057883 | 0.074128 | 0.246471 | 0.405689 | 2 | 2 |
| 0.093582 | 0.118358 | 0.066605 | 0.094394 | 0.328437 | 0.298624 | 2 | 3 |
| 0.078754 | 0.118726 | 0.066145 | 0.110989 | 0.418073 | 0.207314 | 2 | 4 |
| 0.070238 | 0.113561 | 0.075205 | 0.142635 | 0.4586 | 0.139761 | 2 | 5 |

Figure 3.24: Data structure used for training single leakage scenarios

Other factors such as location of vacuum ports, permeability in different regions of the setup and geometry and size of the layup could not be integrated in the training. Because these factors remain constant in a particular setup. That will be a drawback of a data-driven machine learning model since it would not take into account any physics into the model.

It is also possible and even mandatory to use synthesized data instead of experimental data. The importance of simulation and scarcity of data was discussed in Section 3.5 and synthesized data must be used for training. With the analogy introduced in Section 3.5.2, normalized currents could be used as features as an alternative

to normalized flow rates and the high voltage locations in the electrical current analogy will be used as responses. The data structure remain the same but normalized currents that are equivalent term for normalized flow rates will be used as features.

In the next section, different machine learning regression models will be discussed.

### 3.7.2 Regression models used for training

Different regression models can be trained to accurately represent the system. These models may range from simple ones with few parameters, such as linear models, to more complex ones with many parameters, like neural networks. Simpler models are less prone to overfitting and exhibit low variance, but they may suffer from high bias, limiting their ability to capture the flow rate-location function accurately. In contrast, more complex models tend to have higher variance but lower bias, which allows them to achieve low training error. However, they may struggle to generalize to new and unseen data. This creates a trade-off between model complexity and performance. In our case study, we trained various machine learning models to identify the one with the highest performance. A list of these models, along with brief descriptions, is provided in the following.

#### Linear Regression Model

The linear regression model is one of the simplest functions that can be used to predict the outputs. This model has high bias but low variance. Eq. 3.27 shows the linear function for the leakage position prediction system.

$$y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_6 x_6 \tag{3.27}$$

Note that $x_i$ is the normalized flowrate of port $i$. By adding interaction terms, complexity of the linear model can be increased. This will result in a lower bias but higher variance. We will call this interaction linear model, shown in Eq. 3.28.

$$y = b_0 + b_1 x_1 + b_2 x_2 + ... + b_6 x_6 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + ... + b_{56} x_5 x_6 \tag{3.28}$$

Adding interaction terms can be beneficial, because an over simplified model may fail to correctly represent the system. The robust linear model makes the model better equipped at handling outliers. However, the training data may not contain

many outlier points. Eventually, the model can be trained by a stepwise algorithm, in which different terms are added and trained in steps.

### Regression Tree

Regression trees are a type of decision tree in which every leaf represents a numeric value. Regression trees are useful when the input-output relationship is too complicated to be represented as a linear function [152]. Different regression trees with different leaf sizes could be designed and trained for the problem. Obviously, as the leaf size increases, the number of parameters for defining the model increases and an excessive number of parameters lead to overfitting.

### Support Vector Machine(SVM) Regression

Support Vector Machine Regression or Support Vector Regression (SVR) is a supervised learning method that is both powerful and robust. SVM utilizes hyper-planes as decision boundaries [153]. The linear SVM function is defined as:

$$y = b_0 + b_1x_1 + b_2x_2 + ... + b_6x_6 = x'b + b_0 \tag{3.29}$$

The distance of $\epsilon$ in Eq. 3.30 which is the distance to the boundary determines the margin of tolerance that all of the data points will be subjected to:

$$\forall n : |y_n - (x_n'b + b_0)| < \epsilon \tag{3.30}$$

Finally, the loss function would be defined as Eq. 3.31.

$$L_\epsilon = \begin{cases} 0 & if \ \ |y - (x'b + b_0)| < \epsilon \\ |y - (x'b + b_0)| - \epsilon & otherwise \end{cases} \tag{3.31}$$

The equations presented above are the core functions used to implement an SVM. Although an SVM can be initialized with linear kernels, the possibility of specifying different kernels is one of the main advantages. The ability to initialize non linear kernels is what makes an SVM so versatile, especially in problems where linear functions do not adequately represent the problem. For instance, quadratic and cubic functions can be used as kernels in SVM regression. The more polynomial terms in the kernel generally results in a more accurate approximation of the function, however,

this comes at at the expense of being susceptible to overfitting.

### Gaussian Progress Regression (GPR)

Gaussian process regression (GPR) models are non-parametric probabilistic models which, similar to SVM, utilize kernels [154]. A Gaussian Process is a collection of random variables, each with Gaussian distributions. A Gaussian process can be fully specified (described) using the mean and covariance functions as follows:

$$f \sim GP(m, k) \tag{3.32}$$

This means that $f$ is distributed as a Gaussian process with mean function $m$ and covariance function $k$ which is equivalent to a kernel. Four different kernels have been selected and trained for this problem. The squared exponential kernel is one of the most popular kernels used in Gaussian process modelling. It can be computed as[155]:

$$k(x_a, x_b) = \sigma^2 exp(-||x_a - x_b||^2/2l^2) \tag{3.33}$$

where $\sigma^2$ is the overall variance and and $l$ is the lengthscale parameter which specifies the width of the kernel. Using the squared exponential kernel results in a smooth prior on functions sampled from the Gaussian process. Similar to the squared exponential, the rational quadratic function kernel (Eq. 3.34) will result in a smooth prior functions sampled from the Gaussian process. The rational quadratic kernel can be interpreted as an infinite sum of different exponentiated quadratic kernels with different lengthscales. An example of one of these sums can be defined by Eq. 3.34:

$$k(x_a, x_b) = \sigma^2(1 + ||x_a - x_b||^2/2\alpha l^2)^{-\alpha} \tag{3.34}$$

where $\alpha$ is the scale-mixture. The matern ($\nu = 5/2$) kernel benefits from long and short lengthscales which enable extrapolation. Therefore, it can learn from unseen regions of the function's domain[156].

### Neural Networks (NN)

Neural networks were first developed by McCulloch and Pitts in 1943 based on mathematics and algorithms [157] for the purpose of artificial intelligence. However, they have recently seen a resurgence in prominence in the machine learning dialogue. Neural networks have become a popular and useful tool for classification, clustering, pat-

tern recognition and prediction in many disciplines [158]. Neural networks are famous and powerful because of their excellent properties of self-learning, adaptivity, fault tolerance, nonlinearity, and advancement in input to output mapping [159].

Based on the nature of our problem, the input-output function is not too complicated. Thus, a deep and complicated network is not necessarily required, and may, in fact, result in overfitting. Therefore, a shallow layer neural network may be sufficient to model the problem. Based on the architecture of neural network shown in Figure 3.25, five different networks have been designed and trained for the leakage location regression problem. There are six features (inputs) in the first layer and a regression output layer. Note that the neural networks with this architecture must be trained for distance and angle, separately. For single hidden layer network configurations, there are 10 neurons in the middle layer in the narrow network, 25 in the medium and 100 neurons in the so called wide network. The bi-layered and tri-layered networks have two and three hidden layers with 10 neurons respectively.

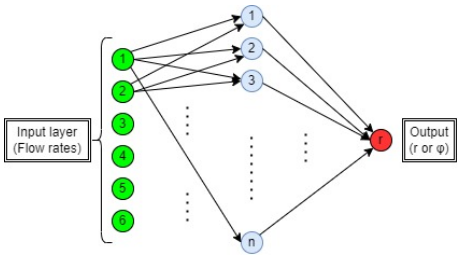


Figure 3.25: General network architecture trained for the problem

## 3.8 Machine learning model for single and multiple leakage Classification

As discussed in Section 3.6, the normalized flow rate generated by a single leakage typically differs from that of multiple leakages. This difference forms the basis for training a classification algorithm to determine the number of leakages in the system, which is a crucial first step toward leakage localization in setups with an unknown number of leakages. If the algorithm classifies the number of leakages as one, the machine learning regression model for single-leakage localization, detailed in the previous section, will be applied. If the number of leakages is classified as two, a new algorithm will need to be developed and implemented. Thus, the initial focus is on

designing and training a classification model. Data collection and data structure as well as the models will be discussed in the following sections.

### 3.8.1 Data collection and generation

To train the classification model effectively, the dataset must be both unbiased and balanced. An unbiased dataset ensures that the data is sampled uniformly across the entire layup, rather than concentrated in a specific region. A balanced dataset contains an approximately equal number of samples from each class, which simplifies the training process and improves the accuracy of the cost functions. Using biased or unbalanced data can lead to training errors and inaccurate predictions.

To achieve an unbiased dataset, a large volume of data needs to be collected or generated from all potential regions of the layup. This often necessitates the use of synthesized data for comprehensive coverage. In this case, a node network with dimensions similar to the 81cm×72cm layup shown in Figure 3.8 has been employed for data synthesis. Simulations can be run on this network, where data for each leakage class is generated by applying a high voltage to the individual nodes.

Synthesizing data for the two-leakage class follows a similar approach, where high voltages are applied to two simultaneous locations. We select a comparable amount of data to maintain a balanced dataset. Leakage pairs are randomly chosen from all possible leakage locations, and with this relatively large dataset, balance is achieved. As discussed in Section 3.6, key findings indicate that, in multiple leakage scenarios, the two leakages must maintain a minimum distance to ensure flow rates distinct from those in single-leakage scenarios. This is especially important for the training dataset. We conducted various sets of dataset generation and training phases with three different Minimum-Distance-Constraints (MDC) for the two-leakage class: 10 cm, 27 cm, and 35 cm. The training results will be discussed in the next chapter. A higher MDC leads to improved accuracy in the training, validation, and test datasets. However, the model becomes less reliable in predicting leakage pairs with shorter distances, reducing its effectiveness in classifying single and multiple leakages in real-world scenarios. On the other hand, selecting a lower MDC may decrease accuracy, increase variance error, and hinder the model's ability to generalize to new data, ultimately diminishing its effectiveness. Therefore, it is crucial to choose an appropriate MDC for the training dataset.

### 3.8.2 Data structure

Data would be pretty much similar to the data structure in Section 3.7.1 and the features or the inputs are the normalized flow rates. The output on the other hand, is the class of data whether it belongs to the class of single leakage (class 1) and two leakage (class 2). Figure 3.26 shows a portion of the data used for training the model. The first six columns of F1 through F6 are normalized flow rates and the last column is the class that each set belongs to.

| F1 | F2 | F3 | F4 | F5 | F6 | class |
|---|---|---|---|---|---|---|
| 0.040044 | 0.048436 | 0.021331 | 0.027574 | 0.111874 | 0.750741 | 1 |
| 0.021331 | 0.048436 | 0.040044 | 0.750741 | 0.111874 | 0.027574 | 1 |
| 0.030141 | 0.058539 | 0.041318 | 0.438021 | 0.379534 | 0.052447 | 2 |
| 0.035152 | 0.058718 | 0.035152 | 0.082257 | 0.706462 | 0.082257 | 1 |
| 0.03494 | 0.061569 | 0.040364 | 0.210315 | 0.565794 | 0.087019 | 2 |
| 0.041585 | 0.06432 | 0.038593 | 0.457314 | 0.188436 | 0.209752 | 2 |
| 0.053036 | 0.065278 | 0.029194 | 0.038621 | 0.163703 | 0.650168 | 1 |
| 0.029194 | 0.065278 | 0.053036 | 0.650168 | 0.163703 | 0.038621 | 1 |
| 0.054832 | 0.069689 | 0.032072 | 0.044368 | 0.206255 | 0.592784 | 1 |
| 0.032072 | 0.069689 | 0.054832 | 0.592784 | 0.206255 | 0.044368 | 1 |
| 0.04631 | 0.070891 | 0.040215 | 0.088394 | 0.314799 | 0.439391 | 2 |
| 0.046587 | 0.071853 | 0.039691 | 0.0768 | 0.631294 | 0.133775 | 1 |
| 0.039691 | 0.071853 | 0.046587 | 0.133775 | 0.631294 | 0.0768 | 1 |
| 0.038144 | 0.073084 | 0.052127 | 0.214083 | 0.548812 | 0.073751 | 2 |
| 0.050605 | 0.073137 | 0.037933 | 0.064664 | 0.590091 | 0.183569 | 1 |

Figure 3.26: Data structure used for single/multiple leakage classification problem

### 3.8.3 Training methods and validation

The training methods and validation for single/multiple leakage classification is similar to Section 3.7. However, instead of training regression models, classification models must be trained. We have used different classification models such as narrow, medium and wide neural networks (NN), SVM methods with quadratic, cubic, fine and medium Gaussian kernels, and fine and weighted KNN (K-nearest neighbors). The output of the model is a class instead of a response as previously seen in regression models. After training the model, the generalizability will be assessed on a test dataset. The data used in test dataset is new and has never been seen by the model. In addition, the performance of the best trained model will be tested on new experimental data to put everything on test for a real-world problem.

The next step after classification of leak numbers is localizing leakage pairs in the multiple leakage scenarios. In order to delve into that, the concept of graph neural networks must be discussed first.

## 3.9 Graph Neural Networks: a More Generalizable Architecture

The proposed framework demonstrates considerable potential in managing diverse layup configurations and accurately predicting leakage locations. By carefully addressing data types and implementing preliminary measures to minimize overfitting, the framework achieves a notable degree of generalizability. However, significant changes in layup size or configuration necessitate restarting critical steps, such as data synthesis and model training, which undermines the desired level of automation.

Furthermore, the regression models discussed in Section 3.7.2 do not incorporate key factors such as layup size. Instead, these models rely solely on data to train a complex input-output function, effectively treating the model as a black box. Consequently, a more automated and versatile framework is required to overcome these limitations. To address this, we propose adopting graph neural network (GNN) designs. GNNs are neural models that capture graph dependencies through message passing between graph nodes.

### 3.9.1 Graph neural network

Graph Neural Networks (GNNs) are a type of deep learning model designed to operate on data structured as graphs [160]. Unlike traditional neural networks that operate on fixed-sized inputs like images, GNNs are capable of handling data with complex relational structures, such as social networks, molecular structures, physics systems, recommendation systems and knowledge graphs [161].

In a graph, the data is represented as a collection of nodes or vertices connected by edges or relationships. GNNs leverage this connectivity information to learn and extract meaningful representations of the graph's nodes and edges. The key idea behind GNNs is to propagate information through the graph by iteratively aggregating and updating features associated with each node based on its neighborhood [162] also called message passing. The main components of GNNs are as following.

1. Node feature: Each node in the graph is associated with a feature vector that represents its initial attributes. These features could be any kind of information associated with the node, such as textual data, numerical values, or categorical labels. In our case of study, nodes will be allocated to the small elements (squares) shown in Figure 3.8 and the node features are their normalized flowrate values.

2. Message passing: GNNs utilize a message passing scheme to propagate information between connected nodes. At each iteration, nodes gather information from their neighbors, perform local computations, and update their own features accordingly. This process allows nodes to incorporate information from their surrounding context.

3. Aggregation function: During the message passing process, an aggregation function is used to combine the features of neighboring nodes. This function can be as simple as averaging or summing the neighboring features or more complex, such as using attention mechanisms to weight the importance of different neighbors. In our case, nodes that contain data could have a more weight compared to nodes without data.

Graphs can be categorized into different types and scales, including directed and undirected graphs. In directed graphs, edges are oriented from one node to another and transmit messages accordingly. On the other hand, each edge in an undirected graph can be viewed as two directed edges. In the context of our specific case study, the direction of message passing is determined by the introduced data, allowing messages to flow in any direction. Thus, an undirected graph is appropriate for our scenario.

Graphs could be homogeneous or heterogeneous. Nodes and edges in homogeneous graphs have same types, while nodes and edges have different types in heterogeneous graphs. When input features or the topology of the graph vary with time, the graph is regarded as a dynamic graph. Otherwise, it's a static graph. For our case study, we will design a homogeneous static graph.

Task type also determine the design of loss function. There are usually three kind of different tasks in graphs. Node-level tasks focus on nodes, which include node classification, node regression and node clustering. Edge-level tasks are edge classification and link prediction, which require the model to classify edge types or

predict whether there is an edge existing between two given nodes. Graph-level tasks include graph classification, graph regression, and graph matching, all of which need the model to learn graph representations. For out specific case, we want the model to classify the correct node as the leakage location, therefore we are doing a node classification and the cost function must be designed accordingly.

Finally, considering the aspect of supervision, our settings can be classified as either supervised or semi-supervised learning, depending on the availability of data. In the case of most layup configurations, there may be an inadequate number of data samples available for all nodes. Consequently, the model needs to be trained using a limited amount of labeled data. Therefore, the approach in such scenarios is typically referred to as semi-supervised training.

### 3.9.2 Grid neural network design

Based on the information presented in Section 3.9.1, it is possible to design an effective GNN. In this design, each node in the GNN corresponds to a square-shaped area defined in the layup. These nodes are arranged in a grid formation to accurately represent our physical system. To refer to this specific design, we will use the term "Grid Neural Network," as illustrated in Figure 3.27. This is similar to the volumetric flow rate matching method used in [137], where flow rate values are stored as an array in the grid. However, there are features added to this method based on message passing in GNNs.

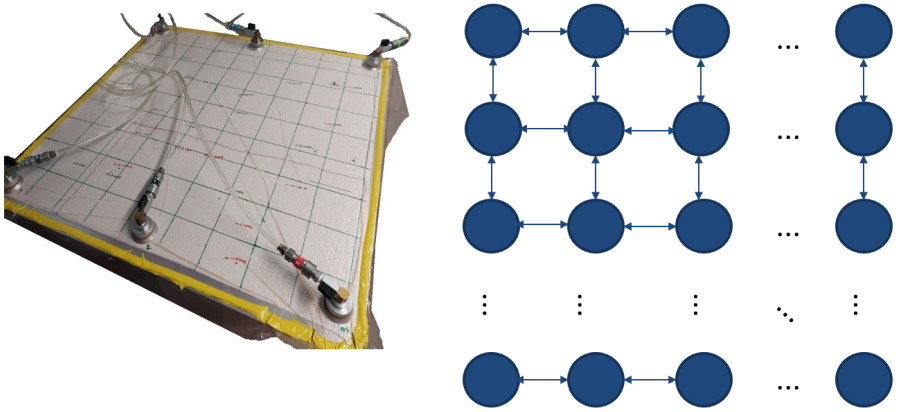

Figure 3.27: Grid neural network design for the physical system

With this design, the connectivity between nodes is already known, eliminating the necessity of explicitly defining an adjacency matrix. An adjacency matrix is a

square matrix used to represent a finite graph. The elements of the matrix indicate whether pairs of vertices are adjacent or not in the graph. In our case, each node corresponds to a small divided area on the layup.

In this design, it is possible to define additional nodes to create a finer mesh, allowing for more detailed representation. However, it is important to note that the prediction accuracy is constrained, and an excessive number of meshes can lead to diminished accuracy or overfitting. Therefore, striking a balance is crucial to maintain an optimal level of accuracy without compromising the model's generalization capability. For the start, one node will be used for an squared-shaped area with the size of $10cm$.

At the beginning, the network is not trained and node feature vector is a uniform vector with all elements having the same value. The node feature vectors could be initialized with physical knowledge that will be discussed in the next section. When a data sample is introduced to the system, first the aggregation function defined in Eq. 3.35 updates the features of the corresponding node with a momentum rate in each iteration and creates a message.

$$M^{x,y}{}_i = pF^{x,y}{}_{i-1} + (1-p)D^{x,y} - F^{x,y}{}_{i-1} \tag{3.35}$$

In the above equation, $M^{x,y}{}_i$ is the message created in $i$-th iteration in location $(x, y)$ which are the row and column number of the corresponding node. $p$ is the momentum rate and a design parameter. $D^{x,y}$ is the input data features (flow rates) which in our case is static since we are averaging the flowrate values. The input value could be a function of time if the user prefers to use dynamic flow rate values. $F^{x,y}{}_{i-1}$ is the feature vector of the node in the previous state. The next step is defining the message passing function.

Message passing facilitates the propagation of information between nodes, enabling iterative updates of the neighboring node values. In our scenario, when introducing input data, it can be iteratively utilized to update the values of adjacent nodes. Subsequently, in the following iteration, a new, albeit weaker, message can be generated to update additional neighboring nodes. This process allows for the efficient utilization of a single input for GNN training. It is crucial to incorporate a discount factor for each new message transmitted to the nodes as it is undesirable to update the nodes far from where the sample was introduced. Without this discount factor, oversmoothing will occur and the representations of nodes in the graph tend

to become indistinguishable from each other. Essentially, the node features lose their individuality and converge toward a similar value or distribution. This hinders the model's ability to differentiate between nodes, leading to a decline in performance. Figure 3.28 represents the message passing process.

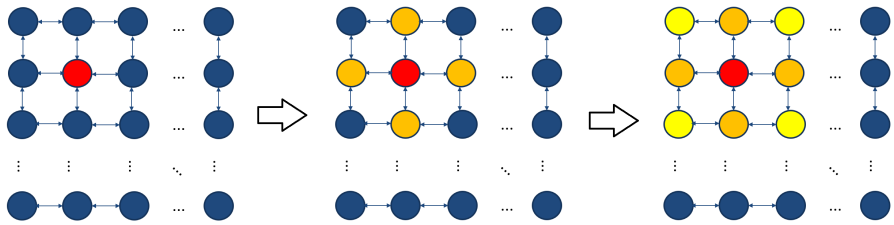

Figure 3.28: Message passing procedure in each iteration from left to right; messages are discounted over time to reduce oversmoothing.

The system's generalizability and adaptability experience a substantial boost with this new design. Users can input the desired layup size, and a grid neural network can be instantly constructed. Following that, synthesized or experimental data can be introduced to the system for training purposes. It is important to note that the model's accuracy improves as the samples become more widely distributed throughout the area.

### 3.9.3 Physics-informed grid neural network design

The proposed grid neural network effectively utilizes input data to train a cluster of nodes and has the potential to outperform traditional machine learning models. To ensure proper training of the network, it is crucial to introduce diverse data from various regions of the graph into the system. However, integrating the model with physical knowledge can elevate its performance beyond its current capabilities. Any relevant physical information can be incorporated into the system. Physically informing the system offers two major improvements. Firstly, it reduces the amount of data required for training. Secondly, the system demonstrates superior performance compared to the original method.

For instance, one simple but incomplete observation is that as the proximity of leakage to a specific vacuum port increases, the flowrate of the vacuum port also increases. This inverse relationship between flowrate and distance can be approximated using a simple equation. It is important to note that the physical information does not need to be entirely accurate, as the model will gradually adjust during the

training process. Nonetheless, informing the system with such knowledge provides a significant advantage during both the training phase and prediction.

The distance-flow rate relationship is primarily effective for uniform layups with simple geometries and has been utilized in prior studies [137]. For more complex layups, incorporating advanced and detailed physical knowledge can enhance model accuracy. The method outlined in Section 3.10.2, which leverages Kirchhoff's law, is well-suited to handle such intricate scenarios.

### 3.9.4 Leakage localization method

Once the system is trained, the COHO system's measured flow rate is used as input during the prediction phase for the node classification problem. The system is designed to accommodate various activation functions, such as Softmax, for node classification tasks. By leveraging Softmax, the system generates prediction probabilities, enabling it to pinpoint the most probable nodes. This approach allows users to prioritize investigating the most likely leakage region first and proceed to subsequent predictions if initial attempts are unsuccessful. To apply the Softmax function, the similarity of each node in the GNN must first be calculated using

$$S^{x,y} = \alpha(1 - \Sigma_{j=1}^{n}|F^{x,y}(j) - f(j)|)\,. \tag{3.36}$$

In Eq. 3.36, $\alpha$ represents the gain, typically a large value around 100; $j$ denotes the index of the vacuum port; $F^{x,y}$ is the flow rate vector computed during the last iteration of the GNN for the coordinates $(x, y)$, and $f$ is the experimentally measured flow rates by sensors. If the experimental flow rates closely match the values of $F^{x_0,y_0}$ at a specific location $(x_0, y_0)$, the similarity score $S^{x_0,y_0}$ will be higher. Consequently, this leads to a higher probability output by the Softmax function, as shown in Eq. 3.37, for that specific location.

$$P^{x,y} = \frac{e^{S^{x,y}}}{\Sigma_{x=1}^{a}\Sigma_{y=1}^{b}e^{S^{x,y}}} \tag{3.37}$$

The Softmax function calculates the probability across all possible grid locations. The values $a$ and $b$ represent the total number of nodes in $x$ and $y$ directions of the layup. For instance, both $a$ and $b$ are nine for the layup illustrated in Figure 3.27. Note that if the layup has a shape other than a rectangle, Eq. 3.37 must include all nodes.

Bayesian updating can also be incorporated at this stage to update the node probabilities. Additionally, expert knowledge can be utilized as priors. For instance, it is commonly known that the probability of finding a leakage at the boundaries is higher than locating a leakage within the vacuum bag due to potential issues with improper sealing of the sealant tape by the user.

### 3.9.5 Multiple leakage localization

Arguably, the most significant advantage of the GNN design lies in its ability to localize multiple leakage scenarios. The classification of the number of leaks on the layup was discussed in Section 3.8. It was explained that if the classification algorithm detects a single leakage, the single-leakage localization models introduced in Section 3.7 can be employed to discover the most probable area of the leak. However, the challenge of localizing leaks remains unresolved when the algorithm identifies two leaks on the layup. This is where the PI-GNN design can be utilized to address the problem effectively.

Since the objective is to identify two nodes as the leakages, this constitutes a multi-label classification problem. One approach to solving this problem is to treat each pair of nodes as a single class. With this method, if there are 81 different nodes on the layup, the total number of classes will be $\binom{81}{2} = 3240$. For a classification problem, it is typically recommended to have at least 10–100 samples per class, which would require a minimum of 30,000 samples for training. Generating such a large dataset is extremely challenging, even with simulations. Moreover, samples belonging to a specific class will exhibit minimal variability, as the flow rates remain consistent when the leakage locations are fixed. This makes it difficult to synthesize diverse data for each class.

To simplify the problem and avoid the complexities of multi-label classification, a different method can be utilized. The PI-GNN structure, resembling a large grid, stores flow rate values and their corresponding locations at each node as a vector. Prior research has demonstrated two critical principles: first, the flow rates caused by multiple leakages adhere to the superposition rule, meaning the effects of individual leaks can be summed together; second, it is assumed that both leakages on the layup are of equal size.

