UNIVERSITY OF CALGARY

Safe Transportation of Blended Hydrogen through Pipeline

by

Josmar Baruffaldi Cristello

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN MECHANICAL ENGINEERING

CALGARY, ALBERTA

OCTOBER, 2023

ABSTRACT

Pipelines are a crucial part of Canada's infrastructure, providing a cost-effective and safe way to transport oil and gas. As the world moves towards a potential hydrogen economy, pipelines will also play a significant role in transporting hydrogen. Rather than building new pipelines, which can be costly in terms of materials, permits, excavation, and labor, this study explores the viability of using existing pipelines for transporting blended hydrogen with natural gas. This is done through two primary aspects, the feasibility of transporting blended hydrogen with existing pipelines and developing a leak detection system for this blended gas.

The feasibility of blended hydrogen transportation is assessed through a steady-state gas hydraulic model, analyzing key operational metrics such as pressure, flow rate, and energy delivery. A virtual pipeline serves as a case study to evaluate the impact of transitioning from pure natural gas to blended hydrogen. It also delves into the trade-offs involved in selecting an optimal blend ratio.

Given the scarcity of leak data for hydrogen-natural gas pipelines, the study introduces a Real-Time Transient Model (RTTM). The model simulates the behaviour of mixed gas through a combination of Equations of State and thermodynamic databases and is built upon the continuity and momentum equations. This model simulates the leak dynamics of blended hydrogen gases, filling a critical data gap. Moreover, a leak detection system (LDS) is developed using a fusion of Convolutional Neural Networks (CNN) and Explainable artificial intelligence (AI) through Adaptive Neuro-Fuzzy Inference Systems (ANFIS). This innovative LDS framework overcomes the “black box” issue common in AI-driven systems, enabling reliable detection and localization of leaks.

The fusion of Explainable Machine Learning models and traditional AI techniques holds

promising implications for blended hydrogen pipelines and other fluid transportation systems. This study contributes to the ongoing efforts to enhance the safety and efficiency of hydrogen transportation, thereby mitigating economic and environmental impacts and addressing public concerns.

PREFACE

I, Josmar Cristello, have developed all contents in this thesis under the supervision of Dr. Simon Park and Dr. Ron Hugo, who have provided critical feedback and review of the publications listed below, which are used as part of the thesis.

Chapter 3 was part of a study conducted collaboratively with Jaehyun Michael Yang in 2021 and was in part submitted to the International Journal of Hydrogen Energy titled "Feasibility Analysis of Blending Hydrogen into Natural Gas Networks." This collaboration has advanced our understanding and added depth to the conclusions of this thesis.

Furthermore, sections of Chapter 4 describing the real-time transient model (RTTM) were developed together with Dr. Zhuoran Dang in a project funded through the Digital Innovations in Clean Energy (DICE) program from Alberta Innovates.

ACKNOWLEDGEMENTS

I extend my gratitude to Dr. Simon Park for his support and guidance at every stage of this research. Not only did he provide the initial opportunity for me to undertake this master's research, but he also remained a steady pillar of support over the past two years, helping me to overcome many challenges with his expertise.

I would also like to thank my co-supervisor, Dr. Ron Hugo, for providing invaluable insights that have considerably enriched the quality of this research. I am particularly thankful for his companionship during our field tests. His willingness to drive long distances and spend long hours has added distinctive and memorable experiences to my educational journey.

My appreciation extends to my colleagues at the Multifunctional Engineering, Dynamics, and Automation Laboratory (MEDAL) and the Pipeline Engineering Center (PEC). Special gratitude goes to Jaehyun Michael Yang, a mentor throughout my first year in the lab. His introduction to the laboratory environment and his continuous guidance during the challenging phases of my research were indispensable. I must also thank Robin Jeong; our afternoon coffees and lab get-togethers have been sources of relief and enjoyment during stressful times. Additional acknowledgments go to Zhuoran Dang and Haobin Peter Chen. Your companionship during countless hours of field tests made those experiences both productive and enjoyable. Finally, my thanks to Sina Rezvani are also in order. His assistance with hardware—often unfamiliar to me—was nothing short of crucial, and for that, I am sincerely grateful.