Leveraging these principles, the flow rate outputs of any two leaks on the layup can be averaged, and this averaged value can then be compared with the measured

flow rates from the sensors. In other words, the system compares the measured flow rate values with the average of the embeddings for each pair of nodes. For any two leakage locations $(x_1, y_1)$ and $(x_2, y_2)$, the total flow rate output can be calculated as

$$\vec{F}^{(x_1,x_2),(y_1,y_2)} = \frac{1}{2}(\vec{F}^{x_1,y_1} + \vec{F}^{x_2,y_2})\,. \tag{3.38}$$

Afterward, the node pairs with the smallest differences from the measured flow rates can be ranked. The subsequent steps follow the same methodology described in the previous section, utilizing Equations 3.36 and 3.37.

The key requirement for using this model is to achieve a precise, well-trained GNN. An advanced physical solution, based on Kirchhoff's node law, will be discussed in the next section. Additionally, a training method for the model's parameters, particularly for non-uniform layups, is presented in Section 3.11. Together, these methods lay the foundation for an enhanced GNN that can be used for multi-leakage localization.

The only remaining part is what if the leaks have different sizes. This will show itself in the Eq. 3.38. Different weights could be assigned to the flow rates from each leakage instead of the $\frac{1}{2}$ coefficient.

$$\vec{F}^{(x_1,x_2),(y_1,y_2)} = \frac{(c_1\vec{F}^{x_1,y_1} + c_2\vec{F}^{x_2,y_2})}{c_1 + c_2} \tag{3.39}$$

If one leakage is significantly larger than the other ($c_1 \gg c_2$), the corresponding flow rates will be dominated by the larger leakage, effectively masking the smaller one. In this case, the single/multiple classification algorithm introduced in Section 3.8 will classify the problem as a single leakage, and the location of the dominant leakage will be identified using the single leakage localization method. After repairing the larger leak, the flow rates from the remaining leakage will become apparent, allowing the user to locate the second leakage.

However, when the leakage sizes are unequal but comparable, the problem becomes much more complex. In this scenario, there could be a continuous distribution of flow rates between the two leaks, meaning that flow rate values will fall within an interval that corresponds to the same leakage location. These flow rate values may overlap with those generated by other leakages at different locations and sizes, making them almost indistinguishable given the current features used.

### 3.9.6 Bayes Theorem on leakage localization

Bayes' Theorem is a fundamental principle in probability theory that describes how to update the probability of a hypothesis based on new evidence [163]. The theorem could be mathematically expressed as

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)} = \frac{P(E|H)P(H)}{P(E|H)P(H) + P(E|H')P(H')}\,. \tag{3.40}$$

Terms of the Eq. 3.40 in the aspect of our leakage localization problem are explained in Table 3.4.

Table 3.4: Terms of Bayes theorem in the leakage localization problem

| Term | Description |
|---|---|
| H: Hypothesis | Presence of leakage in the node |
| E: Evidence | Information acquired from GNN or the user |
| P(H): Prior | Initial probability: leakages are more likely to happen on boundaries |
| P(H\|E): Posterior | Updated probability given the evidence |
| H': Null-hypothesis | Presence of leakage in all of the other nodes |

The initial probability reflects the prior knowledge before measuring flow rates. For instance, it is more likely to find leaks near the boundary or on the tacky tape rather than in the center of the layup. This is due to the fact that the technician could have errors in sealing the vacuum bag. In certain layups, regions containing elements with sharp edges, known to the expert user, may have a higher probability of leaks. After the system measures steady-state flow rates, new evidence from the GNN model can be introduced. This updated information allows the system to adjust the probabilities, further refining the likelihood of leakage in various regions.

Furthermore, the evidence in the leakage localization process refers to any new information acquired during the process. For example, nodes are initially sorted based on the probability of having a leakage. The user begins by investigating the region with the highest probability. If no leakages are detected, new evidence can be incorporated into the system. Specifically, the fact that no leakage was found in that area can be used to update the posterior probabilities.

## 3.10 Vacuum port placement optimization

One challenge in leakage localization, especially in large-scale components, is achieving an optimized configuration for vacuum port placement. Vacuum ports must be strategically placed for two main purposes: (1) to evacuate air from the vacuum bag without leaving air bubbles or voids, and (2) to predict the locations of unknown potential leakages with the highest possible accuracy. Pishvar et al. proposed using magnets to increase consolidation pressure and reduce the likelihood of voids in the structure [164]. Mode et al. developed an active online control system for resin injection using image analysis from cameras and controlled four injection ports with solenoid valves [165]. Maung et al. studied and optimized the number and locations of inlet ports for a fuselage measuring 7900 mm, aiming to minimize fueling time using a genetic algorithm [166].

While there are numerous studies on the optimization of injection ports, there is a lack of research on optimizing vacuum port locations. Unlike resin filling time, which is a critical factor, air evacuation does not depend heavily on vacuum time, as it is not a time-consuming process. However, leakage detection and repair can be time-intensive, especially in industries where reliability and safety are critical, such as aerospace, automotive, and marine sectors.

### 3.10.1 Cost function

The first step in solving optimization problems is to define the cost function. In this case, our goal is to identify optimal locations for vacuum ports to achieve the most accurate predictions after training. Simply put, we are seeking the best positions for vacuum ports that will provide the most valuable data for training our machine learning model.

We will use the initial 90 cm $\times$ 90 cm setup shown in Figure 3.3 as a case study. The layup is divided into a $9 \times 9$ grid, with each cell representing a 10 cm square. Any of these locations can be selected for placing a vacuum port. A finer grid is unnecessary because the current layout already covers the area sufficiently, as seen in the figure. Moreover, refining the grid would only increase computation time by significantly enlarging the state space.

In this study, the number of vacuum ports is not an optimization parameter; we use our standard six vacuum ports. While increasing the number of vacuum ports and flow rate sensors could provide more features for machine learning, leading to

better leakage predictions, using more vacuum ports does not negatively impact the manufacturing process, unlike adding more resin injection ports, which could leave residual resin.

To better illustrate the problem, let's assume we want to place two vacuum ports on a layup. If these ports are placed too close to each other, the flow rate values will likely show little difference, even if there are leaks in different areas of the layup. In other words, the variance in the measured features decreases as the vacuum ports move closer together. This reduced variance makes it harder to train a machine learning model effectively without the risk of overfitting. Therefore, the cost function should aim to maximize flow rate variance across potential leakage locations in the layup configuration [167].

Assuming having a total number of $M$ different leakage locations, the cost function can be defined as Eq. 3.41.

$$y = -\frac{1}{M}\sum_{m=1}^{M}(\sigma^2(f_i))_m \tag{3.41}$$

Here, $f_i$ represents the flow rate of the $i^{th}$ channel. The negative sign is introduced to define "$y$" as a cost function.

For a given port configuration, the flow rates for all potential leakage points are first calculated. Next, the variance of the flow rates across each channel is computed, followed by averaging the variance across all leakages and channels. Finally, the $-y$ value can be used to compare different port configurations.

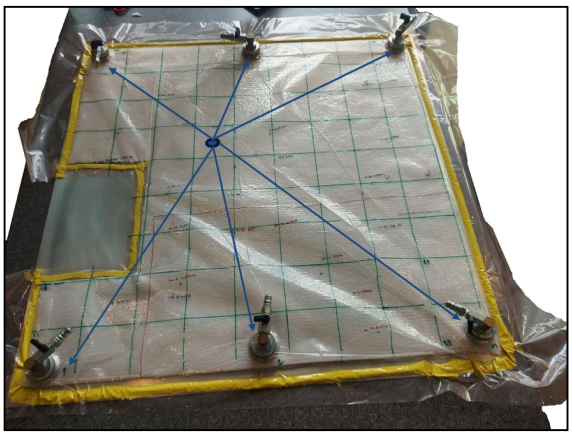

Figure 3.29: Cost function calculation for a single leak for vacuum ports configuration in corners

Figure 3.29 shows a step in calculating the cost function for a leakage event. The six vacuum ports are placed at the corners of the layup, and the flow rates at each of these ports, induced by the leakage, are either measured or modeled. This process is repeated for the remaining 75 regions, which are divided into squares, and the total variance is then calculated. This variance serves as the basis for comparison. The computed value is then compared with the variance from an alternative vacuum port configuration. The next section will discuss the methods used to determine the flow rates.

It is important to note that leakages are more likely to occur along the boundaries of the vacuum bag, specifically on the tacky tape. This is because the tacky tape, which seals the edges of the vacuum bag to the mold, may not adhere perfectly due to factors such as dust, wrinkles, or improper application. Ensuring a proper seal with the tacky tape is essential for maintaining vacuum integrity, and any imperfections or contaminants can lead to leaks. Additionally, edges and corners, where the bag folds or overlaps, are common areas for potential leaks [168].

With this in mind, the cost function in Eq. 3.41 can be adjusted to give more weight to the variance caused by boundary leakages. This adjustment makes it easier to detect leakages in areas where they are more likely to occur. Eq. 3.42 presents the weighted cost function, where $M$ represents the number of leakages in the center of the layup, $K$ represents leakages on the boundary, and $\delta > 1$ is the weight assigned to locations more prone to leakage.

$$y = -\frac{\sum_{m=1}^{M}(\sigma^2(f_i))_m + \delta \sum_{k=1}^{K}(\sigma^2(f_i))_k}{M + \delta K} \tag{3.42}$$

### 3.10.2 Flow rate calculation methods

To compute the flow rates in Eq. 3.41 for the cost function, a robust and generalizable methodology is necessary to accommodate different layup sizes, permeabilities, and varying vacuum port locations. The methodology for determining flow rates should meet the following criteria:

i. **Accuracy**: Ensures that flow rate estimations are reliable and closely match experimental results.

ii. **Generalizability**: Capable of handling various layup sizes and permeabilities, ensuring applicability across different scenarios.

iii. **Leakage Mitigation Cost**: Accounts for the labor and expenses associated with identifying and relocating leaks, impacting overall costs.

iv. **Vacuum Port Configuration Cost**: Considers the labor and expenses involved in modifying vacuum port configurations, contributing to the robustness of the methodology.

In our case study, which involves 81 distinct locations and six vacuum ports, the optimized solution must be selected from combinations of six out of 75 possible configurations. This results in a total of 324 million vacuum port configurations, each with 75 different leakage scenarios. Although symmetries and other methods can potentially reduce this number, they are insufficient to make the experiments feasible. Despite the advantages of the simulation method introduced in Section 3.5, a new model must still be created whenever the vacuum port locations change, making it impractical to simulate such a large number of scenarios within a reasonable timeframe. Therefore, a customized, calculation-based model is essential for accurately estimating flow rates. Given the large number of possible configurations, the ability to quickly iterate through various vacuum port setups becomes more important than achieving high accuracy in flow rate estimation. To address this, we introduce two methods for estimating flow rates: a simplified approach and a more precise method.

### Simplified flow rate calculation

Flow rates could be estimated by applying Darcy's law previously introduced in Eq. 3.4. Based on that, it was shown that the volumetric flow rate could be calculated as Eq. 3.7. The term $g(r)$ then could be written as polynomial function shown in Eq. 3.8. By neglecting internal vacuum port resistance and assuming the leakage isn't extremely close to the vacuum ports, we can estimate normalized flowrates using the first term of Taylor series as follows:

$$F_i = \frac{\frac{1}{r_i}}{\sum \frac{1}{r_i}} \tag{3.43}$$

$F_i$ represents the normalized flow rate for the $i$-th vacuum port and $r_i$ is the distance between the leakage and the $i$-th vacuum port. The equation is based on the fact that generally, the flow rate decreases as the distance between the leakage and vacuum port increases and higher polynomial terms have been neglected. Utilizing

this approach allows for rapid estimation of flow rates across various configurations of vacuum port locations and different leakage positions. However, two key assumptions must be upheld for this estimation to hold validity:

i. Uniform permeability.

ii. A minimum distance between the leakage and the vacuum port.

With these two conditions satisfied, all the criteria in 3.10.2 will be satisfied. Although this simplified methodology for flow rate estimation may introduce some errors, it allows for rapid calculation. Its primary objective is to compute a cost function and not being used for leakage localization purpose.

### Precise Flow rate Calculation

A precise method for calculating flow rates is proposed for scenarios where the assumptions underlying simplified flow rate estimations are invalid or when higher accuracy is required. This approach leverages the electrical-circuit analogy described in [138], without the need for an explicit circuit model. To apply this method, each discrete square section of the vacuum bag in Figure 3.29 is treated as a node, interconnected by resistors representing the viscous resistance of the breather material. A voltage is assigned to each node, as shown in Figure 3.30.

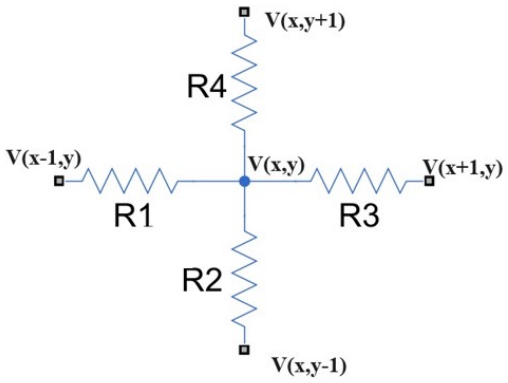


Figure 3.30: Kirchhoff's junction rule applied to the problem for flowrate precise calculation

The governing rule for voltage determination is expressed by Eq. 3.44, which is rooted in Kirchhoff's junction rule [169].

$$\frac{v(x,y)-v(x-1,y)}{R_1}+\frac{v(x,y)-v(x,y-1)}{R_2}+\frac{v(x,y)-v(x+1,y)}{R_3}+\frac{v(x,y)-v(x,y+1)}{R_4}=0 \quad (3.44)$$

By iterating Eq. 3.44 and applying the boundary conditions for leakage and vacuum port values, the voltage distribution for each node can be computed. From this, the currents entering each ground can be determined. In the electrical circuit model, voltage values correspond to the pressure within the physical system, setting the boundary conditions. A high positive voltage is assigned to the leak location, while electrical ground is assigned to the vacuum port locations. This precise calculation method has been experimentally validated and can rapidly compute flow rates for any configuration of vacuum ports and leakage locations. Both the simplified and precise flow rate calculations offer a balance between time efficiency and accuracy. In this method, resistance values can also be adjusted to account for layups with non-uniform permeability.

### 3.10.3 Vacuum Port Location Optimization

One of the primary challenges of this optimization problem is the large state space. Even for relatively simple case studies, the vast number of possible port configurations presents a significant difficulty. While symmetry considerations can help reduce the number of scenarios, this approach has limited applicability, especially in cases involving uniform and symmetrical layups. AS mentioned previously, in our case study with 81 nodes and six vacuum ports, there are over 324 million possible configurations, each requiring flow rate calculations for 75 different leakage locations. This results in an enormous computational cost and time requirement.

Table 3.5 presents the time required to find the optimal solution for various case studies using the proposed simplified and precise methods. The differences in computation time between the two models are evident. For example, when comparing cases 2 and 3, which have a similar number of nodes and vacuum ports, the precise method takes approximately 20 times longer than the simplified approach to determine the optimized configuration. Interestingly, although case 4 involves significantly fewer scenarios than case 2, the computation times are nearly identical. This is due to the higher number of nodes in case 4, which increases the likelihood of potential leakages according to our cost function, thus requiring more calculations. We were unable to measure the time needed to optimize our case study with 81 nodes and six ports

because the process took too long. However, we estimate that it could take roughly 12.67 times the computation time of case 3, approximately 80 hours.

Table 3.5: Time required for finding the optimal solution for different case studies based on simplified (S) or precise (P) method

| Case Number | Nodes | Ports | Scenarios | Model (S/P) | Time (min) |
|---|---|---|---|---|---|
| 1 | 81 | 4 | 1,663,740 | S | 1 |
| 2 | 81 | 5 | 25,621,596 | S | 19 |
| 3 | 81 | 5 | 25,621,596 | P | 380 |
| 4 | 144 | 4 | 17,178,876 | S | 20 |
| 5 | 16 | 6 | 8,008 | S | 1 sec |
| 6 | 16 | 6 | 8,008 | P | 2 |

To address this problem, a bi-level hierarchical optimization approach is proposed [170]. In this approach, the constraint region is implicitly defined by two sets of optimization problems, which are solved in a predetermined sequence. At the first level of the hierarchy, instead of considering all nodes, the layup area is divided into larger regions, each consisting of multiple nodes, as shown in Figure 3.31. The optimization problem at this stage is to determine the best configuration of vacuum ports among these larger regions. In hierarchical optimization, the decisions made at a higher level influence the options available at the lower level [171].

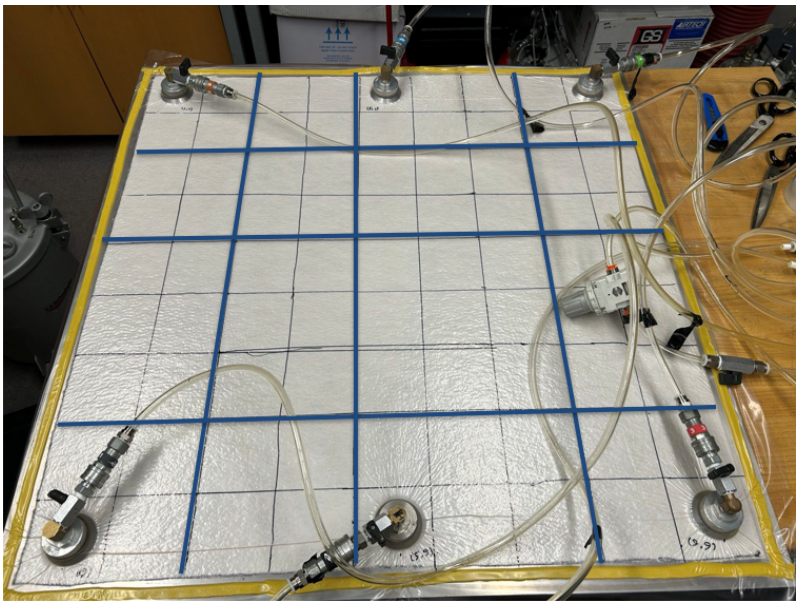

Figure 3.31: Subdividing the layup into smaller areas for the hierarchical optimization approach

Once the optimized regions for placing vacuum ports are identified, the next tier of the hierarchy focuses on refining the resolution to pinpoint the optimal nodes within the designated regions for vacuum bag placement. This illustrates how decisions made at higher levels influence lower-level agents. This approach significantly reduces the number of scenarios, thereby ensuring that the effectiveness, precision, and reliability of the results are maintained. By adopting this method, the number of scenarios is reduced from 324 million to 17,224

The underlying constraint between the two levels of this hierarchy is that a maximum of one vacuum port will be allocated for each area selected by the first level, even though there are multiple nodes within each of these regions. Placing two vacuum ports within a subdivided area would cause their flow rates to be very similar, resulting in a loss of linear independence of the features used for predicting leakage locations. Therefore, implementing this constraint will not affect the optimized solution.

## 3.11 Physical model parameter training

One of the challenges in leakage localization is managing layups with non-uniform permeability. This non-uniformity can occur when different layers of breather material are used in the layup, often to facilitate faster air evacuation in specific areas. Additionally, variations in thickness and the number of reinforcement layers in certain regions further influence the permeability.

The modification of permeability is discussed in detail in Section 3.3.2. Using the electrical circuit analogy introduced in Section 3.5.2, this issue can be addressed by adjusting the resistance values in regions where higher or lower resistance is expected. Since the experiments were designed with prior knowledge of the permeability in different areas, this information is readily available. For instance, regions with higher permeability, as shown in Figure 3.11, are easily identifiable, allowing for a reduction in resistance within the corresponding areas of the electrical circuit. This approach is equally applicable to other cases, such as regions where parts of the breather were removed (Figure 3.10).

A data-driven model must be capable of generalizing to any new setup, even when additional information is not predetermined. Users may also lack knowledge of the permeability settings in their layup. Therefore, we need to design a framework that can learn about areas with non-uniform permeability under the assumption that

configurations and regions are unknown. The system must be self-training, learning the permeability of different regions. We refer to this as "physical model parameter training" since the system must learn parameters such as permeability. This approach is a branch of physics-based machine learning, where both data and physical principles are integrated into the learning process. As a result, generalizability is significantly enhanced while computational costs and data requirements are greatly reduced.

### 3.11.1 Physical model and the case study

To begin with, the physical model and the training parameters must be defined. A simulation based on the grid neural network in Section 3.9.2 is used as the physical model to calculate flow rates as a function of leakage locations. Figure 3.32 shows the case study of a layup with nonuniform permeability. There are two layers of breather in the lower part of the layup and higher permeability in these regions are expected. However, this information will not be used in the training procedure and training results will be compared with the expected results at the end.

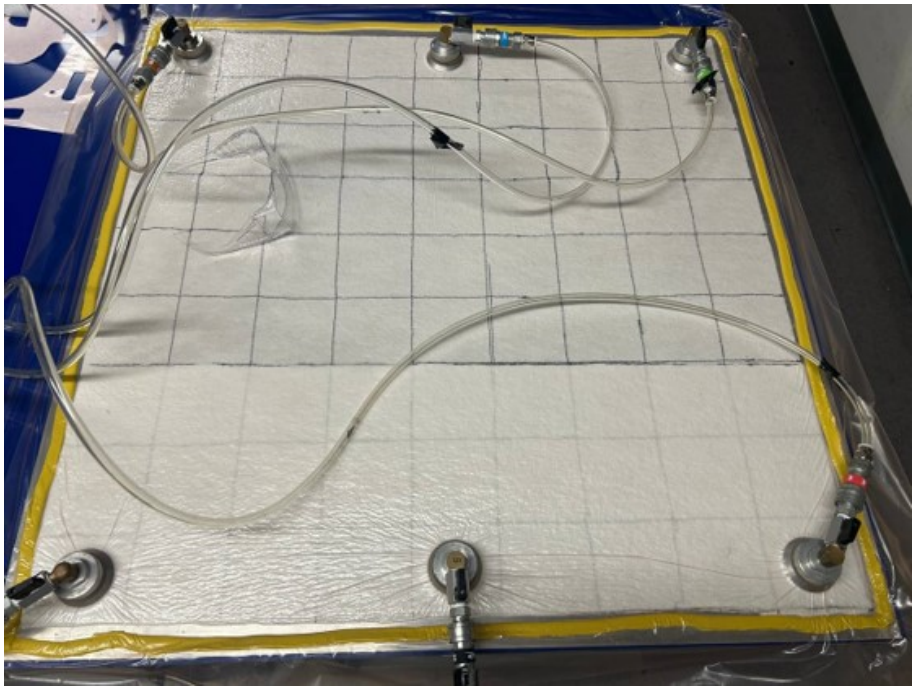

Figure 3.32: Layup with nonuniform permeability as the case study for physical model parameter training

The 90cm×90cm layup will be simulated using a physical model shown in Figure 3.33. The layup is modeled using an 18×18 network of electrical resistors. Each of

these electrical resistors play a small role in defining the permeability of their region. Technically, each of these resistors could be a design parameter.

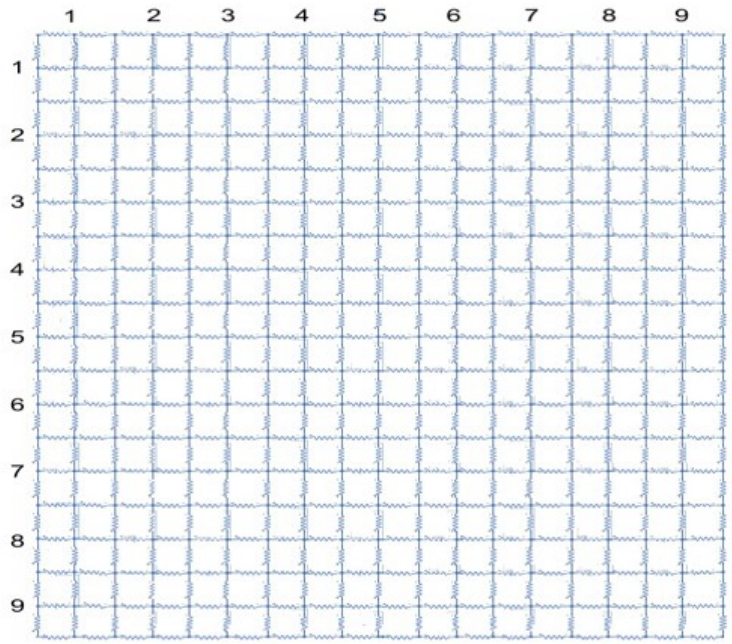


Figure 3.33: Physical model used for simulation the layup

### 3.11.2 Importance of physical model parameter training

In this section, we highlight the importance of using a valid model with appropriate permeability. In our case study, for example, there are two layers of breather material at the bottom of the layup, as shown in Figure 3.32. As a result, the viscous resistance in that region decreases. In the electrical circuit model shown in Figure 3.33, this corresponds to reducing the resistors below the sixth row, since adding an extra layer of breather effectively doubles the porous media for air flow. Consequently, the resistor values in the equivalent regions should be reduced by 50 percent for accurate modeling. We refer to this as the physically designed model, as the parameters are determined based on prior physical knowledge.

In real-world problems, the location and magnitude of non-uniform permeability are often unknown, so the system must learn these from data. We refer to this as the physically trained model, as its physical parameters are trained based on data. The process of physically training the model will be discussed in the next section.

### 3.11.3 Training parameters

The objective is to design a framework that is able to training physical parameters of the model. As previously discussed and showed in Figure 3.33, there is a total of 324 resistors that could be defined as the design parameter. However, this number of parameter is excessive and abundant, and changing the value of a single resistor, does not have a significant effect on the outcome. In addition, there is not enough data to train a model with 324 parameters. If that amount of data is accessible, there is no need to use a physically-informed machine learning model and a regression model or even a simple truth table could be trained instead. Therefore, instead of considering each resistor as a training parameter, a group of electrical resistors in a vicinity will be defined as our training parameter. The layup could be divided into nine smaller regions shown in Figure 3.34. We are assuming a uniform permeability in each of these sub-regions and the relative permeability of each region will be a training parameter. therefore, there are a total of nine training parameters in this case study.

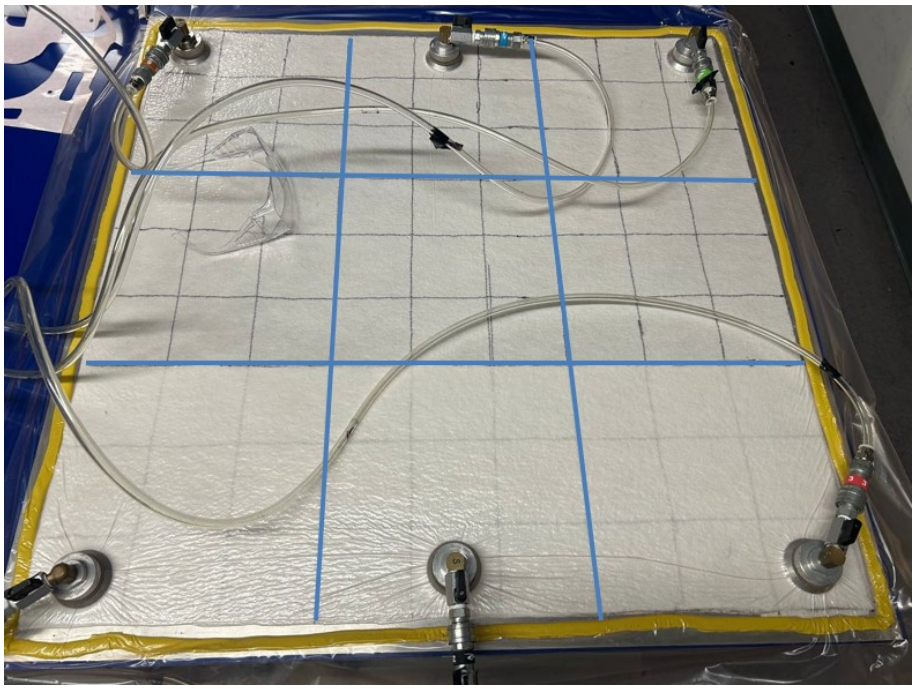

Figure 3.34: Physical model used for simulation the layup

The number of parameters or regions can be adjusted through optimization. Increasing the number of regions or parameters reduces training error but requires more time and data for training. For an easier implementation, we define the inverse of relative permeability, or relative resistance, as our parameter in a matrix. For instance,

a 3×3 matrix of ones would represent the parameters for a completely uniform setup. This 3×3 matrix represents the nine design parameters in our setup. In the setup illustrated in Figure 3.34, the relative resistance of the bottom three regions is half that of the other six regions. The objective, therefore, is to train the parameters to evolve from an initially uniform resistance to the custom resistance pattern shown in Figure 3.35.

The objective values of the training parameters (as shown in the matrix on the right) are based on the assumption that two layers of breather material provide twice the porous medium for air to travel through, thereby reducing resistance by half. While this assumption may not be entirely accurate, the results in the next chapter show it holds to a certain extent. These values were estimated based on experimental observations and physical intuition, which is why we refer to them as "physically designed" parameters. It's important to note that these physically designed values are not incorporated into the model during training but serve as a point of comparison afterward. We use these values to inspect whether training the model can train the parameters to start from the uniform initial values to the objective.

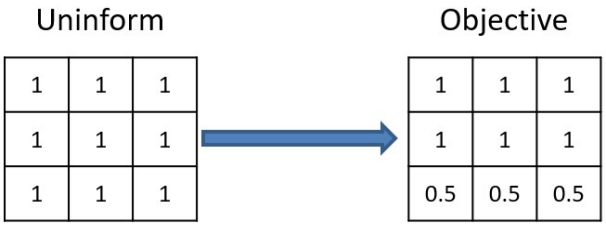


Figure 3.35: Relative resistance (design parameters) initial and objective values

The system must be able to train itself, starting from the uniform flow layup shown on the left, and approximate the objective matrix on the right, which is unknown during the process. However, we will use the objective matrix for comparison purposes. It is important to note that this objective matrix is based on the physically designed layup, assuming that adding two layers of breather material reduces resistance by exactly 50 percent. While this assumption may not be entirely accurate, the uniform setup, physically designed model, and physically trained model will ultimately be compared with experimental results.

### 3.11.4 Cost function and training procedure for physical parameter training

With the training parameters determined, the next step is to define the cost function for the training process. The cost function represents the error between the function's output and the actual data. It can be expressed as the equation

$$y = \frac{1}{M}\Sigma_{m=1}^{M}(\frac{1}{N}\Sigma_{i=1}^{N}(Flow_{calc}(x, y, rp, x_p, y_p) - f_i)^2)_m \tag{3.45}$$

where, $M$ represents the data size, and $m$ is the data index. $N$ is the total number of vacuum ports, which in our case study is six, and $i$ is the vacuum port index. The term $f_i$ denotes the normalized experimental flow rate at the $i_{th}$ port, representing the output to be predicted. Unlike typical machine learning models or neural networks, there is no direct function to calculate outputs (flow rates) based on inputs. The inputs in our problem include the overall layup geometry, vacuum port locations ($x_p$ and $y_p$), leakage location at data index $m$ which is represented by $x, y$, and the training parameter—the relative permeability matrix shown by $rp$.

Unlike a typical machine learning model, an explicit function does not exist to calculate flow rates from the inputs. Instead, a simulation must be used to calculate flow rates as a function of the inputs. This simulation, termed flow rate calculation, is denoted as $Flow_{calc}(input)$ in Eq. 3.45. The $Flow_{calc}$ function takes the leakage location $m$ and other geometrical inputs to calculate the flow rate at each vacuum port. The goal is to train this function to find the best relative permeability ($rp$) that minimizes the error, denoted by $y$.

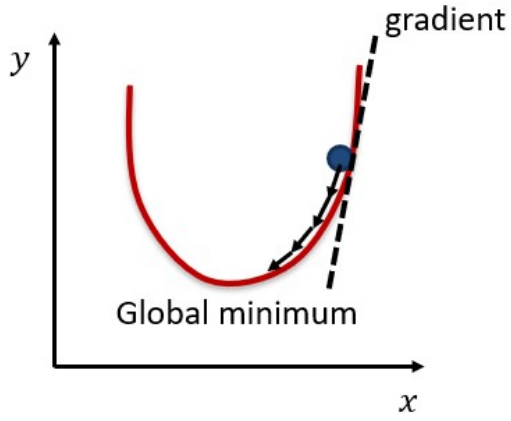


Figure 3.36: Gradient descent optimization algorithm

To minimize the error, a typical optimization algorithm such as gradient descent could be used. However, a few adjustments are required. Figure 3.36 shows a simplified process of gradient descent for minimization when having one design parameter denoted by $x$.

In gradient descent, the gradient of the error as the function of the training parameter will be calculated and the training parameters will be updated with the learning rate $\alpha$ as denoted in the equation

$$rp_j(new) = rp_j(old) - \alpha \frac{\partial}{\partial rp_j} y(rp)\,. \tag{3.46}$$

Note that $j$ is the index number of training parameters. In the setup shown in Figure 3.34, there are a total of nine different zones with different relative permeability, making a total index $j$ to nine. In our problem, the partial derivative of the cost function cannot be explicitly calculated due to the lack of a direct function. However, it could be calculated based on the definition of the derivative. The derivative,

$$\frac{\partial}{\partial rp_j} y(rp) = \frac{y(rp_j + \Delta rp_j) - y(rp_j)}{\Delta rp_j} \tag{3.47}$$

could be calculated numerically by applying a small change in the relative permeability $\Delta rp_j$. This process must be repeated for all of the training parameters (for $j = 1, 2, ..., 9$) and updated simultaneously. With enough number of iterations, the relative permeability will be trained to achieve the lowest error based on the provided data.

### 3.11.5 Challenges in physical parameter model training

Even though it is possible to train physical parameters of the model with this method, there are some challenges associated with it. The first challenge is determining the number of training parameters. For instance, considering the layup with a diagonal non-uniform permeability shown in Figure 3.37, which is similarly divided into nine smaller regions for training.

Unlike the previous example depicted in Figure 3.34, each smaller region in this case does not have uniform permeability. The regions in the top left, bottom right, and center contain both low and high permeability. However, when training the model, the trained relative permeability of these regions will reflect a compromise between the lower and higher permeabilities. Although the trained model will perform better

than a model with uniform permeability, it remains suboptimal. The only solution is to define more regions, thus increasing the number of design parameters. On the other hand, there may not be enough data to properly train a model with more parameters. Given the nature of this case study, achieving an ideal model through training might not be feasible.

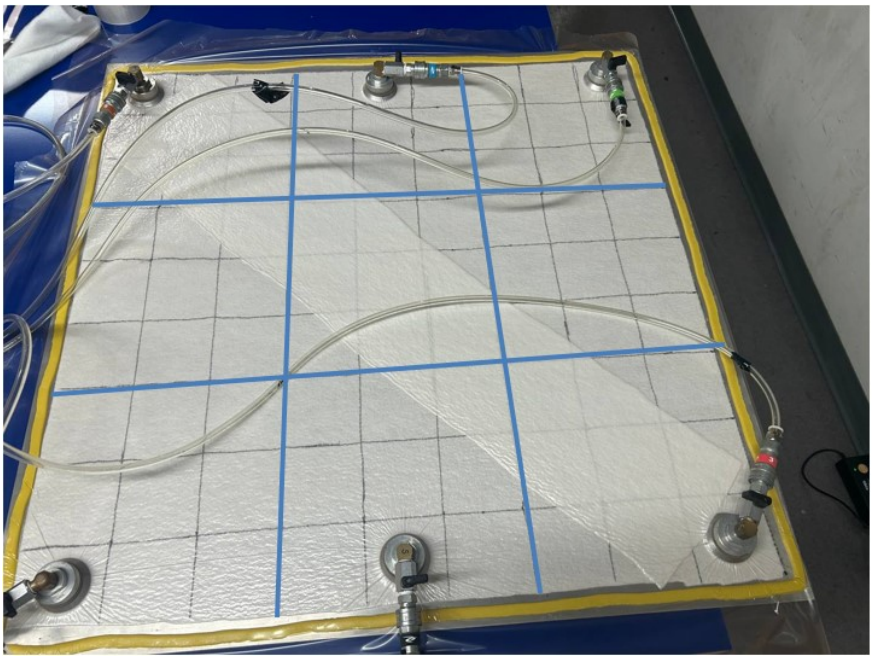

Figure 3.37: Layup with a diagonal non-uniform permeability

Training physical parameters using this framework is not limited to relative permeability; it can be applied to any physical parameter, provided it is properly defined within the flow rate calculation function in Eq. 3.45. However, the challenge lies in the fact that the model is restricted to training only the defined parameters. The algorithm minimizes the cost function error without accounting for other potential sources of error.

For example, consider a scenario where the connecting pipeline for one of the vacuum ports is significantly longer, resulting in lower flow rates at the corresponding sensor. The algorithm, which is focused on training relative permeability, will reduce the permeability of the region near that vacuum port to minimize the error. While this reduction decreases the error, it may not be entirely valid. If the same layup, designed for a specific part, is used with a different pipeline configuration, the system could experience increased error and reduced generalizability. To address this issue, it

is essential to ensure that the other aspects of the model, which are not being trained, are well-designed and accurately represent the system.

To conclude, the methodology outlined in this chapter provides a comprehensive framework for simulation and data synthesis, different training models and algorithms and optimizing the physical parameters of the model. By carefully addressing the challenges of data limitations, and model design, the approaches ensure a balance between accuracy and generalizability. Different case studies and scenarios, including multiple leakages, non-symmetrical geometries and non-uniform permeabilities were considered throughout the process. With these foundational steps in place, the next chapter will present the results and performance analysis of the trained models and simulation, offering insights into their practical application and effectiveness.

# Chapter 4

# Results

This chapter presents the results obtained by applying the methodology described in the previous chapter. Key findings are highlighted, beginning with simulation results across various geometries and configurations. The simulations range from uniform layups to more complex non-uniform setups and case studies involving multiple leakages. Following this, the model training results are discussed, focusing on the performance of the trained models compared to baseline scenarios. The analysis includes both qualitative and quantitative assessments, demonstrating the accuracy, robustness, and potential limitations of the proposed approaches. This chapter also covers training results using different models and methodologies for cases with single and multiple leakages. Finally, various leakage localization techniques, based on measuring volumetric flow rates, are compared with each other. These results offer critical insights into the models' effectiveness and their improvements over state-of-the-art methods.

## 4.1 Simulation results

The importance of process simulation was previously discussed, emphasizing that a well-performing model could be used for data synthesis in training applications. This model must accurately calculate volumetric flow rates based on the layup geometry, configuration, and leakage locations. In this section, we present the results from simulating the process using the electrical-circuit analogy method proposed in Section 3.5.2. Initially, the results for uniform setups where the permeability remains constant throughout the layup will be discussed. Afterwards, more complicated layups such

as layups with non-uniform permeability and cases where multiple leakages exist in the layup will be discussed.

### 4.1.1 Simulation results for uniform setups

Figure 4.1 shows the distribution of error percentages for a 90 cm square-shaped breather. This figure illustrates the percentage error in the simulation when a leakage occurs at any of the nodes marked by squares. The normalized currents from the simulation were compared to the normalized volumetric flow rates measured experimentally at each leakage location, as defined in Eq. 3.6. This comparison was performed for every node, excluding nodes extremely close to the vacuum ports, covering a total of 79 different locations on the layup.

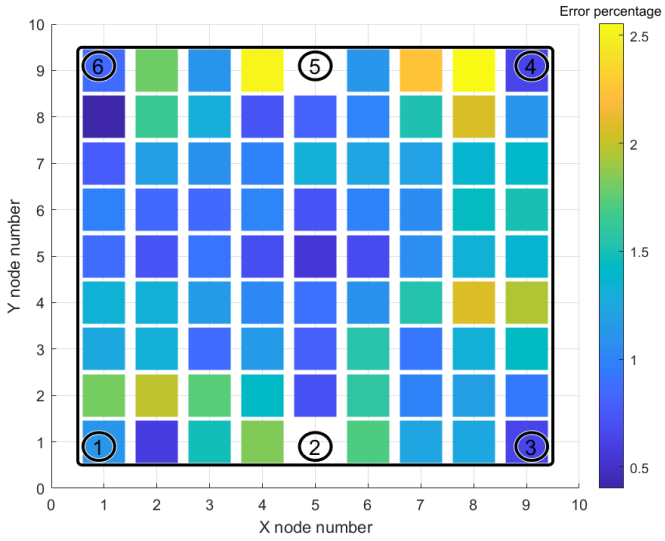


Figure 4.1: Analogy error distribution for a 90-cm square shape layup

The average simulation error across the entire layup is 1.22 %, with an RMS error of 1.30 %, indicating a good consistency and low variance in the predictions. These results suggest that using the electrical-circuit analogy for simulation is promising, at least for setups with uniform permeability and a single breather layer on a flat mold. In this case study, the maximum error occurs at the node with coordinates (X,Y) = (1,4), with an error of approximately 2.5 %. This node is notably close to vacuum port five, resulting in a high flow rate due to the short distance and low viscous resistance. Consequently, the volumetric flow rate at this node exceeds 1 L/min,

surpassing the sensor's upper limit and compromising the reliability of experimental flow rate measurements.

In addition, based on the error distribution illustrated in the Figure 4.1, areas near vacuum ports or boundaries tend to exhibit slightly higher errors. Besides the sensor's limitations, another source of discrepancy between experiments and simulations is the need for a higher number of electrical resistors—or, in other words, a finer mesh—to improve modeling accuracy. This is because the simulation approximates two-dimensional viscous resistance with a limited number of one-dimensional resistors, arranged in a way to recreate the 2D geometry. As the distance decreases, fewer 1D elements are used, which can reduce accuracy.

The simulation must work for other geometries as well and specially when there are leaks on the boundary. Improper bonds between the sealant tape, and the bag or tool, can cause leaks on the boundaries. To evaluate how the proposed method of localization can interpret boundary leaks, an octagon-shaped breather is created and previously represented in Figure 3.9. Leakages can be generated on different sides of this octagon. The procedure of the experiments utilizes the same framework used in previous setups. Leakages are introduced at different locations on the boundary and the measured flowrates have been recorded. Once the physical experiments are complete, the electrical-circuit model is created, and high voltages are introduced at the equivalent circuit locations corresponding to the leaks. The experimental results are then compared with the circuit-analogy outputs, and errors are calculated. The error percentage results are shown in Figure 4.2.

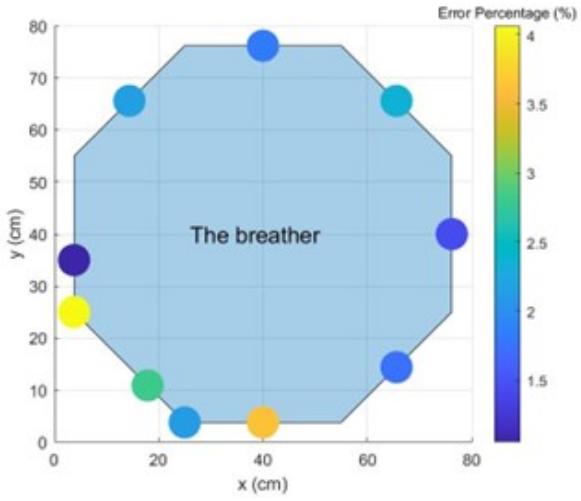


Figure 4.2: Octagonal-shaped breather for the edge (boundary) leakage experiment

There are a total of 10 different experiments and the errors in this case is also acceptable and below 4 %. The leak samples has been taken from the middle and end points of the sides of the layup. This experiments builds up toward a successful localization on the boundaries of the vacuum bags which are more susceptible to leaks. Results in this section indicates that the proposed geometry works well with uniform layups with different geometries. simulations and experiments were done on a rectangular layup as well as an octagonal shaped layup. In the next part, the simulation results for more complicated layups with non-uniform permeability will be discussed.

### 4.1.2 Simulation results for layups with non-uniform permeability

Figure 3.10 illustrates the experimental design in which specific areas of the breather are removed to create a non-uniform media. Since these subtracted regions are not part of the layup, leaks cannot form in these areas. Consequently, even air cannot travel through these subtracted regions, forcing it to flow around them. This results in significant changes in flow rates, rendering previous uniform simulations ineffective. In the electrical circuit model, resistors corresponding to these locations can either be removed or assigned a very high value to simulate reduced air flow. Figure 4.3 presents the error results for four different setups depicted in Figure 3.10. The results indicate that for complex geometries with subtracted areas in the breather, the maximum error remains below 5% in critical locations, such as near the vacuum ports or along boundary-adjacent areas. The average error for these case studies also remain below 2% which is acceptable.

The examples in Figure 4.3 examine cases where the relative permeability decreases in specific regions. Conversely, relative permeability can be increased by adding layers of breather material, as illustrated in Figure 3.11. There is an extra layer of breather in the diagonal of the figure in the left and an extra layer of breather in the bottom part of the figure in the right-hand side and the air could travel more easily in these areas. The increase in air flow occurs because additional breather layers create a more porous medium for air flow. These case studies are simulated by decreasing the resistance in the corresponding regions to represent higher permeability. In regions with two layers of breather material, the resistance values can be reduced by 50%. The simulation results are presented in Figure 4.4 for a total of 74

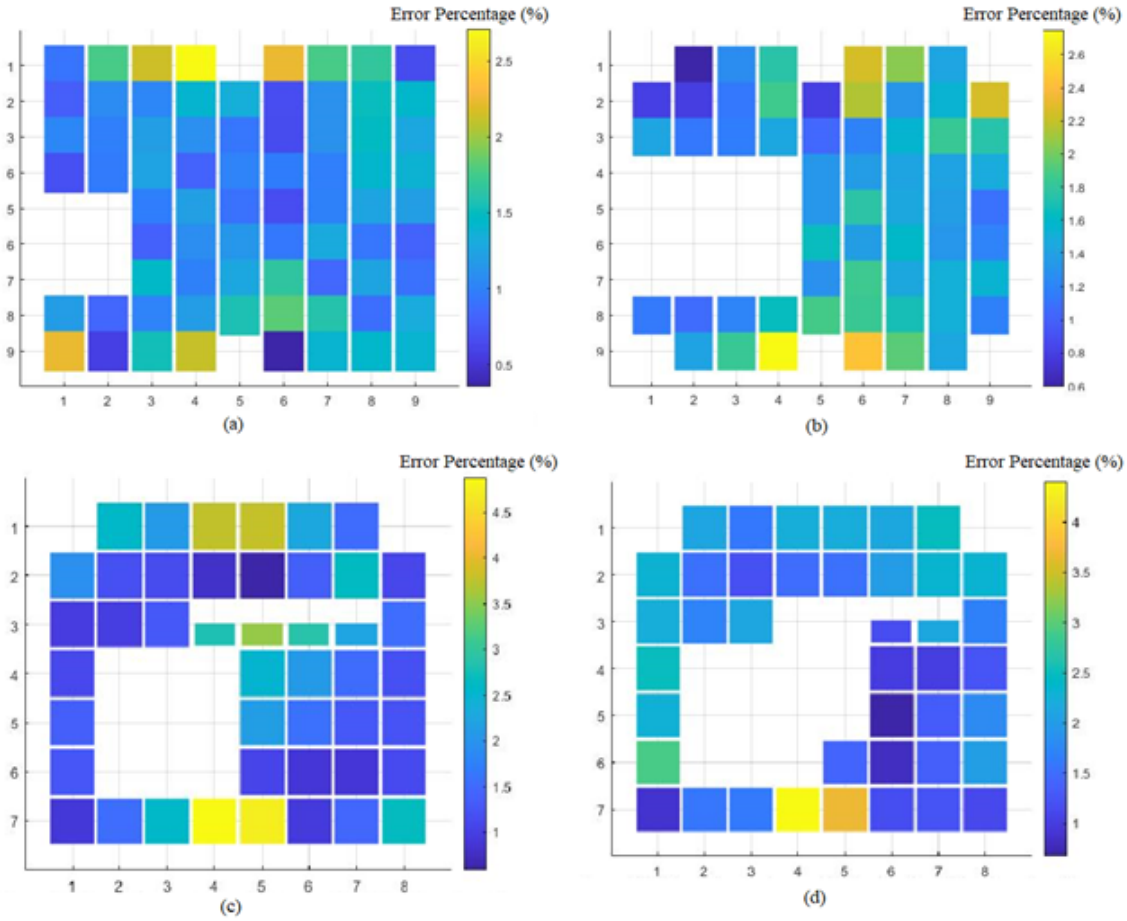


Figure 4.3: Error distribution for the analogy with non-uniform permeability created by deducted breathers; (a) 90 cm × 90 cm breather with an asymmetry 20 cm × 30 cm subtracted area. (b) 90 cm × 90 cm breather with a 40 cm × 40 cm subtracted area (c) and (d) 82 cm × 73 cm breather with complex subtracted areas

different samples.

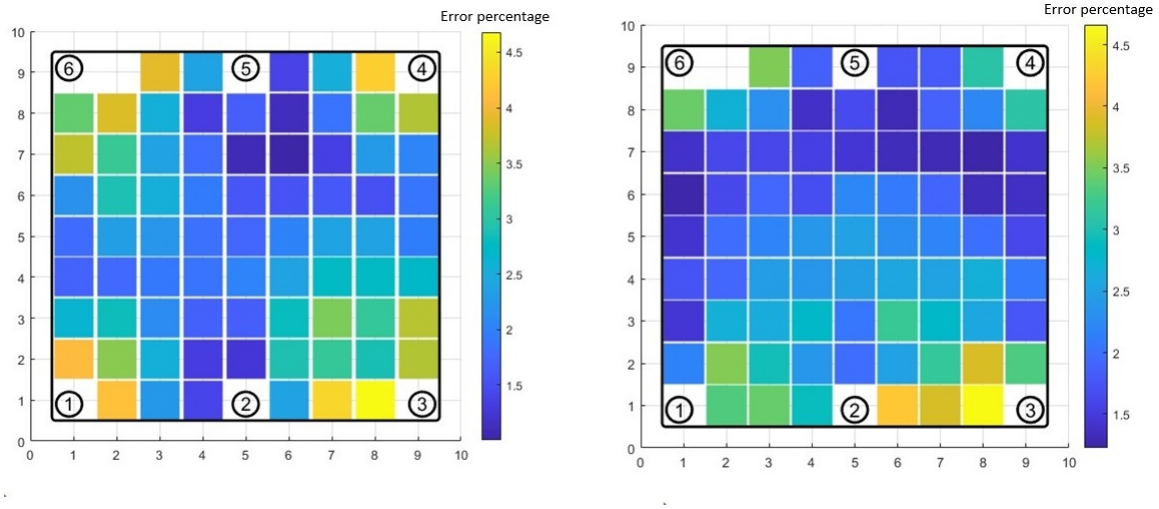


Figure 4.4: Error distribution for the layup analogy with higher permeability in specific areas: (a) higher permeability in the diagonal region of the layup on the left figure and (b) higher permeability in the bottom region of the layup on the right figure.

The average error for the figure on the left, where higher permeability exists along the diagonal of the layup, is 2.44%. Regions with the highest error appear around the corners and near the vacuum ports, with a maximum error of 4.5%. For the figure on the right, where higher permeability is present in the bottom region of the layup, the average error is 2.30%. In both cases, the highest error does not exceed 4.5%, and the simulation results fall within an acceptable range. Thus, it can be concluded that the simulation accurately models cases with regions of varying permeability in the layup, provided these regions are known and their resistance values can be manually adjusted.

Another method to create non-uniform permeability is by using angle iron sections within the vacuum bag. As shown in Figure 3.12 these angle irons form channels for airflow that provide a pathway with significantly lower resistance compared to the breather material. Figure 3.13 illustrates the setup with three angle irons placed horizontally on the layup and beneath the vacuum bag. Similar to previous case studies, the proposed simulation should accurately address this scenario with an appropriate electrical circuit design. Figure 4.5 presents the error distribution results for the setup with three angle irons under the vacuum bag. In this case, the average error is 2.19 %, the RMS error is 2.39 %, and the maximum error at the most critical location,

near vacuum port six, is below 4.5 %.