I express my deepest gratitude to my parents, Josmar and Marina. Your unwavering love and support have been the bedrock upon which all my accomplishments stand. You have given me the freedom to dream and the foundation to realize those dreams; I am eternally grateful for that.

DEDICATION

To my mother, who always believed in me. Your love and wisdom continue to guide me, even in your absence.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS AND NOMENCLATURE

Abbreviation	Definition
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
CFL	Courant-Friedrichs-Lewy
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
EOS	Equation of State
FIS	Fuzzy Inference System
HEE	Hydrogen Environment Embrittlement
ID	Inner Diameter
LDS	Leak Detection System
MAE	Mean Absolute Error
MAOP	Maximum Allowable Operating Pressure
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MoC	Method of Characteristics
NPW	Negative Pressure Wave
NTS	Notched Tensile Stress
OD	Outside Diameter
ODE	Ordinary Differential Equation
PDE	Partial Differential Equation

PID	Proportional Integral Derivative
PSA	Pressure Swing Adsorption
ROW	Right Of Way
RTTM	Real-Time Transient Model
SCADA	Supervisory Control and Data Acquisition
VFD	Variable-Frequency Drive

Symbol	Definition	Units
Q	Volumetric flow rate	m^3/day
f	Darcy friction factor	Dimensionless
P	Pressure	kPa
T	Temperature	°C
G	Gas gravity	Dimensionless
L	Length of pipe segment	km
Z	Gas compressibility factor	Dimensionless
D	Internal diameter	mm
LHV	Lower Heating Value	MJ/m^3
e_d	Absolute pipe roughness	mm
Re	Reynolds number	Dimensionless
J_t	Joule-Thompson Coefficient	°C/Pa
S_h	Hoop Stress	Pa
D_o	External Diameter	mm
t_p	Wall thickness	mm
$MAOP$	Maximum allowable operating pressure	kPa
S	Minimum yield strength	Pa
D_f	Design factor	Dimensionless
T_f	Temperature factor	Dimensionless
E_f	Quality factor of pipe weldment	Dimensionless
H_f	Material performance factor	Dimensionless

C_r	Compression ratio	Dimensionless
$Power$	Compression power	kW
M_f	Mass flow rate	kg/s
γ	Ratio of specific heats of gas	Dimensionless
η_a	Compressor adiabatic efficiency	Dimensionless
ΔH_c	Compressor head	m
η_c	Compressor efficiency	Dimensionless
α_{HE}	Proportional constant (HE Index)	Dimensionless
n	Material decaying constant	Dimensionless
H	Hydraulic head	m
g	Gravitational acceleration	m/s^2
V	Fluid velocity	m/s
ρ	Fluid density	kg/m^3
A	Cross-sectional area of pipe	m^2
z	Elevation	m
C_d	Discharge coefficient	Dimensionless
a	Wave Speed	m/s
Δx	Spatial Step (RTTM)	m
Δt	Temporal Step (RTTM)	s
K_p	Proportional Gain (PID)	Dimensionless
K_i	Integral Gain (PID)	Dimensionless
K_d	Derivative Gain (PID)	Dimensionless

$\bar{x}$	Gas Blend Mixture Composition	Dimensionless
δ	Reduced Mixture Density	Dimensionless
R	Molar Gas Constant	$J.mol^{-1}K^{-1}$
M	Molar Mass of Gas Mixture	g/mol
α	Helmholtz free energy	Dimensionless
$\boldsymbol{\mu}$	Viscosity	Pa.s
f_s	Pseudo-frequency	Hz
a_s	Scale	Dimensionless
f_c	Center Frequency of Wavelet	Hz
T_s	Sampling Period	s