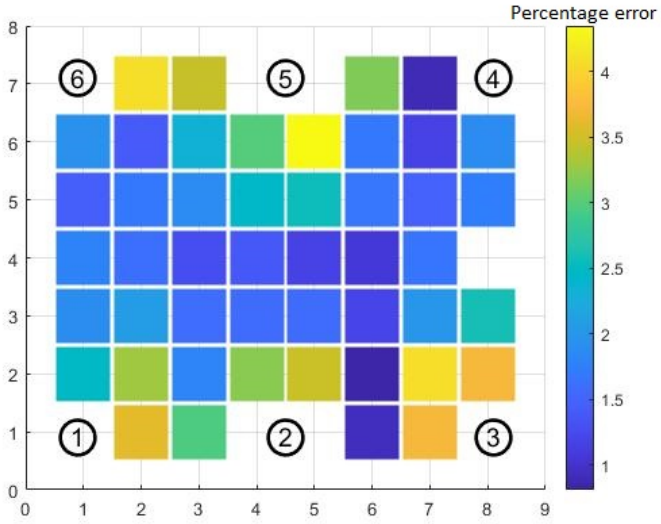


Figure 4.5: Percentage error distribution for horizontal angle iron

The angle iron creates an open air channel with significantly lower resistance compared to the breather material. Therefore, in the electrical circuit model, the resistor value corresponding to the angle iron's location should be reduced to simulate this low-resistance pathway. The main challenge in adjusting the resistor values lies in accurately determining the appropriate resistance to model the angle iron. The following approach addresses this challenge: first, the breather is assigned a baseline resistance of 20 Ω/cm, representing the uniform resistance across the breather area. Second, the low-resistance regions are modeled, noting that resistance depends on both the cross-sectional area of the free space created by the angle iron and the breather's porosity. Since the angle iron provides a direct air channel with minimal viscous resistance, the resistance for the angle iron regions should be considerably lower. To identify an accurate resistance value, three values—2, 0.2, and 0.04 Ω/cm—were modeled and simulated. The average and RMS errors across the breather are summarized in Table 4.1.

Table 4.1: Simulation results of different resistance values for angle irons

| Angle iron resistance value (Ω/cm) | Average error (%) | RMS error (%) |
|---|---|---|
| 2 | 3.08 | 3.25 |
| 0.2 | 2.31 | 2.5 |
| 0.04 | 2.19 | 2.39 |

As shown in Table 4.1, the total error decreases significantly when the resistance is reduced from 2 Ω/cm to 0.2 Ω/cm. However, comparing the error values between 0.2 and 0.04 Ω/cm shows a negligible reduction in error. This suggests that the angle iron model is approximately 100 times more permeable than the N10 breather. Figure 4.5 displays the error distribution between the experiment and the simulation with a resistance of 0.2 Ω/cm.

Additional setups using vertically placed angle irons beneath the vacuum bag were also considered, as shown in Figure 3.14. Figure 4.6 presents the simulation results for setups with five and three angle irons, respectively. The average error for these setups is 2.33% and 2.76%, with the maximum error remaining below 5% in the most critical locations.

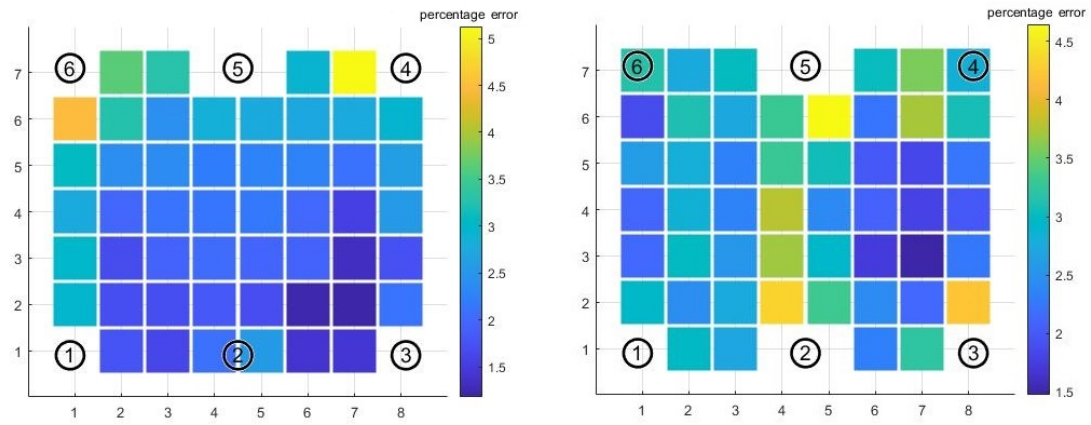


Figure 4.6: Percentage error distribution where (a) five and (b) three vertical angle iron were used in the layup

The previous results indicate a strong correspondence between the proposed simulation model and the experimental outcomes. However, there is potential to enhance the model by fine-tuning certain parameters.

In the earlier model, each 5 cm segment of angle iron was represented as a 2 Ω resistor, connected only to the adjacent resistor for the next 5 cm segment of angle iron (bottom image in Figure 4.7. This setup restricts current/air from flowing from the breather to the angle iron channel, except when entering from the ends of the angle iron. However, in reality, air can enter the channel along the entire length of the angle iron, including at points far from the channel's openings. To more accurately reflect these physical conditions, the resistors representing the angle iron should be interconnected with the resistors representing the breather, similar to the typical

resistor network (top image in Figure 4.7).

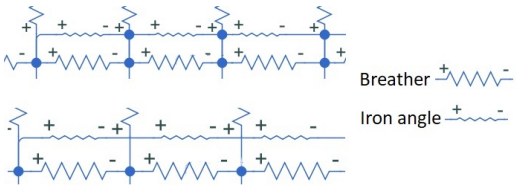


Figure 4.7: Original (bottom) and refined (top) circuit analogy - angle iron resistors are interconnected to the breather resistors in the refined model.

The refined simulation results are shown in Figure 4.8. The refined model reduced error values as follows, the average error decreased to 1.53 percent, the RMS error was reduced to 1.65 percent, and the maximum error was below 4%.

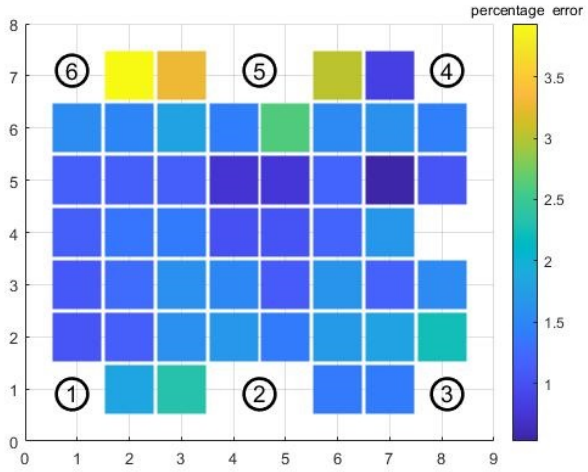


Figure 4.8: Percentage error distribution for horizontal angle iron refined model

### 4.1.3 Simulation results for multiple leakages

The problem of leakage detection can be excruciatingly complicated as an unknown number of leaks with unspecified locations can exist in the system. This problem should be considered when training the system. Therefore, it is imperative that the proposed simulation functions properly when there are simultaneous leakages in the system. The prediction of the location of multiple leakages is significantly more complicated than the single leakage scenarios.

In order to evaluate the reliability and validity of the simulation model, experiments have been conducted for multi leakage scenarios. Same setup in Figure 3.8 has

been used in the experiment but this time, two random and simultaneous leakages are created on the vacuum bag. We have repeated this process two times in order to have two sets of flowrate data resulted from having two simultaneous leaks.

Based on the analogy proposed in Section 3.5.2 the experiment can be modeled using an electric circuit. Voltages in any electrical circuit, provided only linear components exist in the circuit, follow superposition rule. Meaning that the effects of two different voltages could added up as if there are two separate systems. By performing these experiments, we want to make sure the underlying physics of the procedure also follows the superposition rule. To simulate a multiple leakage scenario, two corresponding nodes must be set to high voltage. Figure 4.9 shows the results for the multiple leakage experiment for two different experimental sets. For all of the , the average error is 1.96% and the root-mean-square error is 2.2%. Additionally, the highest error is about 4% which occurs in the vicinity of the vacuum ports. These areas are considered as high-error regions.

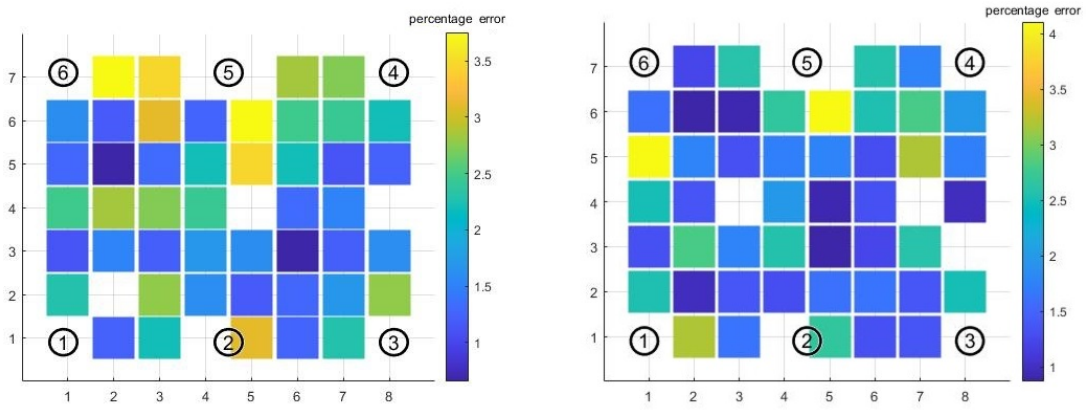


Figure 4.9: Percentage error distribution for two simultaneous leakages

The results suggest that the proposed analogy is able to model cases where two leaks exist in the layup and it follows the superposition rule.

### 4.1.4 Overall simulation results

In this section, results from the proposed simulation method based on an electrical circuit analogy was discussed. The results were verified using experimental data from the real-world setup. The case studies include uniform layups in different geometrical shapes as well as setups with non-uniform permeability in various areas. Eventually, the method was verified when there are more than one leakage on the layup. The

results performed with good accuracy and low error in all of the case studies. However, a good modeling and a proper prior knowledge on the layup is required for a good simulation.

The proposed methodology is rapid and robust and can be used as a data synthesis tool for further applications in machine learning models. In the next section, the verified simulation has been used for data generation and the generated data has been used for model training.

## 4.2 Training results for single leakage localization

The proposed analogy has undergone validation for various configurations, establishing its reliability as a data generation tool. Once a sufficient amount of data has been synthesized, a regression model can be trained as discussed in Section 3.7. The training dataset consists of 554 samples generated from eight distinct configurations. To assess the model's generalizability, a new dataset containing 73 samples is generated from a completely novel configuration, to which the machine learning agent has never been exposed.

Each data entry includes the flowrates of the six vacuum ports (features) and the coordinates of the leaks (labels). To simplify implementation, the polar coordinate system $(r - \phi)$ is employed. Port six (bottom left corner) serves as the origin of this coordinate system, and each point on the breather can be described by two components: $r$ (in centimeters) and $\phi$ (in degrees) shown in Figure 4.10.

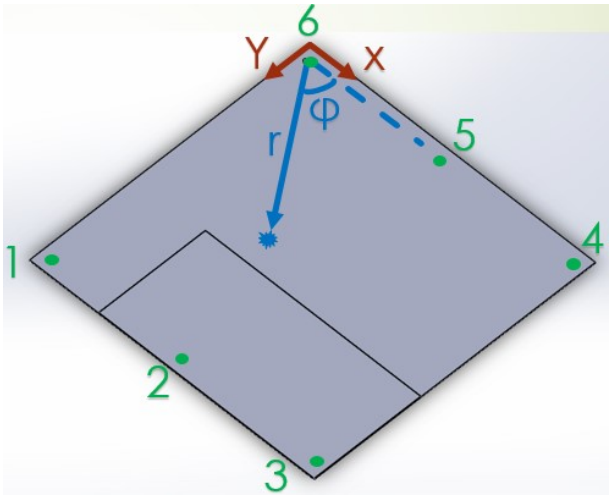


Figure 4.10: Coordinate system for describing the location of the leakages

Regression model training and test errors are represented in the following.

### 4.2.1 Linear regression model

As shown in Table 4.2, the prediction error for the distance ($r$) is below 5 cm for both validation and test datasets. The simplicity and low variance of the model contribute to even lower test errors, with distances predicted within 4 cm. However, due to its simplistic nature, the linear model falls short in accurately predicting the angle ($\phi$) with an acceptable level of error.

Table 4.2: Root-Mean-Square error for linear regression model

| | $r$(cm) RMS error | | $\phi(^\circ)$ RMS error | |
|---|---|---|---|---|
| Linear model | Validation | Test | Validation | Test |
| Linear | 5.4 | 3.25 | 14.85 | 13.22 |
| Interaction linear | 4.2 | 2.5 | 11.15 | 14.38 |
| Robust linear | 5.4 | 2.39 | 19.19 | 20.21 |
| Stepwise linear | 4.3 | 2.39 | 11.36 | 13.50 |

### 4.2.2 Regression Tree

Three different regression trees with leaf sizes of 4 (fine), 12 (medium) and 36 (coarse) are designed and trained for the problem. Results in Table 4.3 show as the leaf size increases, the variance increases and leads to overfitting. Based on the results, increasing the leaf size has a negative effect on the accuracy of the model.

Table 4.3: Root-Mean-Square error for regression tree

| | $r$(cm) RMS error | | $\phi(^\circ)$ RMS error | |
|---|---|---|---|---|
| Tree type | Validation | Test | Validation | Test |
| Fine tree | 4.4 | 4.1 | 11.52 | 6.64 |
| Medium tree | 4.8 | 4.7 | 12.27 | 5.91 |
| Coarse tree | 7.1 | 5.1 | 13.76 | 9.67 |

The regression tree model is accurate in predicting the angle ($\phi$) even in the unseen test dataset. The model has an adequate level of generalization. This is supported by the test error being lower than the validation set error.

### 4.2.3 SVM Regression

Table 4.4 shows the training results for SVM model with different kernels. Based on the results, the linear kernel has a high bias and is unsuitable for modelling the problem, especially in the prediction of the angle. In contrast, the cubic kernel tends to overfit to the data because it is too complicated and has a high test error. The quadratic model, in essence, has a proper trade-off between adequately representing the problem and not overfitting to the test data. This kernel has about 10° of test error for the angles.

Table 4.4: Root-Mean-Square error for Support Vector Machine (SVM)

| | $r$(cm) RMS error | | $\phi(^\circ)$ RMS error | |
|---|---|---|---|---|
| SVM model | Validation | Test | Validation | Test |
| Linear | 5.4 | 6.8 | 16.18 | 15.26 |
| Quadratic | 4.2 | 5.1 | 11.59 | 9.81 |
| Cubic | 3.1 | 4.3 | 9.09 | 17.61 |
| Fine Gaussian | 4.7 | 5.9 | 9.92 | 8.58 |
| Medium Gaussian | 3.3 | 3.8 | 10.11 | 10.64 |
| Coarse Gaussian | 4.5 | 5.3 | 13.81 | 11.90 |

The fine Gaussian kernel shows the best prediction over validation and test sets among all of the SVM kernels. This kernel results in an error of less than 10° for angle predictions. The medium Gaussian kernel has the best prediction over distance ($r$). The coarse Gaussian kernel is prone to overfitting for this application because it is too complicated for the current model and datasets.

### 4.2.4 Gaussian progress regression (GPR)

Table 4.5 shows the training results using Gaussian process regression (GPR) with different kernels.

Table 4.5: Root-Mean-Square Error for Gaussian Progress Regression (GPR)

| | $r$(cm) RMS error | | $\phi(^\circ)$ RMS error | |
|---|---|---|---|---|
| GPR model | Validation | Test | Validation | Test |
| Rational quadratic | 2.3 | 4.4 | 9.02 | 10.07 |
| Squared exponential | 2.3 | 4.6 | 9.03 | 12.12 |
| Matern 5/2 | 2.3 | 4.3 | 8.45 | 11.42 |
| Exponential | 2.6 | 3.4 | 8.6 | 10.36 |

Based on the results, almost all of the kernels converge towards similar error results. The $r$ validation error for GPR is the smallest among all other models tested. Unfortunately, the model is overfitting, this is supported by the fact that the test error is roughly two times that of the validation error. The GPR models work fine on predicting angle because the angle response requires a more complicated model than the model for predicting distance($r$). The Exponential and Rational Quadratic kernels, which are among the most simple kernels in GPR, perform better on the angle predictions.

### 4.2.5 Neural network

The results of the neural network model are show in Table 4.6.

Table 4.6: Root-Mean-Square error for Neural Network(NN)

| | $r$(cm) RMS error | | $\phi(^\circ)$ RMS error | |
|---|---|---|---|---|
| NN model | Validation | Test | Validation | Test |
| Narrow | 3.4 | 4.1 | 11.68 | 13.89 |
| Medium | 2.7 | 4.7 | 9.15 | 12.20 |
| Wide | 2.5 | 4.8 | 9.03 | 8.68 |
| Bi-layered | 2.8 | 5.5 | 10.06 | 10.48 |
| Tri-layered | 2.7 | 4.2 | 9.00 | 9.03 |

Almost all of the networks except the narrow one-hidden layered network are susceptible to overfitting as they have quite low validation error and higher test error. The narrow model which has the least neurons in the only existing hidden layer performs best in predicting distance ($r$). However, this network performs poorly on the angle prediction as this output is more complicated compared to distance. All other designed networks especially the wide and tri-layered models perform well for predicting the angle.

### 4.2.6 Localization performance on experimental setup

In all previous trainings, the synthesized data was used for training, validation and test. Even though the test dataset was new never exposed to the model during the training process, it is still required to test the localization method on experimental case studies. In other words, the model must be put on test in the real world. Figure 4.11 shows the localization results on real-world data. The figure shows the actual

location of leaks as well as the predicted locations on the 90cm×90cm layup. The model has been trained using the GPR model with an exponential kernel which was one the best performing models in this section.

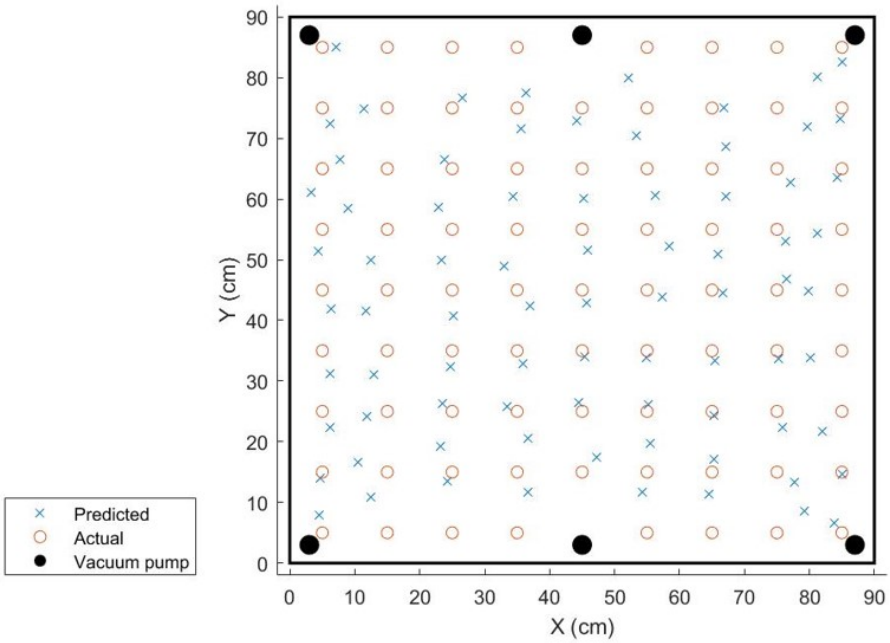


Figure 4.11: Single leakage localization performance of a trained model using exponential GPR for real-world data

The errors on $r$-$\phi$ components were added together to make one single average error which is equal to 4.2cm of leakage location prediction error. This value is dependent on the layup size and must always be expressed with the layup size. This means that the user must search in a circle with a radius of 4.2cm for a potential leakage instead of the whole layup area. This will reduce the search area by 97% and user has to search for only 3% of the layup. We will call this by "search area shrinkage" throughout this thesis.

## 4.3 Single/multiple leakage classification training results

This section discusses the findings from Section 3.8, where the system must accurately determine the number of leaks present by interpreting flow rate inputs. We have fo-

cused on cases where there are specifically two leaks in the system, excluding systems with more leaks from our study. The simulation results in Section 4.1.3 confirmed that the proposed approach is effective for cases with two simultaneous leaks and is a viable method for generating the data needed for model training. Using this approach, 200 samples of the second class (two leaks) were synthesized, and these, along with 554 previously generated samples from the first class (one leak), were used to train various classification models. Model performance was then evaluated on a new, unseen dataset containing 72 samples from class 1 and 50 samples from class 2. Table 4.7 presents the classification results from different models, including SVM with various kernels, three distinct neural networks, and a fine and weighted K-nearest neighbors (KNN) model, which uses proximity to predict classifications [172]. The models used for training are pretty much similar to the models introduced in Section 3.7.2, with the difference that the outputs are class labels instead of a continuous response as we are dealing with classification in this part.

Table 4.7: Training results for single and multiple leakage classification

| | | Accuracy (%) | |
|---|---|---|---|
| Model | Kernel/Type | Validation | Test |
| SVM | Cubic | 87 | 44.3 |
| | Quadratic | 84.5 | 60.7 |
| | Fine Gaussian | 82.6 | 68.9 |
| | Medium Gaussian | 81.5 | 65.6 |
| NN | Narrow | 84.5 | 55.7 |
| | Medium | 85.1 | 54.1 |
| | Wide | 85 | 58.2 |
| KNN | Fine | 80.7 | 70.5 |
| | Weighted | 81.3 | 72.1 |

The results in the table indicate that many of the trained models exhibit signs of overfitting. SVM models with cubic and quadratic kernels, as well as all of the neural networks, demonstrate high validation accuracy on the training dataset but have limited generalizability, with substantially lower accuracy on the test dataset. For models with only two output classes, achieving accuracy in the range of 40-60% is inadequate. However, models with fewer features, such as SVMs with fine and medium Gaussian kernels and KNN models, show more promise, achieving test accuracy rates of 65-70%.

A primary reason for this overfitting and high test error may be the unbalanced

dataset used for training, where class 1 contains more than twice the number of samples as class 2. Due to this unbalanced dataset, it is better to measure the performance of the trained models using the F-score factor. Table 4.8 shows the confusion matrix for the wide neural network model on both validation and test datasets. A confusion matrix is a table that is commonly used to evaluate the performance of a classification algorithm. It compares the predicted classifications of a model with the actual classifications from a dataset, giving insight into the accuracy and types of errors the model is making. Each row of the matrix represents the instances of an actual class, while each column represents the instances of a predicted class. Diagonal entries represent correct predictions for each class, while off-diagonal entries show where the model misclassified instances as other classes.

Table 4.8: Wide neural network confusion matrix for validation and test datasets

| Wide NN confusion matrix | | Predicted class | | | Predicted class | |
|---|---|---|---|---|---|---|
| Dataset | True class | 1 | 2 | Dataset | 1 | 2 |
| Validation | 1 | 496 | 57 | Test | 33 | 39 |
| | 2 | 64 | 151 | | 12 | 38 |
| F-score | | 0.891 | | F-score | 0.569 | |

The wide neural network model achieves a high F-score on the validation set but a low F-score on the test dataset, indicating poor generalizability. While the model accurately predicted multiple leakage scenarios, it misclassified more than half of the single-leakage instances as multiple leakage. Table 4.9 presents the confusion matrix for the best-performing model, the weighted KNN.

Table 4.9: Weighted KNN confusion matrix for validation and test datasets

| Weighted KNN confusion matrix | | Predicted class | | | Predicted class | |
|---|---|---|---|---|---|---|
| Dataset | True class | 1 | 2 | Dataset | 1 | 2 |
| Validation | 1 | 529 | 24 | Test | 64 | 8 |
| | 2 | 107 | 108 | | 26 | 24 |
| F-score | | 0.890 | | F-score | 0.788 | |

The model shows a similar F-score on the validation set compared to the wide neural network model, but it outperforms it on the test set, as expected. However, the model misclassified half of the multiple leakage instances as single-leakage. The wide neural network model is more likely to classify new data as double-leakage, while the weighted KNN model tends to classify inputs as single-leakage. The misclassifications in the tables reflect a systematic error in the training process.

It is possible that imbalance of data hindered the training process. That could be a reason that the model is tend to misclassify new data systematically. To address this issue, additional samples with two leaks were synthesized, increasing the dataset size for class 2 from 200 to 500. Table 4.10 presents the training results after balancing the dataset to include 554 samples for class 1 and 500 for class 2. Since the dataset is balanced in this stage, there is no need to use F-score for performance assessment.

Table 4.10: Training results for single and multiple leakage classification for a more balanced dataset

| | Accuracy (%) | |
|---|---|---|
| Model | Validation | Test |
| Fine Gaussian SVM | 77.8 | 68 |
| Fine KNN | 79 | 73 |
| Weighted KNN | 78.8 | 71.3 |

The table results show that, while using a more balanced dataset for training provided a slight improvement in accuracy, this adjustment did not significantly enhance model performance. This suggests that data size or balance is not the primary source of error. Various machine learning models were tested, and although the best-performing models were selected, none could achieve accuracy beyond 75%. The issue likely stems from the nature of the problem itself. This aligns with findings in Section 3.6.1, which demonstrated that "as the distance between two leakages decreases, the flow rates become more similar, resembling the pattern of a single leakage between them." This similarity in flow rates can lead to overfitting and high variance errors, particularly when the training data includes instances where leaks are very close together.

In the class 2 dataset, where flow rate readings result from two simultaneous leaks, the minimum distance between leaks is 10 cm—a relatively short distance that likely contributes to prediction errors. The models were trained on data where leaks could be as close as 10 cm, resulting in flow rates that closely mimic single-leak scenarios, thereby reducing prediction accuracy. To test this hypothesis, the trained models were evaluated on a customized test set with leakages spaced 40 cm apart, rather than 10 cm. On this test set, model accuracy improved to 83%. Notably, the models were still trained on a dataset without distance constraints, allowing leaks as close as 10 cm. This suggests that to further enhance accuracy, it would be necessary to impose a minimum distance constraint between leaks in the training dataset.

### 4.3.1 Classification results with dataset constraint

Based on the previous findings, a minimum distance constraint of 35 cm between each two leaks were applied to the dataset. This includes all of data in training, validation and test sets. Table 4.11 shows the training results where a minimum distance exists between the leaks in the data.

Table 4.11: Training results for single and multiple leakage classification with a 35 cm minimum distance of leaks constraint

| | | Accuracy (%) | |
|---|---|---|---|
| Model | Kernel/Type | Validation | Test |
| SVM | Cubic | 93.5 | 77 |
| | Quadratic | 82 | 66.5 |
| | Fine Gaussian | 90.4 | 85.7 |
| | Medium Gaussian | 81.6 | 74.5 |
| NN | Narrow | 89.7 | 72.7 |
| | Bi-layered | 93.4 | 89.4 |
| | Wide | 95.1 | 85.7 |
| KNN | Fine | 94.9 | 91.3 |
| | Weighted | 93.4 | 90.1 |

The results show significant improvements. Although models like SVM with cubic and quadratic kernels and narrow neural networks still suffer from overfitting, most other models exhibit notable improvement, achieving accuracy in the 85-91% range. The fine KNN model achieves the best validation and test accuracy with minimal error.

These results suggest that, with the current features—six normalized flow rates—it is nearly impossible to train a model capable of accurately classifying leaks that are extremely close to one another. In addition to implementing a minimum distance constraint, increasing the amount or diversity of information (i.e., adding more features) could also enhance accuracy. The next section introduces a method for augmenting the system with additional features.

### 4.3.2 Increasing total features for improved accuracy

The objective is to introduce additional features to potentially improve prediction accuracy. Although adding more vacuum ports and sensors to the layup could achieve this, it is generally undesirable, as it reaches the system's practical limits and users typically prefer a consistent setup for their layups. The primary advantage of using

flow rates for leak detection and localization is that it enables automation without altering the layup configuration. Therefore, adding physical components to increase features is not ideal.

An effective way to increase features without modifying the setup is to selectively close some of the vacuum port valves. When a valve is closed, airflow through that port is blocked, requiring the outbound air to exit through alternative ports, thus generating new flow rate values independent of the original features. This is analogous to removing one of the grounds in an electrical circuit: a new equivalent resistance is created, resulting in new current distributions, as though the system's geometry had been altered.

We designed a simple experiment to validate this approach, conducting two sets of trials in the setup shown in Figure 4.12. In the first set, two simultaneous leaks were introduced at the locations indicated by green circles. The distance of these two leaks is 20 cm. After starting the vacuum pump, the flow rates were recorded. Next, one of the vacuum port valves was closed, and the flow rates from the five remaining ports were recorded. This process was repeated twice more, each time closing a different valve.

In the second set, the leaks were sealed, and the entire procedure was repeated with a single leak positioned between the previous two leaks, marked in red. This position was chosen as it would produce flow rates most similar to those from the two-leak scenario. After completing both sets of experiments, the flow rates were compared to assess the validity of the proposed feature enhancement method.

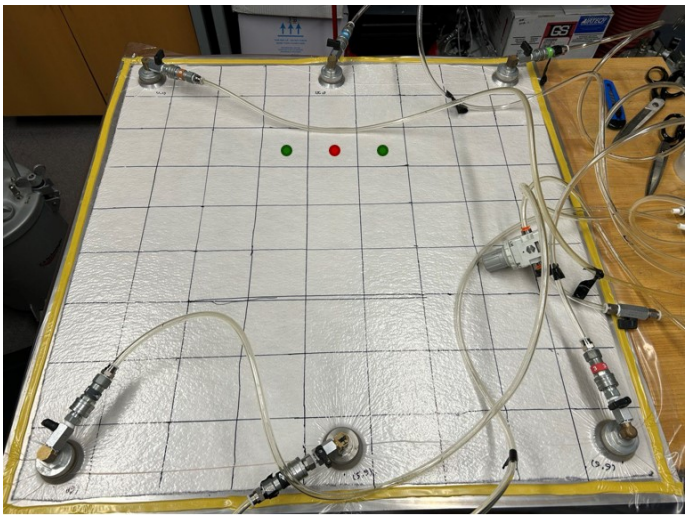

Figure 4.12: Single and multiple leakage test for feature increase hypothesis

Table 4.12 shows the comparison results for the scenarios. The columns $F1$ through $F6$ show the normalized flow rate values. Note that zero values in the table represents cases where the ports are closed.

Table 4.12: Similarity check for single and multiple leakage scenarios after feature increase

| Scenario | F1 | F2 | F3 | F4 | F5 | F6 | Sum |
|---|---|---|---|---|---|---|---|
| Two-leaks | 0.093<br>0.107<br>0<br>0 | 0.147<br>0.175<br>0.245<br>0 | 0.091<br>0.116<br>0<br>0 | 0.163<br>0<br>0.188<br>0.258 | 0.347<br>0.423<br>0.383<br>0.488 | 0.160<br>0.180<br>0.184<br>0.253 | 1 |
| One-leak | 0.092<br>0.104<br>0<br>0 | 0.147<br>0.174<br>0.244<br>0 | 0.090<br>0.114<br>0<br>0 | 0.154<br>0<br>0.179<br>0.249 | 0.366<br>0.439<br>0.402<br>0.506 | 0.151<br>0.168<br>0.176<br>0.245 | 1 |
| Difference | 0.001<br>0.003<br>0<br>0 | 0.000<br>0.001<br>0.001<br>0 | 0.001<br>0.002<br>0<br>0 | 0.009<br>0<br>0.009<br>0.009 | 0.019<br>0.016<br>0.019<br>0.018 | 0.008<br>0.011<br>0.009<br>0.009 | 0.039<br>0.032<br>0.037<br>0.037 |

The last rows show the differences between the two scenarios with various valves open. As indicated, the total flow rate difference between these two cases is 0.039, with only a 4% difference between them. This example illustrates the challenge of classification, particularly when leaks are close to each other, as their flow rates closely resemble a single-leak scenario. The results also show that, when some vacuum port valves are closed, flow rate differences persist. By altering the system structure through valve closures, independent features with minimal variance are generated, making them suitable for training purposes.

### 4.3.3 Training Results with Enhanced Feature Set

Table 4.13 presents the classification training results when using an enhanced feature set. The table compares results between training with 6 features versus 15 features. For the 6-feature setup, only the normalized flow rates from the six ports are used. In the 15-feature setup, after recording the initial flow rates, the valve corresponding to the highest flow rate is closed, and the flow rates from the remaining five ports are recorded. This process is repeated by closing the next highest-flow-rate port and recording the flow rates from the remaining four ports. The resulting 15 flow rate measurements are then normalized and combined as a single feature set for training.

The training dataset consists of 565 samples with approximately equal representation from each class. Around 20% of this set is set aside as a validation set. An additional unseen dataset of 140 samples is used as a test set to assess generalizability. Various SVM, KNN, and neural network models are employed for training. The minimum distance between each pair of leaks is 27 cm, which is a stricter constraint compared to the 35 cm minimum distance used previously, as shown in Table 4.11.

Table 4.13: Training results for single and multiple leakage classification with enhanced feature set

| | | 6 features accuracy (%) | | 15 features accuracy (%) | |
|---|---|---|---|---|---|
| Model | Kernel/Type | Validation | Test | Validation | Test |
| SVM | Cubic | 80.7 | 78.6 | 87.7 | 90 |
| | Quadratic | 71.5 | 70.7 | 81.1 | 75.7 |
| NN | Narrow | 80.6 | 70 | 85.1 | 83.6 |
| | Medium | 81.4 | 78.6 | 86.8 | 87.9 |
| | Wide | 83.3 | 83.6 | 85.1 | 85 |
| | Bi-layered | 84.5 | 86.4 | 82.3 | 85.7 |
| | Tri-layered | 81.8 | 82.1 | 81.9 | 85.7 |
| KNN | Fine | 72 | 70 | 79.2 | 77.1 |
| | Weighted | 71.4 | 73.6 | 77.4 | 74.3 |

In almost all models, using additional features significantly improves accuracy, as expected, since more features provide richer information during training. More complex models, such as Cubic SVM and bi-layered and tri-layered neural networks, show substantial improvement, as these models are better suited to handle larger feature sets. However, with the stricter minimum distance constraint, the KNN models do not perform as well as before. Overall, the results demonstrate that with an increased number of features, it is possible to accurately predict the number of leaks as high as 90 %, even when leak distances are as close as 27 cm. Table 4.14 shows the confusion matrix for the cubic SVM model using the enhanced feature set.

Table 4.14: Cubic SVM confusion matrix for validation and test datasets

| Cubic SVM confusion matrix | | Predicted class | | | Predicted class | |
|---|---|---|---|---|---|---|
| Dataset | True class | 1 | 2 | Dataset | 1 | 2 |
| Validation | 1 | 274 | 11 | Test | 69 | 1 |
| | 2 | 60 | 231 | | 13 | 57 |
| F-score | | 0.885 | | F-score | 0.908 | |

Since the dataset is balanced between the classes, the F-score values are approx-

imately equal to the accuracy shown in Table 4.13. The results indicate that most of the misclassified examples are false positives, where the true class was the second class (multiple leaks) but the predicted class was single-leak. This typically occurs when two leaks are close to each other, and their flow rates are similar, making it appear as if only one leak exists.

### 4.3.4 Classification results on experimental results

The next step is to test the trained models on experimental data. In the setup shown in Figure 4.13, a total of 30 experiments were conducted, including 22 single-leak and 8 double-leak scenarios. Both models trained with the normal and enhanced feature sets were used for leakage classification. The results indicate that the model achieved correct predictions in 26 out of 30 experiments, yielding an accuracy of 87%. When using the model trained with the enhanced feature set, accuracy improved further, with two of the four initial misclassifications corrected. In Figure 4.13, correct predictions are marked with green circles, incorrect predictions with red, and improved predictions with the enhanced feature set are highlighted in yellow.

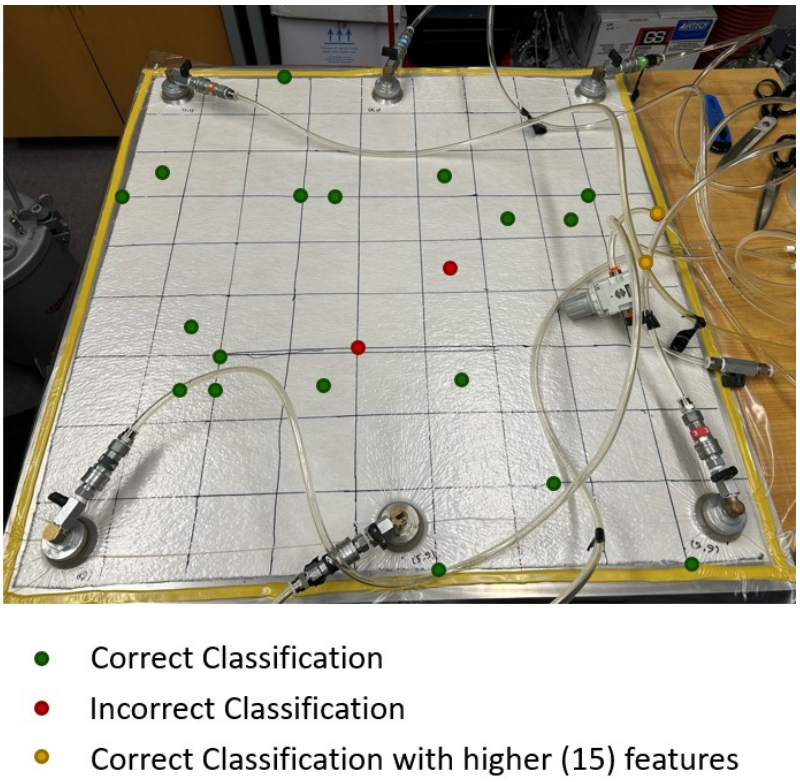


Figure 4.13: Single/multiple leakage classification results on experimental data

It is also notable that some leaks have appeared at the boundaries, where leaks are more likely to occur. Across experiments, the average distance between leakage pairs is 52 cm, while the average distance for misclassified cases is 34 cm. This further supports the observation that as leakage distance decreases, flow rates resemble those of single-leak scenarios.

In conclusion, the problem of classifying the number of leaks can be addressed by generating sufficient, accurately labeled data. Achievable accuracy rates, verified through simulations and experiments, are around 90%. Classifying the number of leaks is the first step toward comprehensive leak localization. If the classification algorithm identifies a single leak, the single-leak localization method introduced in Section 3.7 is applied. For cases with two leaks, a different algorithm—discussed in the following section—will be required to determine their locations.

## 4.4 Physics-informed grid neural network

The GNN framework was evaluated on a $90cm \times 90cm$ breather, and its performance was assessed. Three different frameworks were compared: the regression models from Section 3.7.2, the GNN model, and the PIGNN model, all tested under the same setup and using identical input data. To evaluate performance under data scarcity, only a limited dataset of 11 samples was used for training. The results, summarized in Table 4.15, demonstrate the models' accuracy when even a portion of this small dataset (about 55%) is used for training. The accuracy was estimated on unseen data from different locations.

Table 4.15: GNN, PI-GNN and regression model performance comparison

| Model | Data used (%) | Error (%) | Classification accuracy (%) | Average error (cm) | Probe area shrinkage (%) |
|---|---|---|---|---|---|
| GNN | 55 | 9.43 | 18.5 | 21.6 | 33.9 |
| | 100 | 6.95 | 34.5 | 14.1 | 22.2 |
| PI-GNN | 55 | 1.81 | 86.4 | 6.5 | 10.2 |
| | 100 | 1.20 | 95.1 | 5.5 | 8.6 |
| Regression | 55 | NA | NA | 22.2 | 34.9 |
| | 100 | NA | NA | 17.9 | 28.2 |

The GNN and previous regression methods perform similarly when a small amount of data is used, requiring the user to search within a relatively large area for the leakage. However, as the amount of training data increases, the GNN outperforms the

regression model. The PI-GNN demonstrates significant superiority over both methods by integrating physical knowledge with data-driven learning. By incorporating the geometry of the system, the GNN models surpass the regression model, which relies solely on data to train a black-box function.

Figure 4.14 illustrates the prediction results for three different leakage experiments using the PI-GNN model trained on a limited dataset. In the examples shown, the system accurately identifies the actual leakage location with the highest probability.

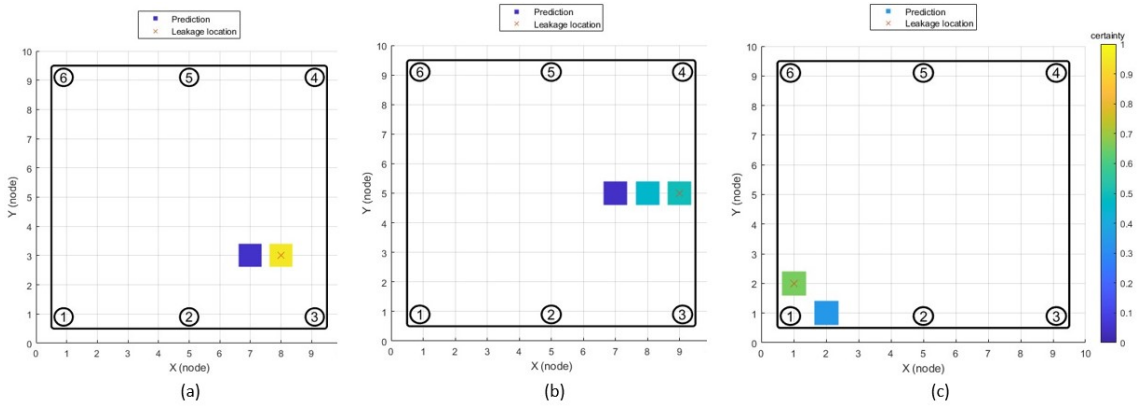


Figure 4.14: Prediction results of PIGNN model for three different leakage locations

Using more data for training, especially diverse data from various locations, significantly enhances the model's performance. This not only improves prediction accuracy but also increases the system's reliability and confidence. Figure 4.15 presents the prediction results for the same locations shown in Figure 4.14, this time with additional training data. In this case, the system's confidence in its predictions improves.

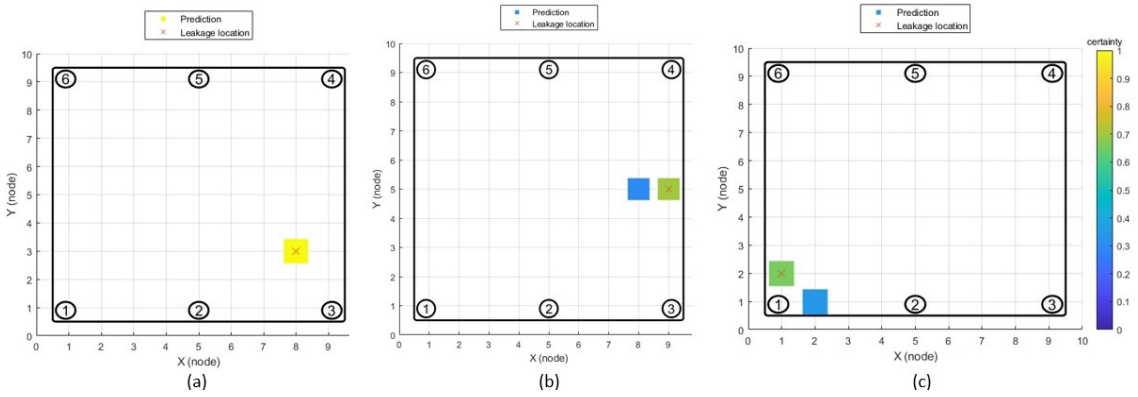


Figure 4.15: Prediction results of the PIGNN model trained on more data

The results further demonstrate the system's efficiency in terms of data require-

ments.

### 4.4.1 Results on more challenging layup with non-uniform permeability

The flexibility of the system within the GNN framework surpasses that of the regression method. This superiority was demonstrated through a challenging case study involving a non-uniform layup with angle irons. In this case, shown in Figure 3.13, the regions parallel to the length of the angle irons exhibit identical flowrate values, making accurate leakage localization particularly difficult. When applying regression training to this scenario, the model would likely predict one of these regions as the leakage location. However, due to the similarity in flowrates, such predictions are prone to inaccuracy. Increasing the training data does not resolve this issue, as the flowrates remain nearly identical along the angle iron length.

In contrast, the PIGNN model effectively addresses this challenge by predicting all possible leakage regions rather than generating incorrect results. As depicted in Figure 4.16, the system successfully identifies all potential leakage locations, including along the length of the angle iron. While the location with the highest probability may not correspond to the true leakage site, the system ensures comprehensive identification of all possible leakage areas.

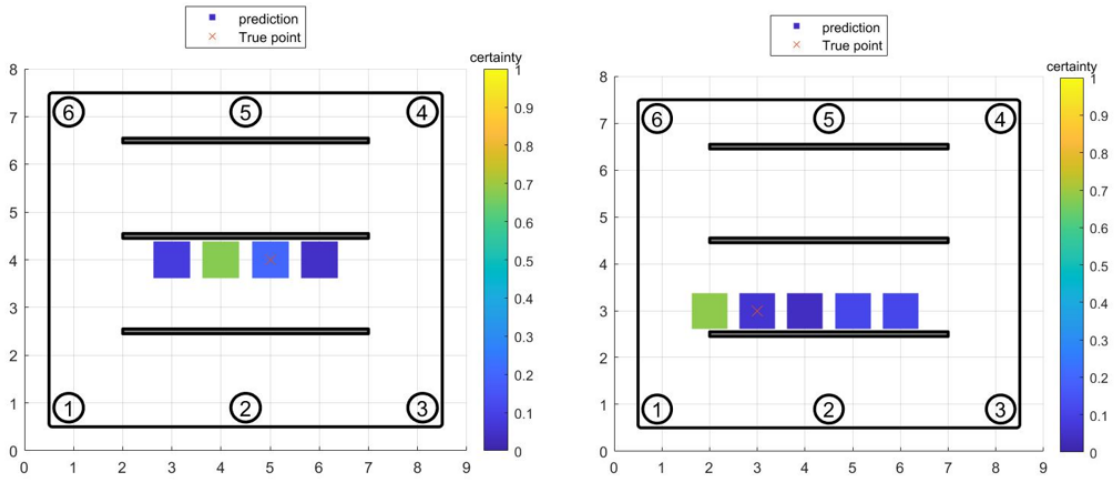


Figure 4.16: PIGNN prediction results for challenging locations with similar flow rates in a non-uniform layup

If a leakage occurs in a region sufficiently distant from the angle irons, the system

accurately predicts the leakage location with high confidence. This is because areas farther from the angle iron channels exhibit low variance in flow rates. An example of this scenario is shown in Figure 4.17.

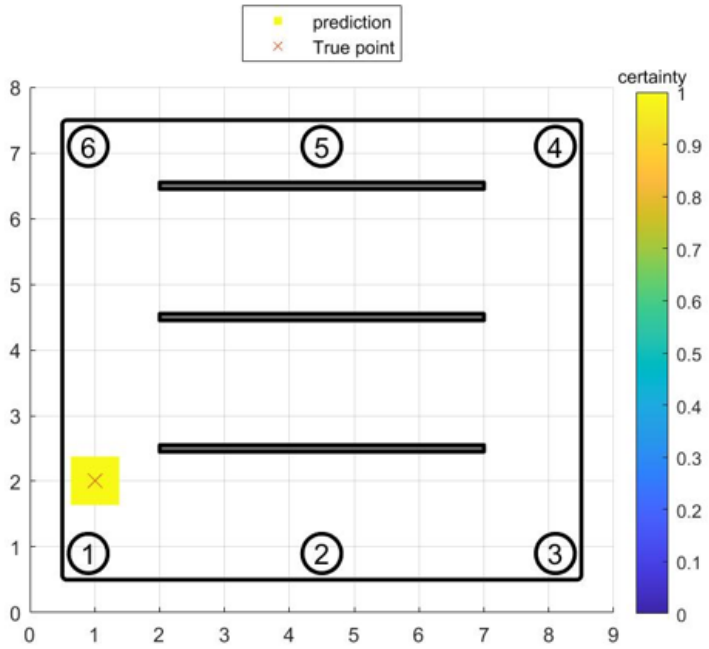


Figure 4.17: PIGNN prediction result for a sample location far from the angle irons

In many scenarios, accurately pinpointing the exact leakage location with high confidence is challenging. However, this new design enables the system to predict challenging leakages with lower confidence and probability while incorporating the ability to update itself based on new evidence—an improvement that is not achievable with conventional regression model training.

### 4.4.2 Results of applying Bayesian inference

Bayesian updating can be employed to leverage any evidence within the system. For instance, users can begin their search in the regions with high probabilities as depicted in Figure 4.16. If they are unable to locate the leakage, the probabilities can be updated, allowing the system to predict new leakage locations. This iterative process enables the system to adapt and refine its predictions based on the acquired evidence.

Figure 4.18 illustrates the prediction results in a scenario where the user was initially unable to locate the predicted leakage, as shown in the left-hand image

of Figure 4.16. Initially, the user investigates the area with the highest predicted probability but fails to find the leakage. This new evidence is then fed back into the system, which updates the probability distribution using Bayesian inference to incorporate the new information.

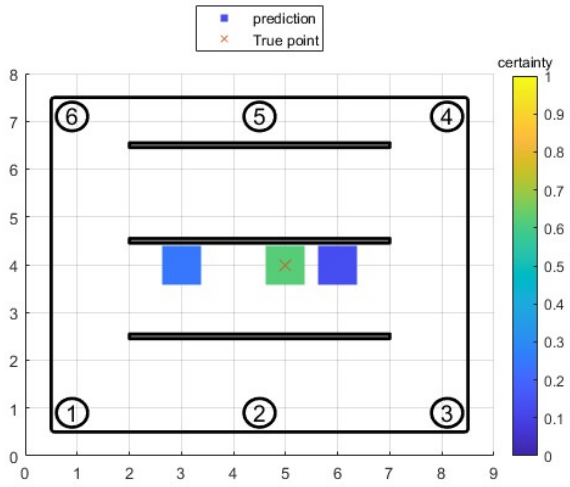


Figure 4.18: PIGNN prediction result after a Bayesian update with a new evidence

As discussed in Section 3.9.6, Bayesian inference can be utilized to increase the probability of regions more susceptible to leakages, such as boundaries. This information is incorporated into the system as prior knowledge. Figure 4.19 demonstrates how including priors affects the probability of leakage at the boundaries, as expected. In the single leakage localization results shown in Figure 4.11, leakages near the boundaries tend to be predicted closer to the center of the layup. Incorporating Bayesian inference could enhance the accuracy of these predictions.

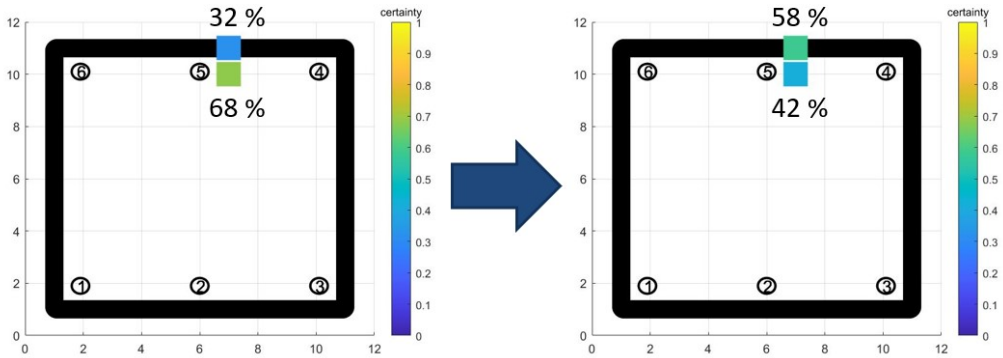


Figure 4.19: PIGNN prediction result for leakages close to the boundary

### 4.4.3 Multiple leakage localization results

The method for localizing two leakages was presented in Section 3.9.5. This approach minimizes the total error in flow rate combinations derived from node pairs in the GNN model and assumes that both leaks are approximately the same size. The model has been tested on approximately 30 experimental multi-leakage scenarios, achieving an average prediction error of 13.8$cm$ on a 90$cm \times 90cm$ layup. Figure 4.20 illustrates the prediction results alongside the actual leak locations for a case with two simultaneous leakages. It is noteworthy that this model uses a finer mesh and a greater number of nodes compared to previous models. Additionally, the Softmax function and Bayesian inference can also be applied to the multiple leakage case.

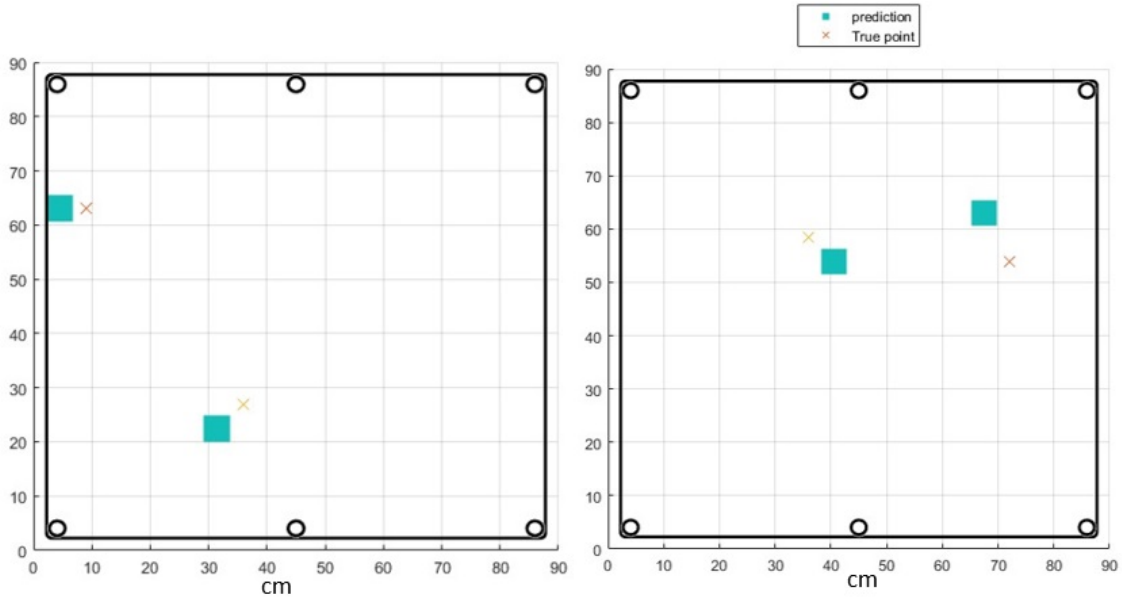


Figure 4.20: PIGNN prediction result for leakages close to the boundary

Although the system achieves acceptable predictions in the previous example, there are cases where it fails to identify locations close to the leaks. As discussed in Section 4.3.2, using an enhanced feature set was proposed to address this limitation. By closing the vacuum valves, it is possible to generate more independent features. This enhanced feature set has been shown to be effective in classifying the number of leakages. Figure 4.21 compares the results obtained using the normal feature set and the enhanced feature set for prediction. One of the leaks in this example, is extremely close to one of the vacuum ports and to the boundary, making the prediction challenging. The normal feature set uses six normalized flow rates as

input, whereas the enhanced feature set incorporates 15 flow rates by closing the port with the highest flow rate after measuring the initial flow rates for two times.

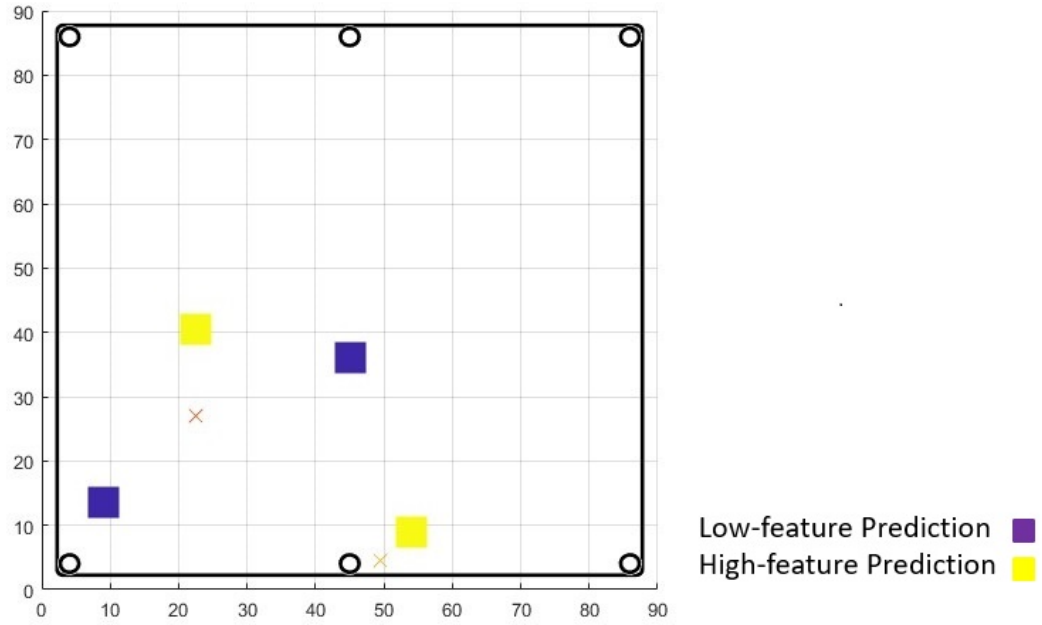


Figure 4.21: PIGNN prediction result for leakages close to the boundary

The results in the figure demonstrate that using the enhanced feature set improves accuracy. However, it is not always necessary to use the enhanced feature set, as it increases the time and effort required. The normal feature set can still provide reliable predictions, as shown in Figure 4.20.

One of the most significant advantages of the GNN model is its ability to address the problem of multiple leakage localization—a challenge that has not been solved by any other method. Additionally, generating datasets for multiple leakages is only required for the classification problem. Once the number of leaks is determined, single-leakage examples are sufficient for constructing the model, and further to be used for multiple leakage localization.

## 4.5 Vacuum port placement optimization results

The importance of using an optimal vacuum port placement was discussed in Section 3.10. In this section, the results of port optimization for the case study consisting of 81 nodes and six vacuum ports are shown in Figure 4.22. Four different configurations

have been compared with each other in the terms of their score, optimization time and prediction average error. Configuration (a) is the original setup where ports were located in corners and middle of top and bottom rows. Configurations (b) and (c) are the optimization results for simplified and precised solutions, respectively. Configuration (d) shows a configuration when all vacuum ports are aligned in a row as a sample for comparison purposes. In this particular configuration, vacuum ports are close to each other and therefore flowrates will be less independent, resulting in a more penalized cost function.

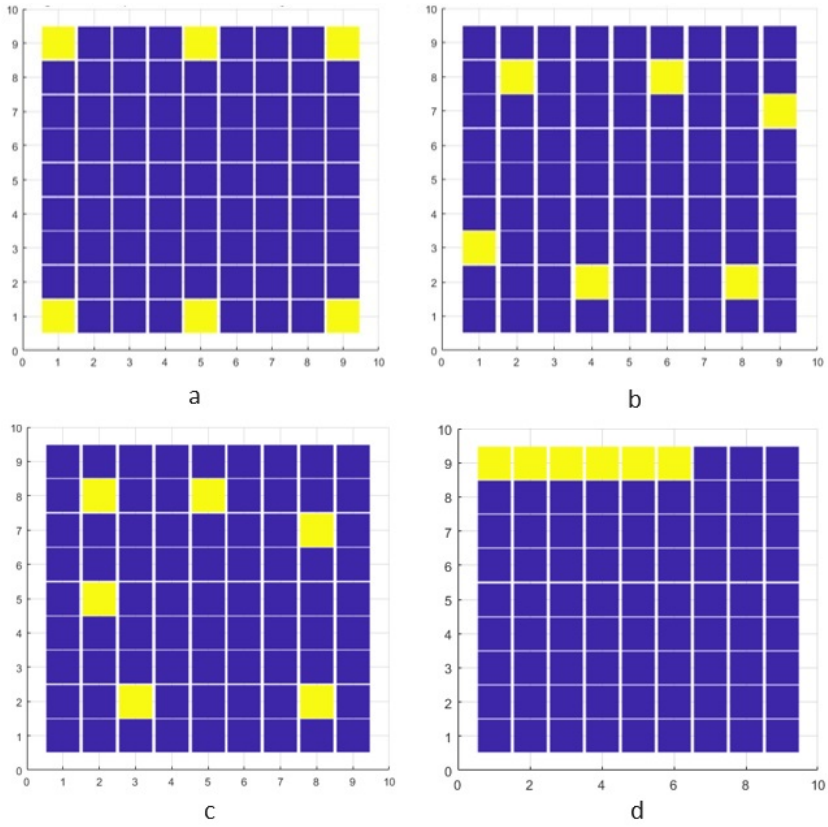


Figure 4.22: Optimization results for a 90cm×90cm layup consisting 81 nodes and six vacuum ports; a) initial setup, b) results from the simplified solution, c) results from the precise solution and d) a sample configuration placing vacuum ports in a same row

### 4.5.1 Performance evaluation of the optimized port placements

The ultimate objective of the vacuum port placement optimization is to improve the system on the terms of prediction accuracy. Therefore, system's performance must be evaluated. The model's leakage location prediction accuracy for a particular configuration could be assessed using these four steps:

**Step 1** Synthesized data generation: generation of synthesized data for the particular vacuum port placement using simulation or the precise flowrate calculation method.

**Step 2** Model training: a regression model will be trained using a proper regression machine learning model such as exponential Gaussian process regression (GPR).

**Step 3** Experimental data acquisition: An experimental setup with the same vacuum port configuration will be made and manual leakages will be created in different locations and the flowrates will be recorded. This experimental data will be used as a test set that has been never seen by the trained model.

**Step 4** Using experimental flowrates from step 3 as input to the trained model in step 2 and comparing the predicted location with the actual location of the leaks to calculate the prediction error. The average error then will be calculated for the system.

### 4.5.2 Comparison results

Using this procedure, prediction errors of all four configurations could be calculated and compared. Table 4.16 shows the results of the four configurations. The score for the optimized results outperform the original and sample configurations; case (a) and (d) respectively. The optimized solution resulted from precise method slightly outperform solution from simplified method with the price of more calculation time. The score is calculated through the summation of flowrate variances. Higher score means flowrates or the machine learning features are more independent and therefore training will be less likely to suffer from overfitting. Configurations with higher scores end up having a lower average error and higher accuracy; meaning that the user will have to search for a smaller area for a possible leakage. This smaller area is shown in the last column of the table as "Search area". Using a proper vacuum port placement

can decrease the possible leakage area from 18 percent to 8 percent. Search area for all four configurations are visualized in Figure 4.23.

Table 4.16: Results for four different configurations in Figure 4.22

| Configuration | Score | Time | Prediction error (cm) | Search area (%) |
|---|---|---|---|---|
| Original | 9.58 | - | 8.4 | 12 |
| Simplified | 14.13 | 1 sec | 7.4 | 9.5 |
| Precise | 14.91 | 18 min | 6.8 | 8 |
| Sample | 4.37 | - | 10.3 | 18.5 |

Case (d) where all vacuum ports are placed horizontally in a row has a good accuracy for predicting leakages on the horizontal axis but the prediction has a high error on the vertical axis perpendicular to the vacuum ports line. This is due to lacking enough features and flowrate information among different vacuum ports as moving a leakage across the vertical axis will not change flowrates measured on vacuum ports.

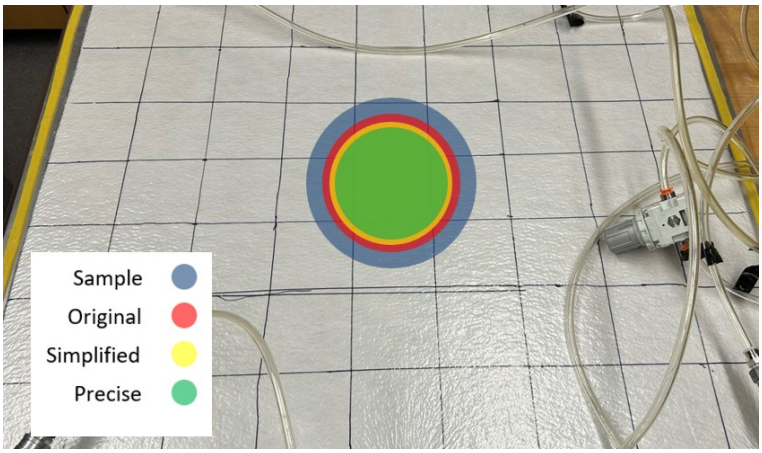


Figure 4.23: Search areas for a possible leakage after training the model with different port configurations; Optimized port placements outperform other methods and achieve a higher accuracy and require a smaller area for leakage search

In this section, the results of vacuum port placement were analyzed in terms of efficient and accurate leakage localization. Two different flow rate calculation methods—simplified and more precise—were introduced. The simplified method, which is based on the distance between the leakage point and the vacuum port, is effective when there is uniform permeability in the layup. This method is more applicable in cases where the product thickness remains relatively constant and equal layers of breather material are used throughout the layup. For scenarios where these

conditions do not hold, such as the examples in Figure 3.11, a more precise method based on Kirchhoff's law was introduced. It was shown in Table 4.17 that assuming uniform resistance in a nonuniform setup may cause errors to increase by 300% and this significant error reduces the reliability of optimization results. Although the precise method is more powerful, the improvement is negligible in uniform layups, making the simplified method preferable due to its much lower computation time.

## 4.6 Results of physical model parameter training

This section presents the results of training the physical parameters of the model, as discussed previously in Section 3.11. We begin with a detailed explanation of the importance of training these physical parameters, using our case study as an example.

### 4.6.1 Significance of physical model parameter training: a case study

In this section, we compare the physically designed model with the uniform model to demonstrate the importance of using a valid electrical circuit model. The case study is a 90cm×90cm layup with a more permeability in the bottom section shown in Figure 3.32. The network of uniform electrical resistors previously shown in Figure 3.33 is compared with a network where the resistors below row six are reduced by 50 percent to represent the non-uniform layup. Afterwards, the flow rates are compared with experimental results. Figure 4.24 shows the error distribution among different leakage locations resulted from both models.

Higher error values are observed in the uniform setup shown in Figure 4.24 (a) compared to the valid model results in Figure 4.24 (b). Using a model with lower accuracy for data generation leads to an inadequate and inaccurate machine learning model. Table 4.17 shows the comparison results of the two models. Using a uniform layup increases flow rate errors by more than 50 percent compared to a properly modeled system. This value was obtained by comparing simulation results with experimental flow rates using Eq. 3.6. Due to this higher flow rate error, the prediction error also increases by approximately 50 percent. Prediction accuracy is improved by training the model with synthesized data. This larger prediction error leads to a wider area on the layup that the user must search for potential leakage, thereby increasing leakage localization time. The table illustrates the importance of using

an accurate physical model for simulation, and thus the significance of training the model's physical parameters.

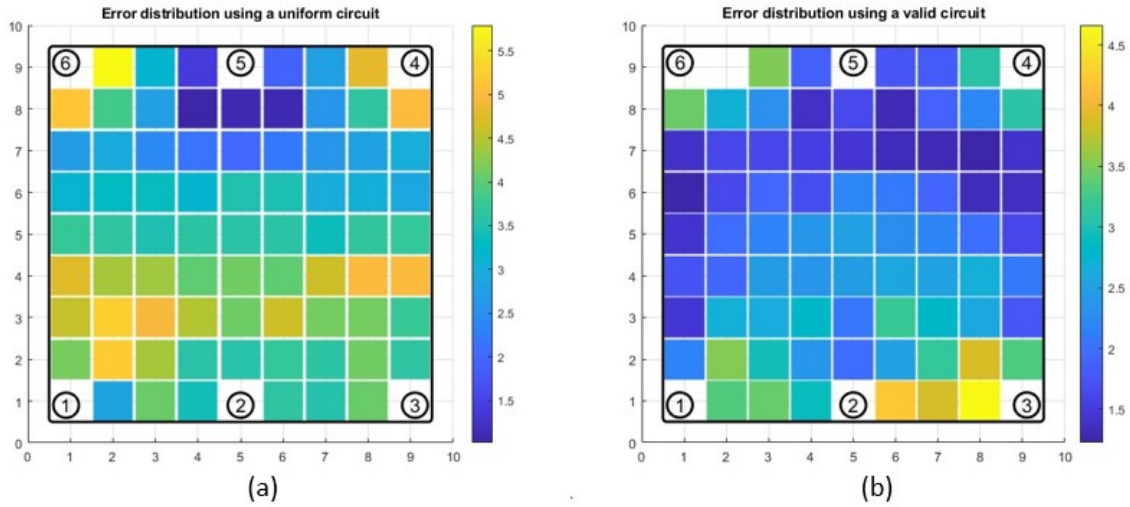


Figure 4.24: Error distribution when (a) uniform setup (b) valid electrical circuit are used for simulating a non-uniform layup

Table 4.17: Comparison of using a uniform or valid electrical circuit design on a layup with non-uniform permeability

| Model | Flowrate error (%) | Prediction error (cm) | Area shrinkage(%) |
|---|---|---|---|
| Uniform setup | 3.57 | 14.1 | 35 |
| Valid electrical circuit | 2.30 | 7.5 | 10 |

The primary objective of this section is to utilize the available data to construct a more accurate model for data generation.

### 4.6.2 Training results using synthesized data

The physical parameter training procedure and associated training parameters were thoroughly discussed in Section 3.11. We trained the relative resistance of different sections of the layup, or the training parameters, using various data subsets to evaluate their sensitivity to the data. Figure 4.25 illustrates the cost function across different iterations. The cost function, defined in Eq. 3.45, quantifies the error between the model's output and the actual data.

In each iteration, the algorithm must go through the entire simulation multiple times for each data point (leakage location). According to Eq. 3.47, the derivatives

for all parameters need to be calculated individually and summed across all available leakage locations in the dataset. This process is time-consuming and could be optimized using mini-batch training. The results in Figure 4.25 show that after approximately five iterations, further improvement in the cost function or simulation outputs is minimal. The remaining error is attributed to other governing factors not accounted for in the training parameters that influence the simulation and experimental results.

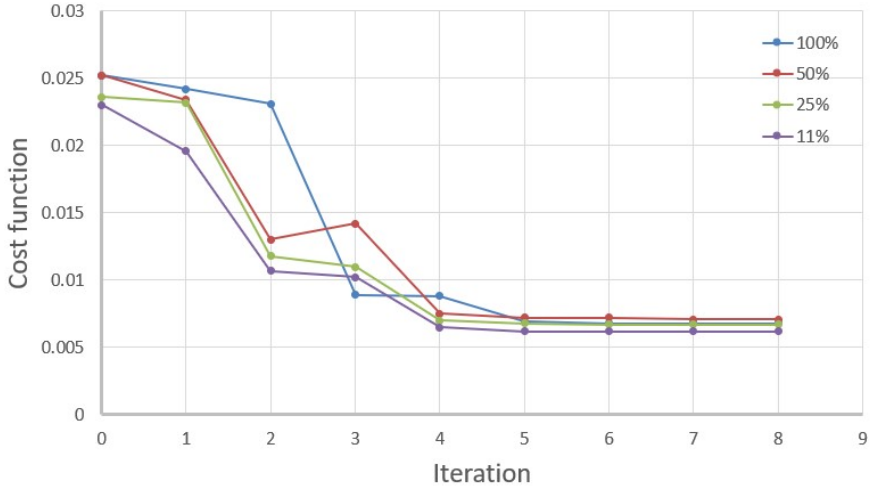


Figure 4.25: Iterative cost function trained on different subsets of data

In Figure 4.25, although different subsets of data converge to similar cost function values after sufficient iterations, this does not imply they have similar performance. The cost function reflects the average error across the available data, and the algorithm optimizes based on this metric. Models trained on smaller subsets of data may still exhibit high error on unseen data, potentially suffering from high bias. To assess each model's performance, the trained parameters can be compared with the objective parameters shown in Figure 3.35, which are based on prior knowledge from experimental design. The RMS error between each model's trained parameters and the objective parameters is provided in Table 4.18.

Table 4.18: Comparison of using a uniform or valid electrical circuit design on a layup with non-uniform permeability

| Data usage | No data | 100% | 50% | 25% | 11% |
|---|---|---|---|---|---|
| RMS error | 0.289 | 0.105 | 0.111 | 0.110 | 0.134 |
| Improvement (%) | - | 64 | 62 | 61 | 54 |

Five different models, utilizing 11%, 25%, 50%, 100%, and no data (the uniform model), are compared, with their RMS errors relative to the objective and corresponding improvements shown in the table. As the amount of training data decreases, the RMS error increases, leading to a reduction in improvement. As expected, the untrained model with no data exhibits the highest error. The results indicate that reducing the data has a minimal impact on performance as long as it is well-distributed across the layup to cover all sections. When data usage is reduced by 50%, the model's performance decreases by only about 2%. In a more significant reduction, with approximately 90% less data, performance drops by only 10%. Data scarcity is a common challenge in leakage localization, and these results demonstrate that a data-driven, physics-informed model can enhance performance with limited but valuable data.

Figure 4.26 shows the training outputs on the design parameters. Each element of the matrix, shows the relative resistance of each region corresponding to Figure 3.34. The initial configuration was the uniform setup, and the synthesized data used for training was generated by configuration shown in the "objective" matrix. The bottom matrices represent the trained parameters. The normalized trained matrix is then compared with the objective matrix. The relative resistance parameters in the middle and bottom rows are similar and the errors mostly come from the top row.

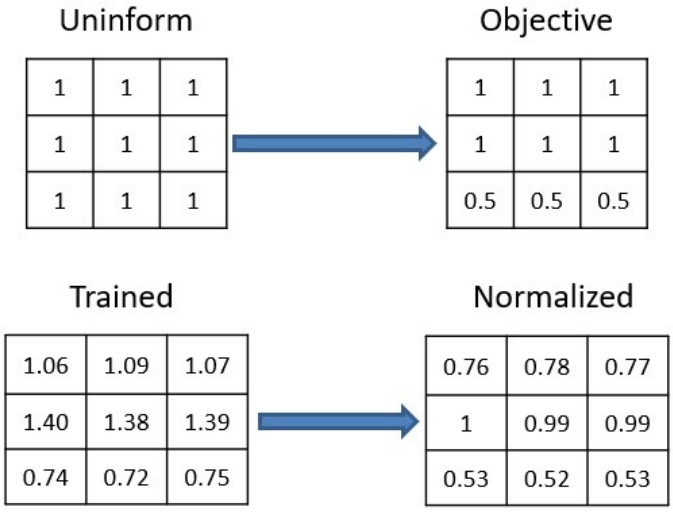


Figure 4.26: Uniform, objective and trained outcomes on synthesized data for the physical parameter training of relative resistance

### 4.6.3 Training results on experimental data

In Section 4.6.2, the models were trained using synthesized data derived from the validated electrical circuit model. In this section, however, experimental data will be utilized for training. Collecting experimental data, particularly on production lines, is a more challenging and labor-intensive process. Nonetheless, without such data, it is difficult to develop a reliable simulation model, as users may lack accurate information regarding the relative permeability distribution of their layup and other relevant settings. Therefore, experimental models will be employed to enhance the simulation model, which can subsequently be used to generate synthetic data. This approach addresses the typical insufficiency of experimental data for comprehensive model training. Furthermore, a validated simulation model is necessary to cover multiple leakage scenarios, as the volume of data required for these complex cases cannot feasibly be obtained through experiments alone.

To test this methodology in our case study, a total of 10 experiments were conducted to train the model on the relative resistance of various regions. The locations of these leakages are indicated by red circles in Figure 4.27. The flow rates measured by the sensors, along with the coordinates of each corresponding leakage, were recorded. This data will then be used for training according to the proposed algorithm.

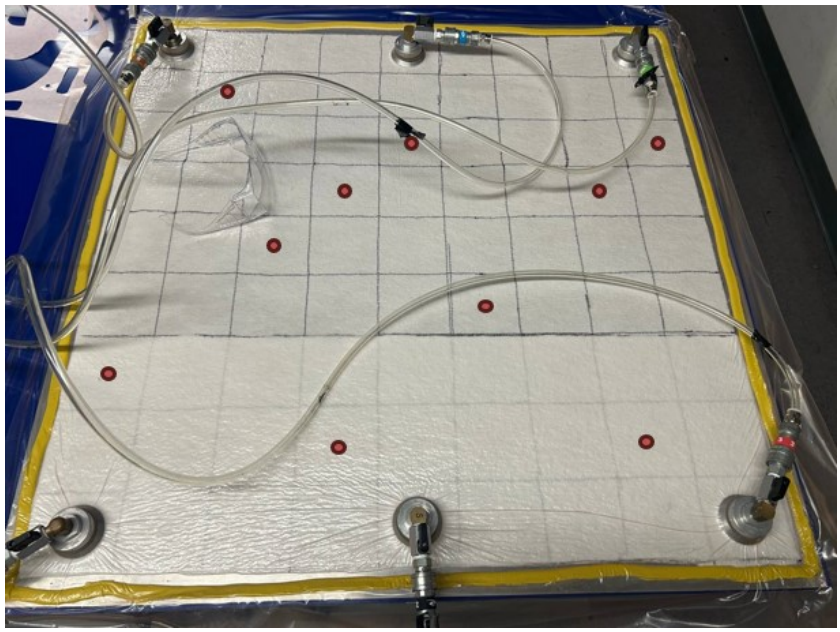

Figure 4.27: Non-uniform layup with data collected from intentionally created leaks at the indicated locations

After training the parameters, the normalized relative resistances were obtained and are shown in Figure 4.28. The bottom rows display the lowest resistance values, as expected. The data indicates a relatively lower resistance on the right side of the layup, likely due to wrinkles in the bag. We refer to this model as the "Physically trained model".

| | | |
|---|---|---|
| 0.97 | 1 | 0.91 |
| 0.87 | 0.86 | 0.86 |
| 0.52 | 0.52 | 0.41 |

Figure 4.28: Normalized relative resistance of regions trained by experimental data for a non-uniform layup

The performance of three models—the uniform layup model, the physically designed model, and the physically trained model—in leak localization is compared in Table 4.19. Each of these models was used for data synthesis, with the generated data subsequently used to train a machine learning model. In the uniform model, the relative resistance is equal across all regions and have been presented in the table for comparison purpose. The physically designed model assumes that using two layers of breather material halves the resistance, and its relative resistance matrix is presented in Figure 4.26 as the objective matrix.

Table 4.19: Experimental comparison of three trained models on a layup with non-uniform permeability

| Model | Flowrate error (%) | Prediction error (cm) | Area shrinkage(%) |
|---|---|---|---|
| Uniform setup | 4.25 | 15.0 | 39 |
| Physically designed | 2.67 | 10.5 | 19 |
| Physically trained | 1.91 | 7.5 | 7.5 |

It was previously shown using synthesized data that the uniform layup assumption leads to significant errors in flow rate calculations, and similar results were observed with experimental data. The average prediction error for a square-shaped layup with the size of 90 cm is 15.0 cm. This error is considered substantial, requiring the user to search over a large area for potential leaks. In contrast, the physically designed model outperforms the uniform setup. By incorporating prior information from the

user into the model, the flow rate error is reduced by 40%, and the prediction error is decreased by 30%. achieving an average prediction error of 10.5 cm. With this model, the user needs to search a smaller area, nearly half the size of the uniform model.

However, it is important to note that the assumption made in the physically designed model—specifically, that areas with two layers of breather experience a 50% reduction in resistance—may not be entirely valid. This is why the physically trained model performs better than the physically designed model. The physically trained model achieves a flow rate error of 1.91%, which is considered low, given the non-linearity of the layup. As a result, the prediction error is reduced to 7.5 cm, and the user only needs to search a very small area compared to the previous models.

Finally, even the most accurate model will still have a minimum prediction error of 5.0 cm, as leakages are sampled from small squares, similar in size to those shown in Figure 4.27.

### 4.6.4 Training results on challenging layups

The challenges of using this framework were discussed previously in Section 3.11.5. One particularly challenging layup involves the use of an extra diagonal breather, as shown in Figure 3.37. The challenging part in this training is that each smaller region in this case does not have uniform permeability and the trained relative permeability of these regions will reflect a compromise between the lower and higher permeabilities. The same algorithm described in the previous section was applied here, comparing three models: the uniform setup, the physically designed model, and the physically trained model. The results are presented in Table 4.20.

Table 4.20: Comparison of different trained models on a layup with non-uniform diagonal permeability

| Model | Flowrate error (%) | Prediction error (cm) | Area shrinkage(%) |
|---|---|---|---|
| Uniform setup | 3.82 | 12.1 | 26 |
| Physically designed | 3.16 | 11.4 | 23 |
| Physically trained | 1.99 | 9.7 | 16 |

As in the previous case study, the physically designed model outperforms the uniform setup, and the physically trained model outperforms both. Due to the limited number of training parameters in this layup, higher errors are observed compared to

the previous case study. Figure 4.29 illustrates the normalized relative resistance of regions for each model, with lower resistance observed along the main diagonal in the trained model.

Uninform

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

Designed

| | | |
|---|---|---|
| 0.5 | 1 | 1 |
| 1 | 0.5 | 1 |
| 1 | 1 | 0.5 |

Trained

| | | |
|---|---|---|
| 0.75 | 1 | 0.9 |
| 0.83 | 0.80 | 0.83 |
| 0.74 | 0.68 | 0.43 |

Figure 4.29: Uniform, physically designed and physically trained outcomes on experimental data for the physical parameter training of relative resistance for non-uniform diagonal breather

One approach to address this issue is to increase the number of design parameters. In the original design, there were a total of 9 training parameters. Figure 4.30 shows a similar setup with 16 areas featuring trainable relative permeability. This increased number of parameters provides greater flexibility for training, though it requires additional computational time. However, the system may also become more susceptible to overfitting.

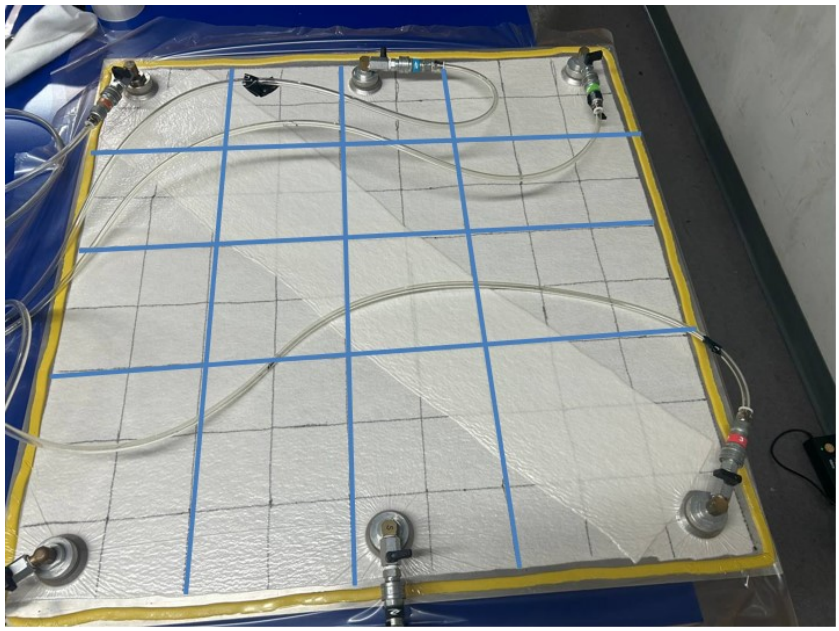

Figure 4.30: Layup with diagonal non-uniform permeability, featuring an increased number of training parameters for enhanced performance

The system with 16 distinct areas was trained, and Figure 4.31 shows the normalized relative resistance of regions based on experimental data. Lower resistance is observed in regions close to the main diagonal.

| | | | |
|---|---|---|---|
| 0.8 | 0.83 | 1 | 0.88 |
| 0.79 | 0.80 | 0.81 | 0.83 |
| 0.81 | 0.79 | 0.79 | 0.76 |
| 0.88 | 0.79 | 0.86 | 0.61 |

Figure 4.31: Normalized relative resistance of regions trained by experimental data for the layup with diagonal non-uniform permeability

The model exhibited a slightly higher flow rate error on experimental samples that were not seen during the training process, with a flow rate error of 2.13%. This could be due to greater discrepancies between some of the experiments and the trained model. However, the new design achieved a lower prediction error of 8.4 cm, outperforming the previous model with fewer parameters. This improved performance was anticipated, as the newly defined regions demonstrate greater consistency with the actual permeability.

# Chapter 5

# Conclusions

The objective of this research was to develop a robust method for detecting and localizing leakage in vacuum bags used in composite manufacturing through machine learning-based approaches. The study encompassed experimental design, the creation of a simulation model for data synthesis and validation, and the training of machine learning models. This chapter summarizes the key findings, highlights the novelties and contributions to the state of the art, and discusses the limitations of the study. Recommendations for future work are also presented.

## 5.1 Recap of key findings

In the first stage, we proposed a simulation method for modeling gas flow rates in VARTM systems, enabling the synthesis of large datasets. By introducing the concept of using relative port contributions instead of absolute values, the simulations remain valid regardless of leakage size. This approach was validated through a two-way ANOVA performed on 64 different examples across four needle sizes.

The electric-based analogy was tested on various configurations and setups. Overall, the error in all experiments was generally less than 3%, except near the ports, where it increased slightly but remained below 4.5%. Each simulation had a runtime of under 2 minutes on a Core-i7 2.4 GHz CPU, demonstrating sufficient speed for generating large synthetic datasets. The presented framework is capable of simulating multi-leakage scenarios, provided the leakages are of the same size. Since the framework relies on an electrical circuit analogy, which adheres to the superposition principle, multi-leakage scenarios can be recreated by introducing positive voltages

at multiple nodes.

Here's a revised version of your text for improved clarity and flow:

The synthesized data generated from various configurations was utilized for training regression models to predict leakage locations based on flow rate inputs. Among the models tested, Gaussian SVM, exponential GPR, and decision trees performed best for predicting leakage locations. Shallow neural networks also achieved a suitable bias-variance trade-off. The predictions were validated on real-world applications, yielding satisfactory results. However, despite the accuracy and some degree of generalizability, the models may fail when faced with a new layup featuring significant changes. This limitation arises because the regression models do not incorporate the geometry and configurations of the layup and vacuum ports, omitting crucial information.

To enhance generalizability and adaptability, a grid neural network (GNN) framework was introduced. This approach significantly reduced the need for training data while simultaneously improving model performance by integrating physical knowledge. The GNN framework demonstrated exceptional accuracy in identifying potential leakage locations and their associated probabilities, particularly in cases with identical flow rates. Moreover, it enables dynamic updating of probabilities by incorporating new evidence provided by users or experts through Bayesian updating. This capability allows the model to flexibly adapt to new information and refine its predictions.

The challenge of addressing multiple leakages in the layup has been approached under the assumption that the leakages are of the same size. Initially, a large dataset comprising two labeled classes—scenarios with either one or two leaks—is generated. A classification model is then trained to identify the number of leaks present on the layup. However, the task of classifying the number of leaks becomes more complex and prone to overfitting when the two leaks are very close to each other in multiple leakage scenarios. To mitigate this issue, a minimum distance criterion between the leaks is introduced.

Once the number of leaks is determined using the classification algorithm, the corresponding leakage localization algorithm is applied to localize the leaks. The localization approach for multiple leakages relies on minimizing combinations of flow rates. Additionally, a method for enhancing the input feature set was introduced by temporarily closing the vacuum port valve with the highest flow rate. This enhanced feature set outperformed the standard feature set by leveraging additional information

for improved predictions.

For complex layups with non-uniform permeability that are unknown to the user, a novel method for training the physical parameters of the model was introduced. This approach enables the system to infer permeability variations across different regions using the available data. Remarkably, this method performs effectively even with limited data. The physical parameter training algorithm provides a pathway to achieving a more accurate and reliable model, which is essential for handling multiple leakage scenarios.

Table 5.1 compared different algorithms and models and their ability to handle different scenarios. Additionally, accuracy, generalizability, and training dataset size for each of the models are compared. Models such as classification learner and physical parameter training are required for handling some of the scenarios.

Table 5.1: Comparison of the developed model performance on different scenarios

| Model | Leakage | Permeability | Accuracy | Gen | Data size |
|---|---|---|---|---|---|
| Truth Table | Single | Uniform | M | VL | H-VH |
| Regression Learner | Single | Uniform | H | M | M |
| | Single | Non-uniform | M | L | H |
| PIGNN | Single | Uniform | VH | VH | VL |
| | Single | Non-uniform | H | M | M |
| | Multiple | Uniform | M-H | L | L-M |
| Physical Parameter Training | Single | Non-uniform | VH | VH | L |
| | Multiple | Non-uniform | H | H | L-M |
| Classification learner | Multiple | All | H | M | Synthesized |

Gen: Generalizability
VL: Very Low, L: Low, M: Medium, H: High, VH: Very High

The Truth Table model has been used as a baseline for comparison with the developed models. The regression model is versatile and effective when sufficient data is available for a specific layup. In such cases, a black-box model can be trained, avoiding the limitations and potential errors associated with simulations. Figure 5.1 illustrates the average error of the regression model compared to the physical training model for a non-uniform layup.

When data is limited, the physical training model outperforms the regression

model, as it leverages additional information about the geometry and configuration of vacuum ports and is initialized with physical knowledge. However, the performance of the physically trained model is inherently capped, as achieving an ideal simulation model is often infeasible. In contrast, the regression model has no theoretical performance limit when sufficient data is available, allowing it to eventually surpass the physical training model.

Despite these strengths, the regression model has notable limitations. It cannot handle multiple leakage scenarios, even with ample data, and its generalizability is inherently weaker compared to the PIGNN model.

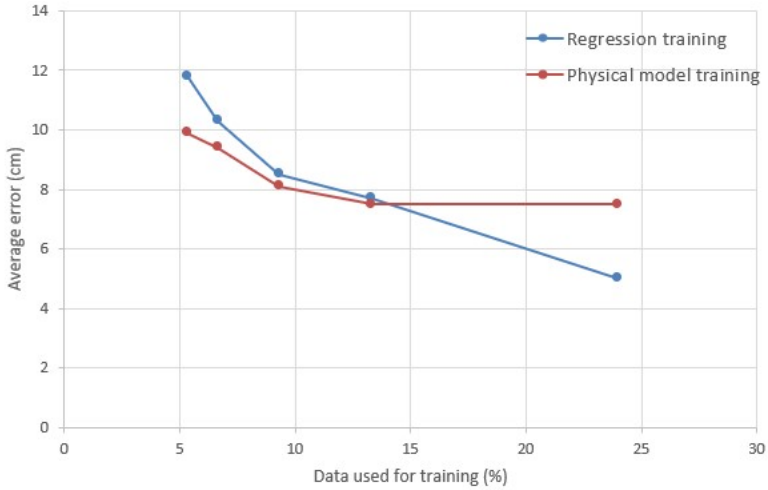


Figure 5.1: Performance comparison of regression model and physical parameter training based on training data availability

The qualities that truly illustrate the potential of this framework fall into three main categories:

- Generalizability: Despite the experiments and training have been conducted at a laboratory scale, the leakage prediction corresponds well with the experimental results across all geometries tested. The statement is also valid when dealing with layups with non-uniform permeabilities. Notably, the versatility of the GNN framework allows its implementation across assemblies of various sizes and shapes, as the network can be designed and customized to suit each specific use case.

- Efficiency: There are a plethora of commercially available fluid flow analytical simulation software suites, but they are generally not well suited for data

synthesis. The traditional numerical simulation methods are far too hardware intensive to be of any value for the type of simulations required for data synthesis. Instead of comprehensive flow simulations that would be required for studying resin flow, the framework needed a simple, reliable, and reproducible simulation method for computing flowrate values. Moreover, the proposed PI-GNN network exhibits efficiency in terms of data requirements. It demonstrates exceptional performance even when provided with a limited amount of data.

- Reliability: When comparing the outputs of the two systems, the error values are typically less than 3%. This demonstrates a high level of correlation between the experimental outputs and the results of the simulation. Thus, flowrate values can be used interchangeably with the electrical current values for training a machine learning agent. Additionally, the system excels in providing answers along with their corresponding probabilities, ensuring a reliable mechanism for avoiding incorrect outputs.

Hence, the output of the proposed framework has a satisfactory level of generalizability, efficiency, and reliability for future development of a data-dependent machine learning-based leakage localization method for use in composite manufacturing when dealing with real-size products with multiple leakages.

## 5.2 Contribution to knowledge

The main contributions to the knowledge could be categorized in four main aspects.

1. **Introducing a reliable data synthesis tool**

   In previous studies, flow rate data from different locations were obtained experimentally and extended using symmetry. The newly proposed simulation method, based on an electrical circuit analogy integrated with the physical model, enables the rapid and efficient synthesis of large datasets. This method, validated across diverse layups, serves as a viable alternative to the time-consuming processes of conducting experiments, recording, and post-processing data.

2. **Addressing the problem for layups with non-uniform permeability**

Most previous studies focused on uniform, symmetrical, and regularly shaped layup geometries due to the complexities introduced by elements with non-uniform permeability. While this issue could theoretically be addressed through experiments, such an approach lacks a generalizable solution. In our study, we tackled various forms of non-linearity and non-uniform permeability within the system, as well as non-symmetrical and irregular geometries. Furthermore, our proposed framework for physical parameter training enables the model to learn and adapt to non-uniform locations effectively.

3. **Addressing the problem for multiple leakage scenarios**

   The possibility of multiple leakages on the layup has been acknowledged in previous studies but has not been systematically addressed. In our research, we have explored this problem by first classifying the number of leakages on the layup. Subsequently, the system employs a tailored algorithm for precise leakage localization. To enhance the accuracy of the approach, we have also proposed and investigated the integration of additional input features, enabling the model to better capture the complexities of multi-leakage scenarios. This comprehensive framework offers a significant step forward in addressing multiple leakage challenges in composite manufacturing setups.

4. **Vacuum port optimal placement**

   Determining the optimal locations of vacuum ports on a layup is a problem that has not been addressed in previous research. While several studies have explored the placement and number of inlet ports, none have focused specifically on vacuum ports. This question is particularly relevant for large-scale parts manufactured using prepreg, where effective vacuum port placement is critical for ensuring proper consolidation and minimizing defects.

   In our study, we developed a generalizable solution for optimal vacuum port placement, accommodating any number of ports on various layup geometries. The primary objective was to achieve the most reliable leakage localization while maintaining system efficiency. Given the high dimensionality and complexity of this problem, we employed a hierarchical optimization approach to systematically reduce the search space and identify the best configurations. This methodology provides a robust framework for optimizing vacuum port placement in practical manufacturing scenarios.

## 5.3 Limitations of the study

Despite the contributions and innovations of this work, certain limitations remain. Most of these stem from the constraints of the technology employed for leakage detection and localization. Due to the inherent nature of the technology, pinpointing the exact location of a leak is not feasible; instead, the system can only localize the leakage. This limitation arises from the machine learning approach, as achieving high accuracy while maintaining generalizability is challenging. Consequently, the results are not entirely reliable, and errors may persist within the system. This unreliability affects all stages, including single or double leakage localization, simulation methods, and the classification of the number of leakages. While it may not be possible to completely eliminate these errors, efforts have been made and can continue to be made to minimize them.

The system also has limitations when detecting leaks in specific locations. For instance, its reliability decreases for leaks near vacuum ports or close to the boundaries. When leaks are extremely close to vacuum ports, the flow rate values may exceed the sensor's capacity. As the distance between the leak and the port decreases, modeling 2D viscous resistance with 1D resistors becomes more challenging. Consequently, a finer mesh of electrical resistors is required to address this issue. Similar challenges occur in regions near the boundaries, where accurate modeling becomes more difficult.

Various algorithms and methods, such as the electrical circuit analogy, GNN, and physical parameter training, have been introduced to enhance generalizability and reduce the amount of experimental data needed for training. However, when a new layup is introduced, much of the process must be repeated from the beginning. Although the framework is applicable to most layups, experimental data is still necessary to build the model, meaning the system is not fully independent of data. New simulations, data, and training are required for each new layup.

This research addresses the problem of leakage detection and localization for multiple leakages for the first time. However, it is based on the key assumption that two leakages exist in the layup, both of roughly equal size. While the solutions remain useful when one leakage is significantly larger than the other, they fall short when the leakages are unequal and comparable in size. Our conclusion is that solving the multiple leakage problem in its general form using the current method and technology may not be feasible. To enhance the system's analytical capabilities, additional methods and technologies need to be integrated. The challenge of handling more than

two leaks will be explored in future work.

## 5.4 Future work and recommendations

There are possible future improvements that are recommended in the following.

1. Even though multiple leakage scenarios were studied in this research, the results and solutions are based on the assumption of having two leaks of the same size on the layup. To enhance the generalizability of the findings, future studies are recommended to address scenarios involving leaks of varying sizes and more than two leaks on the layup. Furthermore, incorporating additional sensors or devices for this purpose is also suggested.

2. The proposed method, based on flow rate measurement in evacuation lines, can predict the approximate location of a leak but cannot pinpoint its exact position. Addressing this limitation may require incorporating additional technologies, such as sensors with different detection mechanisms. For example, while scanning the entire layup area with a thermo-camera may not be feasible, using the camera to inspect a smaller, targeted region could help locate the leak more precisely. Therefore, integrating complementary technologies with the current method is recommended for future work to enhance leak detection accuracy.

3. All subsystems of the existing framework, including the data synthesis tool, system training, and prediction modules, have been thoroughly tested and validated in various real-world laboratory setups. This robust foundation positions the system to potentially handle industrial-scale components. Additionally, the number of vacuum channels can be increased to support larger-scale applications. A crucial future step is validating the system with large-scale components to further enhance its reliability and applicability. A designed framework is also required for technicians dealing with the industrial large-scale components.

4. The model's accuracy and efficiency could be improved by incorporating additional dynamic components, such as inductors or capacitors, into the electrical circuit. This approach would better emulate real-time data utilization, rather than relying solely on averaged flow rates, thereby introducing more training

features. However, due to the nature of the problem, improving accuracy cannot rely solely on the current features, as adopting a more complex model with additional parameters risks overfitting. Addressing this challenge may involve exploring the inclusion of alternative features.

Additionally, investigating system features to enhance accuracy presents an avenue for further research. Another potential area of exploration is developing an automatic mesh generator algorithm for regions requiring finer mesh resolution, which could further improve system performance.

5. The data synthesis tool, based on the electrical circuit, is created by an expert in the software. Currently, in the GNN design, users are required to manually input the layup size, geometry, and other settings. This process could be automated by developing a system that generates simulations directly from an image or a layup data sheet using image perception algorithms.

# Bibliography

[1] Aytac Goren and Cesim Atas. Manufacturing of polymer matrix composites using vacuum assisted resin infusion molding. *Archives of Materials Science and Engineering*, 34, 12 2008.

[2] K. Balasubramanian, Mohamed T.H. Sultan, and N. Rajeswari. *4 - Manufacturing techniques of composites for aerospace applications*. Woodhead Publishing Series in Composites Science and Engineering. Woodhead Publishing, 2018.

[3] K Balasubramanian, Mohamed TH Sultan, and N Rajeswari. Manufacturing techniques of composites for aerospace applications. In *Sustainable composites for aerospace applications*, pages 55–67. Elsevier, 2018.

[4] A McIlhagger, E Archer, and R McIlhagger. Manufacturing processes for composite materials and components for aerospace applications. In *Polymer composites in the aerospace industry*, pages 59–81. Elsevier, 2020.

[5] LA Khan and AH Mehmood. Cost-effective composites manufacturing processes for automotive applications. In *Lightweight composite structures in transport*, pages 93–119. Elsevier, 2016.

[6] Jeong-Min Lee, Byung-Min Kim, and Dae-Cheol Ko. Development of vacuum-assisted prepreg compression molding for production of automotive roof panels. *Composite Structures*, 213, 04 2019.

[7] Cristiano Fragassa. Marine applications of natural fiber-reinforced composites: A manufacturing case study. *Advances in Applications of Industrial Biomaterials*, pages 21–47, 2017.

[8] J Summerscales. Composites manufacturing for marine structures. In *Marine Applications of Advanced fiber-Reinforced Composites*, pages 19–55. Elsevier, 2016.

[9] Mihaela Raluca Condruz, Ion Malael, Ionut Sebastian Vintila, and Mihail Puscas Cernat. Manufacturing of advanced composite wind turbine blades for counter rotating vertical wind turbine. *Mater. Plast*, 57:45–56, 2019.

[10] Peter J Schubel. Cost modelling in polymer composite applications: Case study–analysis of existing and automated manufacturing processes for a large wind turbine blade. *Composites Part B: Engineering*, 43(3):953–960, 2012.

[11] SA Hadigheh and Sima Kashi. Effectiveness of vacuum consolidation in bonding fiber reinforced polymer (frp) composites onto concrete surfaces. *Construction and Building Materials*, 187:854–864, 2018.

[12] Bilal Zahid and Xiaogang Chen. Manufacturing of single-piece textile reinforced riot helmet shell from vacuum bagging. *Journal of composite materials*, 47(19):2343–2351, 2013.

[13] M Sreejith and RS Rajeev. Fiber reinforced composites for aerospace and sports applications. In *Fiber Reinforced Composites*, pages 821–859. Elsevier, 2021.

[14] Bethany Middleton. Composites: manufacture and application. *Design and manufacture of plastic components for multifunctionality: structural composites, injection molding, and 3D printing*, page 53, 2015.

[15] P Balakrishnan, Maya J John, L Pothen, MS Sreekala, and S Thomas. Natural fibre and polymer matrix composites and their applications in aerospace engineering. In *Advanced composite materials for aerospace engineering*, pages 365–383. Elsevier, 2016.

[16] Mauricio Cabrera-Ríos and José M Castro. The balance between durability, reliability, and affordability in structural composites manufacturing. *Polymer composites*, 28(2):233–240, 2007.

[17] Costas Soutis. fiber reinforced composites in aircraft construction. *Progress in aerospace sciences*, 41(2):143–151, 2005.

[18] A Brent Strong. *Fundamentals of composites manufacturing: materials, methods and applications*. Society of manufacturing engineers, 2008.

[19] T. Centea, L.K. Grunenfelder, and S.R. Nutt. A review of out-of-autoclave prepregs – material properties, process phenomena, and manufacturing considerations. *Composites Part A: Applied Science and Manufacturing*, 70:132–154, 2015.

[20] K-T Hsiao and D Heider. *Vacuum assisted resin transfer molding (VARTM) in polymer matrix composites*. Elsevier, 2012.

[21] Timotei Centea and Steven R Nutt. Manufacturing cost relationships for vacuum bag-only prepreg processing. *Journal of Composite Materials*, 50(17):2305–2321, 2016.

[22] James R. Krove Duane M. Rachal. Vacuum bagging process for fiber reinforced thermoplastics, US4915896A patent 1987.

[23] Gaurav Nilakantan and Steven Nutt. Reuse and upcycling of aerospace prepreg scrap and waste. *Reinforced Plastics*, 59(1):44–51, 2015.

[24] John M Hodgkinson. *Mechanical testing of advanced fiber composites*. Woodhead publishing, 2000.

[25] Anja Haschenburger and Niklas Menke. Sensor-based leakage detection in vacuum bagging. *The International Journal of Advanced Manufacturing Technology*, 2021.

[26] sang yoon Park, Chi Choi, Won Choi, and Seong Hwang. A comparison of the properties of carbon fiber epoxy composites produced by non-autoclave with vacuum bag only prepreg and autoclave process. *Applied Composite Materials*, 26, 02 2019.

[27] G Fernlund, J Wells, L Fahrang, J Kay, and A Poursartip. Causes and remedies for porosity in composite manufacturing. In *IOP conference series: materials science and engineering*, volume 139, page 012002. IOP Publishing, 2016.

[28] Anja Haschenburger and Jan Stüve. Influence of leaky vacuum bags on the quality of composite parts made from prepreg material. *Kunststoffe*, 2021.

[29] Flake C Campbell Jr. *Manufacturing processes for advanced composites*. elsevier, 2003.

[30] Philip E Irving and Costas Soutis. *Polymer composites in the aerospace industry*. Woodhead Publishing, 2019.

[31] Zvi Hashin. Analysis of composite materials—a survey. *Journal of applied mechanics*, 1983.

[32] Zvi Hashin and FW Wendt. Theory of composite materials. *Mechanics of composite materials*, pages 201–242, 1970.

[33] J Praveenkumara, P Madhu, TG Yashas Gowda, MR Sanjay, and Suchart Siengchin. A comprehensive review on the effect of synthetic filler materials on fiber-reinforced hybrid polymer composites. *The Journal of the Textile Institute*, 113(7):1231–1239, 2022.

[34] Victoria A Gafarova, Alexander Babin, Elvira R Gareeva, Karina N Abdrakhmanova, and Liliya N Lomakina. Influence of a filler on strength characteristics of the properties of a composite material based on epoxy resin. *Materials Today: Proceedings*, 11:252–257, 2019.

[35] Holm Altenbach, Johannes Altenbach, Wolfgang Kissing, Holm Altenbach, Johannes Altenbach, and Wolfgang Kissing. Classification of composite materials. *Mechanics of composite structural elements*, pages 1–14, 2004.

[36] Nikhilesh Chawla and Y-L Shen. Mechanical behavior of particle reinforced metal matrix composites. *Advanced engineering materials*, 3(6):357–370, 2001.

[37] Maria Cristina Tanzi, Silvia Farè, and Gabriele Candiani. *Foundations of biomaterials engineering*. Academic Press, 2019.

[38] Josh Kelly and Mohsen Mohammadi. Uniaxial tensile behavior of sheet molded composite car hoods with different fibre contents under quasi-static strain rates. *Mechanics Research Communications*, 87:42–52, 2018.

[39] Muhammad Pervaiz and Mohini M Sain. Sheet-molded polyolefin natural fiber composites for automotive applications. *Macromolecular materials and engineering*, 288(7):553–557, 2003.

[40] Jack R Vinson and Robert L Sierakowski. *The behavior of structures composed of composite materials*, volume 105. Springer, 2006.

[41] BD Agarwal, LJ Broutman, BD Agarwal, and LJ Broutman. *Analysis and performance of fiber composites Second edition*. John Wiley & Sons, 1990.

[42] AB Nair and R Joseph. Eco-friendly bio-composites using natural rubber (nr) matrices and natural fiber reinforcements. In *Chemistry, manufacture and applications of natural rubber*, pages 249–283. Elsevier, 2014.

[43] Kim L Pickering, MG Aruan Efendy, and Tan Minh Le. A review of recent developments in natural fibre composites and their mechanical performance. *Composites Part A: Applied Science and Manufacturing*, 83:98–112, 2016.

[44] Maria Ernestina Alves Fidelis, Thatiana Vitorino Castro Pereira, Otávio da Fonseca Martins Gomes, Flávio de Andrade Silva, and Romildo Dias Toledo Filho. The effect of fiber morphology on the tensile strength of natural fibers. *Journal of Materials Research and Technology*, 2(2):149–157, 2013.

[45] Costas Soutis. Carbon fiber reinforced plastics in aircraft construction. *Materials Science and Engineering: A*, 412(1-2):171–176, 2005.

[46] Han Gi Chae, Bradley A Newcomb, Prabhakar V Gulgunje, Yaodong Liu, Kishor K Gupta, Manjeshwar G Kamath, Kevin M Lyons, Sushanta Ghoshal, Chandrani Pramanik, Lucille Giannuzzi, et al. High strength and high modulus carbon fibers. *Carbon*, 93:81–87, 2015.

[47] Deborah Chung. *Carbon fiber composites*. Elsevier, 2012.

[48] Kenton R Kaufman and Kathie Bernhardt. Functional performance differences between carbon fiber and fiberglass prosthetic feet. *Prosthetics and Orthotics International*, 45(3):205–213, 2021.

[49] Philip F Chu, Shigeo Iwasawa, Philip L Schell, and Chia-Yu Lin. Carbon fiber versus glass fiber reinforcements: A novel, true comparison in thermoplastics. *Polymer Composites*, 42(11):6173–6181, 2021.

[50] M Jassal and S Ghosh. Aramid fibers-an overview. *Indian Journalof fiber and TExtile Research*, 27:290–306, 2002.

[51] Karlheinz Hillermeier. Prospects of aramid as a substitute for asbestos. *Textile research journal*, 54(9):575–580, 1984.

[52] Maria Mrazova. Advanced composite materials of the future in aerospace industry. *Incas bulletin*, 5(3):139, 2013.

[53] Deborah DL Chung. Polymer-matrix composites: Structure and processing. *Carbon Composites*, pages 161–217, 2017.

[54] Michael Elkington, D Bloom, C Ward, A Chatzimichali, and K Potter. Hand layup: understanding the manual process. *Advanced manufacturing: polymer & composites science*, 1(3):138–151, 2015.

[55] Syed Waheedullah Ghori, Ramengmawii Siakeng, Masrat Rasheed, Naheed Saba, and Mohammad Jawaid. The role of advanced polymer materials in aerospace. In *Sustainable composites for aerospace applications*, pages 19–34. Elsevier, 2018.

[56] Mohammad RM Jamir, Mohammad SA Majid, and Azduwin Khasri. Natural lightweight hybrid composites for aircraft structural applications. In *Sustainable composites for aerospace applications*, pages 155–170. Elsevier, 2018.

[57] FWJ Van Hattum, F Regel, and M Labordus. Cost reduction in manufacturing of aerospace composites. *Plastics, Rubber and Composites*, 40(2):93–99, 2011.

[58] D Cripps, TJ Searle, and J Summerscales. Open mold techniques for thermoset composites. *Comprehensive Composite Materials*, 2000.

[59] Alessia Serena Perna, Antonio Viscusi, Antonello Astarita, Luca Boccarusso, Luigi Carrino, Massimo Durante, and Raffaele Sansone. Manufacturing of a metal matrix composite coating on a polymer matrix composite through cold gas dynamic spray technique. *Journal of Materials Engineering and Performance*, 28:3211–3219, 2019.

[60] R Gauvin, M Chibani, and P Lafontaine. The modeling of pressure distribution in resin transfer molding. *Journal of reinforced plastics and composites*, 6(4):367–377, 1987.

[61] Akbar Shojaei, S Reza Ghaffarian, and S Mohammad-Hossien Karimian. Modeling and simulation approaches in the resin transfer molding process: A review. *Polymer Composites*, 24(4):525–544, 2003.

[62] Daniel C Davis and Thomas O Mensah. Fabrication and fatigue of fiber-reinforced polymer nanocomposites–a tool for quality control. *Nanotechnology Commercialization: Manufacturing Processes and Products*, pages 335–368, 2017.

[63] Dipen Kumar Rajak, Durgesh D Pagar, Pradeep L Menezes, and Emanoil Linul. Fiber-reinforced polymer composites: Manufacturing, properties, and applications. *Polymers*, 11(10):1667, 2019.

[64] Michael Moeller Jr, Conner Shane, and Ratneshwar Jha. Fabrication of composite laminates using a vacuum bag process. *Annual SSymposium on UUndergraduate RResearch EExperiences*, page 214, 2007.

[65] Christopher Williams, John Summerscales, and Stephen Grove. Resin infusion under flexible tooling (rift): a review. *Composites Part A: Applied Science and Manufacturing*, 27(7):517–524, 1996.

[66] Felice Rubino and Pierpaolo Carlone. A semi-analytical model to predict infusion time and reinforcement thickness in vartm and scrimp processes. *Polymers*, 11(1):20, 2018.

[67] Kerang Han, Shunliang Jiang, Chuck Zhang, and Ben Wang. Flow modeling and simulation of scrimp for composites manufacturing. *Composites Part A: Applied Science and Manufacturing*, 31(1):79–86, 2000.

[68] TH Hou and BJ Jensen. Evaluation of double-vacuum-bag process for composite fabrication. In *SAMPE 2004 Symposium and Exhibition*, 2004.

[69] Yasir Mujahid, Nabihah Sallih, and Mohamad Zaki Abdullah. A comparison of single-vacuum-bag and double-vacuum-bag methods for manufacturing high-quality laminated composites. In *Advances in Manufacturing Engineering: Selected articles from ICMMPE 2019*, pages 457–467. Springer, 2020.

[70] Muhammad Razlan Zakaria, Mohd Zharif Ahmad Thirmizir, M Shukur Zainol Abidin, Hazizan Md Akil, Mohd Firdaus Omar, Aslina Anjang Ab Rahman, Norlin Nosbi, and Norfaradina Ab Ghafar. A comparison process between wet lay-up, single vacuum bagging and double vacuum bagging toward natural fibre (palm, coconut and kenaf) reinforced epoxy composite laminates. *Fibers and Polymers*, 24(6):2107–2116, 2023.

[71] Nur Md Hanafiah Hafzareen, Abdul Rahim Othman, and Mark Ovinis. Significant effect of vacuum bagging processing on inter-laminar shear strength and voids of composite in oven cure. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 37(1):69–81, 2024.

[72] Soo Hyun Eum, Yun Hae Kim, Joong Won Han, Kook Jin Kim, Do Hoon Shin, Cheolmun Yim, and Riichi Murakami. A study on the mechanical properties of the honeycomb sandwich composites made by vartm. *Key Engineering Materials*, 297:2746–2751, 2005.

[73] Vigneshkumar Ramdas Tamakuwala. Manufacturing of fiber reinforced polymer by using vartm process: A review. *Materials Today: Proceedings*, 44:987–993, 2021.

[74] Tengfei Chang, Lihua Zhan, Wei Tan, and Shujian Li. Void content and interfacial properties of composite laminates under different autoclave cure pressure. *Composite Interfaces*, 24(5):529–540, 2017.

[75] Quazi TH Shubhra, AKM Moshiul Alam, and M Adul Quaiyyum. Mechanical properties of polypropylene composites: A review. *Journal of thermoplastic composite materials*, 26(3):362–391, 2013.

[76] Jose M Castro and Richard M Griffith. Press molding processes. In *Composites Engineering Handbook*, pages 493–526. CRC Press, 1997.

[77] J Ramôa Correia. Pultrusion of advanced fibre-reinforced polymer (frp) composites. In *Advanced Fibre-Reinforced Polymer (FRP) Composites for Structural Applications*, pages 207–251. Elsevier, 2013.

[78] SC Joshi. The pultrusion process for polymer matrix composites. In *Manufacturing techniques for polymer matrix composites (PMCs)*, pages 381–413. Elsevier, 2012.

[79] Niklas Minsch, Falko H Herrmann, Thomas Gereke, Andreas Nocke, and Chokri Cherif. Analysis of filament winding processes and potential equipment technologies. *Procedia CIRP*, 66:125–130, 2017.

[80] Pankar K Mallick. *Fiber-reinforced composites: materials, manufacturing, and design*. CRC press, 2007.

[81] YW Leong, S Thitithanasarn, K Yamada, and H Hamada. Compression and injection molding techniques for natural fiber composites. In *Natural Fibre Composites*, pages 216–232. Elsevier, 2014.

[82] Mei-po Ho, Hao Wang, Joong-Hee Lee, Chun-kit Ho, Kin-tak Lau, Jinsong Leng, and David Hui. Critical factors on manufacturing processes of natural fibre composites. *Composites Part B: Engineering*, 43(8):3549–3562, 2012.

[83] Carmen M González-Henríquez, Mauricio A Sarabia-Vallejos, and Juan Rodriguez-Hernandez. Polymers for additive manufacturing and 4d-printing: Materials, methodologies, and biomedical applications. *Progress in Polymer Science*, 94:57–116, 2019.

[84] Chee Kai Chua and Kah Fai Leong. *3D Printing and additive manufacturing: Principles and applications (with companion media pack)-of rapid prototyping.* World Scientific Publishing Company, 2014.

[85] JS Hayward and B Harris. The effect of vacuum assistance in resin transfer moulding. *Composites Manufacturing*, 1(3):161–166, 1990.

[86] Hoda Koushyar, Seyed Alavi-Soltani, Bob Minaie, and Melanie Violette. Effects of variation in autoclave pressure, temperature, and vacuum-application time on porosity and mechanical properties of a carbon fiber/epoxy composite. *Journal of Composite Materials*, 46(16):1985–2004, 2012.

[87] Nishant Verma, Rajeev Kumar, Sunny Zafar, and Himanshu Pathak. Vacuum-assisted microwave curing of epoxy/carbon fiber composite: an attempt for defect reduction in processing. *Manufacturing Letters*, 24:127–131, 2020.

[88] Flake C Campbell, Andrew R Mallow, and Clarles E Browning. Porosity in carbon fiber composites an overview of causes. *Journal of advanced materials*, 26(4):18–33, 1995.

[89] Michael R Wisnom, Tom Reynolds, and Nigel Gwilliam. Reduction in inter-laminar shear strength by discrete and distributed voids. *Composites Science and Technology*, 56(1):93–101, 1996.

[90] Sérgio Frascino Müller de Almeida and Zabulon dos Santos Nogueira Neto. Effect of void content on the strength of composite laminates. *Composite structures*, 28(2):139–148, 1994.

[91] Peter R Ciriscioli, George S Springer, and Woo Il Lee. An expert system for autoclave curing of composites. *Journal of composite Materials*, 25(12):1542–1587, 1991.

[92] Yizhuo Gu, Min Li, Zuoguang Zhang, and Zhijie Sun. Void formation model and measuring method of void formation condition during hot pressing process. *Polymer Composites*, 31(9):1562–1571, 2010.

[93] LK Grunenfelder and SR Nutt. Void formation in composite prepregs–effect of dissolved moisture. *Composites Science and Technology*, 70(16):2304–2309, 2010.

[94] FYC Boey and SW Lye. Effects of vacuum and pressure in an autoclave curing process for a thermosetting fibre-reinforced composite. *Journal of materials processing technology*, 23(2):121–131, 1990.

[95] FYC Boey and SW Lye. Void reduction in autoclave processing of thermoset composites: Part 1: High pressure effects on void reduction. *Composites*, 23(4):261–265, 1992.

[96] Vishwanath R Kedari, Basil I Farah, and Kuang-Ting Hsiao. Effects of vacuum pressure, inlet pressure, and mold temperature on the void content, volume fraction of polyester/e-glass fiber composites manufactured with vartm process. *Journal of composite materials*, 45(26):2727–2742, 2011.

[97] AP Mouritz. Ultrasonic and interlaminar properties of highly porous composites. *Journal of composite materials*, 34(3):218–239, 2000.

[98] MM Grigoriev, J Han, R Woo, and D Cheng. Effect of atmospheric pressure plasma treatment on surface porosity of ooa composites. In *Proceedings of the SAMPE 2013 Conference and Exhibition: Education and Green Sky—Materials Technology for a Better World, Long Beach, CA, USA*, pages 6–9, 2013.

[99] L Hamill, T Centea, G Nilakantan, and S Nutt. Surface porosity in out-of-autoclave prepreg processing: causes and reduction strategies. In *International SAMPE Technical Conference*, 2014.

[100] Lee Hamill, Timotei Centea, and Steven Nutt. Surface porosity during vacuum bag-only prepreg processing: Causes and mitigation strategies. *Composites Part A: Applied Science and Manufacturing*, 75:1–10, 2015.

[101] American Society for Testing and Pennsylvania) Materials (Filadelfia. Astm d3171-15: Standard test method for constituent content of composite materials. In *ISBN*. ASTM, 2015.

[102] Kirk D Tackitt and Shawn M Walsh. Experimental study of thickness gradient formation in the vartm process. *Materials and manufacturing processes*, 20(4):607–627, 2005.

[103] Bekir Yenilmez, Murat Senan, and E Murat Sozer. Variation of part thickness and compaction pressure in vacuum infusion process. *Composites Science and Technology*, 69(11-12):1710–1719, 2009.

[104] Murat Senan. *Monitoring of part dimensions and fiber content in Vacuum Assisted Resin Transfer Molding (VARTM) process*. PhD thesis, Koç University, 2007.

[105] Stephanie Gonçalves Nunes, Wanderley Ferreira de Amorim Jr, Andrea Manes, and Sandro Campos Amico. The effect of thickness on vacuum infusion processing of aramid/epoxy composites for ballistic application. *Journal of Composite Materials*, 53(3):383–391, 2019.

[106] Marcos Yutaka Shiino, Maria Odila Hilário Cioffi, Herman Cornelis Jacobus Voorwald, and Edmundo Cabral Ortiz. Tricot stitched carbon fiber reinforced polymer composite laminates manufactured by resin transfer molding process: C-scan and flexural analysis. *Journal of composite materials*, 47(14):1695–1703, 2013.

[107] ASTM Committee D-30 on Composite Materials. *Standard test method for short-beam strength of polymer matrix composite materials and their laminates*. ASTM International, 2006.

[108] Lawrence C Bank, T Russell Gentry, and Aaron Barkatt. Accelerated test methods to determine the long-term behavior of frp composite structures: environmental effects. *Journal of Reinforced Plastics and Composites*, 14(6):559–587, 1995.

[109] A.H. Turnbull. Leak detection and detectors. *Vacuum*, 15(1):3–11, 1965.

[110] M Köber and U Eberth. Laminieren für profis—ein trainingshandbuch für faserverbundwerkstoffe. *Auflage, Airbus Deutschland GmbH, Hamburg*, 2007.

[111] Albert Nerken. History of helium leak detection. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 9(3):2036–2038, 1991.

[112] Mass spectrometer – helium leak test. `https://www.applus.com/ca/en/what-we-do/sub-service-sheet/mass-spectrometer-helium-leak-test`. Accessed: 2024-07-29.

[113] Robert Brockmann. Ultra sniffer–new leak detection method. In *19th World conference on non-destructive testing*, 2016.

[114] Arne Tiddens, Marc Röger, Hannes Stadler, and Bernhard Hoffschmidt. A tracer gas leak rate measurement method for circular air circuits. *Flow Measurement and Instrumentation*, 47:45–53, 2016.

[115] T Gurunathan and A NoorulHaq. Innovative reliability excellence in helium gas leak testing by vacuum method. *International Journal of Engineering and Management Research (IJEMR)*, 6(1):503–511, 2016.

[116] Werner Große Bley. Methods of leak detection. *Handbook of vacuum technology*, pages 907–942, 2016.

[117] Hemi Sagi. Advanced leak test methods. *no date, ATC Inc*, pages 1–6, 2016.

[118] Jeff Sloan. Camx 2017 preview: Convergent manufacturing technologies, 2017.

[119] EDWARD Naranjo and SHANKAR Baliga. Expanding the use of ultrasonic gas leak detectors: A review of gas release characteristics for adequate detection. *International Gases & Instrumentation*, 3(6):24–29, 2009.

[120] Jan Steckel and Herbert Peremans. Ultrasound-based air leak detection using a random microphone array and sparse representations. In *SENSORS, 2014 IEEE*, pages 1026–1029. IEEE, 2014.

[121] Distran Switzerland. Ultra ce , speed up vacuum leak inspection. https://distran.swiss/en/ultra-ce-composite-edition/. Accessed: 2024-08-12.

[122] Xiao Liu, Jun Li, Jianjian Zhu, Yishou Wang, and Xinlin Qing. Cure monitoring and damage identification of cfrp using embedded piezoelectric sensors network. *Ultrasonics*, 115:106470, 2021.

[123] Junxiao Zhu, Liang Ren, Siu-Chun Ho, Ziguang Jia, and Gangbing Song. Gas pipeline leakage detection based on pzt sensors. *Smart Materials and Structures*, 26(2):025022, 2017.

[124] F Okosun, P Cahill, B Hazra, and V Pakrashi. Vibration-based leak detection and monitoring of water pipes using output-only piezoelectric sensors. *The European Physical Journal Special Topics*, 228:1659–1675, 2019.

[125] Jayant Sirohi and Inderjit Chopra. Fundamental understanding of piezoelectric strain sensors. *Journal of intelligent material systems and structures*, 11(4):246–257, 2000.

[126] Oliver Radestock Raino Petricevic. Messvorrichtung und verfahren zur ermittlung und ortung von undichtkeiten beim vakuumsackverfahren, DE102015010815A1 patent 2017.

[127] Suresh G. AdvaniFuping ZhouJustin Beech AlmsCharlotte C. Corlay. System and method of detecting air leakage in a vartm process, US7762122B2 patent 2010.

[128] Bölke J. Ucan H. Stefaniak D. Krombholz C. leakage detection, DE102011100096B4 patent 2014.

[129] Anja Tripmaker and Hakan Uçan. Detection of leakages on composite high performance structure parts. *Lightweight Design worldwide*, 10(3):46–49, 2017.

[130] Brett Aaron Cosco. *A novel application of Schlieren imaging: leakage localization in composite manufacturing vacuum bags*. PhD thesis, University of British Columbia, 2023.

[131] Gary S Settles and Michael J Hargather. A review of recent developments in schlieren and shadowgraph techniques. *Measurement Science and Technology*, 28(4):042001, 2017.

[132] James T Heineck, Daniel W Banks, Nathanial T Smith, Edward T Schairer, Paul S Bean, and Troy Robillos. Background-oriented schlieren imaging of supersonic aircraft in flight. *AIAA Journal*, 59(1):11–21, 2021.

[133] Finley B. MillerMichael E. BenneKenneth M. Dull. Device and method for detecting an air leak in a tool, US8438909B2 patent 2013.

[134] Finley B. MillerMichael E. Benne. Leak detection in vacuum bags, US20110079174A1 patent 2013.

[135] Anja Haschenburger and Clemens Heim. Two-stage leak detection in vacuum bags for the production of fiber-reinforced composite components. *CEAS Aeronautical Journal*, 10(3):885–892, 2019.

[136] Arne Hindersmann. Confusion about infusion: An overview of infusion processes. *Composites Part A: Applied Science and Manufacturing*, 126:105583, 2019.

[137] AI Haschenburger, L Onorato, MS Sujahudeen, DS Taraczky, A Osis, ARS Bracke, MD Byelov, FI Vermeulen, and EHQ Oosthoek. Computational methods for leakage localisation in a vacuum bag using volumetric flow rate measurements: Delft university of technology, german aerospace center. *Production Engineering*, 16(6):823–835, 2022.

[138] Yussuf Reza Esmaeili, Brett Cosco, and Homayoun Najjaran. Vacuum bag leak detection for resin infusion: an electric current–based analogy. *The International Journal of Advanced Manufacturing Technology*, 124(5-6):1775–1786, 2023.

[139] Yussuf Reza Esmaeili, Brett Cosco, Yifan Pan, and Homayoun Najjaran. Intelligent vacuum bagging leakage location prediction. *The International Journal of Advanced Manufacturing Technology*, pages 1–15, 2025.

[140] Christoph Brauer, Dirk Lorenz, and Lionel Tondji. Group equivariant networks for leakage detection in vacuum bagging. In *2022 30th European Signal Processing Conference (EUSIPCO)*, pages 1437–1441, 2022.

[141] Eric Brill, Jimmy Lin, Michele Banko, Susan T Dumais, Andrew Y Ng, et al. Data-intensive question answering. In *TREC*, volume 56, page 90, 2001.

[142] Xiudong Sun, Shoujie Li, and L James Lee. Mold filling analysis in vacuum-assisted resin transfer molding. part i: Scrimp based on a high-permeable medium. *Polymer Composites*, 19(6):807–817, 1998.

[143] MK Kang, WI Lee, and HT Hahn. Analysis of vacuum bag resin transfer molding process. *Composites Part A: Applied science and manufacturing*, 32(11):1553–1560, 2001.

[144] NC Correia, F Robitaille, AC Long, CD Rudd, P S˘ima´ c˘ek, and Suresh G Advani. Use of resin transfer molding simulation to predict flow, saturation, and compaction in the vartm process. *J. Fluids Eng.*, 126(2):210–215, 2004.

[145] Malcolm et al. Lane. Detection, monitoring, and management of gas presence, gas flow and gas leaks in composites manufacturing, US9862144B2 patent 2021.

[146] Gesang Nugroho, Indro Pranoto, and Novata Zaka Rohmana. Effect of breather type and vacuum pressure on the manufacturing of an unmanned aerial vehicle fuselage using vacuum bagging method. *AIP Conference Proceedings*, 1983(1), 07 2018. 040005.

[147] Timothy George Peter Gutowski. *Advanced composites manufacturing*. John Wiley and Sons, 1997.

[148] Henry Darcy. *Les fontaines publiques de Dijon*. Publiques de Dijon, 1856.

[149] Clayton R Paul. *Fundamentals of electric circuit analysis*. John Wiley & Sons, 2001.

[150] A Nerken. Versuche über die strömung von gasen durch lecke. *Vak. Techn*, 7:111, 1958.

[151] C Benjamin Nakhosteen and Karl Jousten. *Handbook of vacuum technology*. John Wiley & Sons, 2016.

[152] Wei-Yin Loh et al. Classification and regression tree methods. *Encyclopedia of statistics in quality and reliability*, 1:315–323, 2008.

[153] Vladimir N Vapnik. The nature of statistical learning. *Theory*, 1995.

[154] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer, 2003.

[155] David Duvenaud. *Automatic model construction with Gaussian processes*. PhD thesis, University of Cambridge, 2014.

[156] Stefanos Eleftheriadis, Tom Nicholson, Marc Peter Deisenroth, and James Hensman. Identification of gaussian process state space models. In *NIPS*, pages 5309–5319, 2017.

[157] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[158] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, and Humaira Arshad. State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11):e00938, 2018.

[159] Ding Wang, Haibo He, and Derong Liu. Intelligent optimal control with critic learning for a nonlinear overhead crane system. *IEEE Transactions on Industrial Informatics*, 14(7):2932–2940, 2017.

[160] Lingfei Wu, Peng Cui, Jian Pei, Liang Zhao, and Le Song. *Graph neural networks*. Springer, 2022.

[161] Jie Zhou, Ganqu Cui, Shengding Hu, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. Graph neural networks: A review of methods and applications. *AI open*, 1:57–81, 2020.

[162] Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

[163] James Joyce. Bayes' theorem. 2003.

[164] Maya Pishvar, Mehrad Amirkhosravi, and M Cengiz Altan. Magnet assisted composite manufacturing: A flexible new technique for achieving high consolidation pressure in vacuum bag/lay-up processes. *JoVE (Journal of Visualized Experiments)*, 135:e57254, 2018.

[165] Dhiren Modi, Nuno Correia, Michael Johnson, Andrew Long, Christopher Rudd, and Francois Robitaille. Active control of the vacuum infusion process. *Composites part A: Applied science and manufacturing*, 38(5):1271–1287, 2007.

[166] Pyi Phyo Maung, Tun Lin Htet, and GV Malysheva. Simulation and optimization of vacuum assisted resin infusion process for large-sized structures made of carbon fiber-reinforced plastics. In *IOP Conference Series: Materials Science and Engineering*, volume 709, page 022041. IOP Publishing, 2020.

[167] Yussuf Reza Esmaeili and Homayoun Najjaran. Optimal vacuum port placement in vacuum-assisted composite manufacturing for leakage detection and

localization: a hierarchical optimization approach. *Production Engineering*, pages 1–11, 2024.

[168] Ever J Barbero. *Introduction to composite materials design*. CRC press, 2010.

[169] Pierre-Marie Robitaille. Kirchhoff's law of thermal emission: 150 years. *Progr. Phys*, 4:3–13, 2009.

[170] G Anandalingam and Terry L Friesz. Hierarchical optimization: An introduction. *Annals of Operations Research*, 34:1–11, 1992.

[171] Omar Ben-Ayed and Charles E Blair. Computational difficulties of bilevel linear programming. *Operations Research*, 38(3):556–560, 1990.

[172] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer. Knn model-based approach in classification. In *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings*, pages 986–996. Springer, 2003.