CHAPTER 1: INTRODUCTION

As the world grapples with climate change, hydrogen emerges as a promising alternative to fossil fuels due to its zero carbon emissions during combustion. However, efficient distribution poses economic challenges, especially as constructing new pipelines is cost-prohibitive. Leveraging existing pipeline infrastructure, predominantly used for natural gas, offers a practical solution for large-scale hydrogen transportation. Although hydrogen has a lower volumetric energy density compared to natural gas, blending the two can facilitate a gradual transition to a more sustainable, hydrogen-based energy system, achieving a balance between energy delivery and greenhouse gas reduction. This approach, however, raises questions about pipeline compatibility and introduces new challenges in leak detection and safety, underlining the need for comprehensive solutions.

1.1 Background

Fossil fuels, currently a dominant energy source [1], contribute to greenhouse gas (GHG) emissions during combustion for heat generation, significantly impacting climate change. In contrast, hydrogen emerges as a promising alternative energy source due to its zero carbon emissions during combustion and electrochemical processes and high gravimetric energy density [2]. This positions hydrogen as a viable means to mitigate climate change by reducing GHG emissions and supplying efficient energy.

Hydrogen can be generated through electrolysis [3], wherein electricity breaks water into hydrogen and oxygen. This process can be reversed to produce electricity via the reaction between hydrogen and oxygen in fuel cells. Additionally, hydrogen can be converted to electrical energy through combustion [4]. These multiple advantages of hydrogen utilization have the potential to

significantly decrease carbon emissions and manage energy balance by harnessing excess energy from power generators or renewable energy production to generate hydrogen. This hydrogen can then be distributed as an energy carrier to meet energy demands dynamically.

Transporting hydrogen presents economic challenges for efficient distribution [5,6]. Pipelines are widely regarded as the most cost-effective method for transporting large volumes of hydrogen, with cost estimates ranging from $0.05 to $3 per ton, depending on distance [5]. However, constructing new pipelines is substantially expensive, posing a prohibitive barrier to entry. Moreover, a vast number of pipelines already exist. Canada alone has approximately 840,000 km of pipelines [7], with a significant majority carrying natural gas. Specifically, 57% of Alberta's 440,000 km of pipelines transport natural gas [8]. Consequently, leveraging the existing pipeline infrastructure emerges as the optimal solution for the large-scale distribution of hydrogen.

Table 1.1: Comparison of Key Physical Properties of Hydrogen and Methane. [10,11]

Properties	Hydrogen	Methane	Unit
Liquid Density	71	430-470	kg/m^3
Gas Density	0.08	0.65	kg/m^3
Gravimetric Energy Density	120	50	MJ/kg
Lower Heating Value (LHV)	10.2	34	MJ/m^3

The discussion of hydrogen transportation must be contextualized with respect to methane, the primary component of natural gas [9]. Transporting pure hydrogen entails several challenges, including hydrogen embrittlement and end-user compatibility. One of the most significant economic hurdles is that, despite its high gravimetric energy density, hydrogen possesses a volumetric energy density 6 to 8 times lower than natural gas due to its lower heating value (LHV).

This comparison is illustrated in Table 1.1.

As a result, maintaining the energy delivery capacity of natural gas would require increasing operating pressure and flow rates to levels that often surpass the operational limits of the existing pipelines [12]. A promising solution to this barrier lies in blending hydrogen with natural gas within existing pipeline networks. This approach facilitates a trade-off between energy delivery and GHG emissions reduction while capitalizing on the current infrastructure and allowing a gradual transition to hydrogen energy. While this is a beneficial approach and tradeoff, the transportation of blended hydrogen and natural gas through existing pipeline infrastructure brings new leak detection challenges.

The unique properties of hydrogen, such as its low density, high diffusivity, and small molecule size, can lead to increased leakage rates when compared to natural gas [13]. Furthermore, hydrogen leaks present additional safety hazards, as hydrogen has a wider flammability range [14], lower ignition energy [15], and higher flame speed [16]. Public perception of hydrogen safety is also negatively influenced by historical incidents, such as the Hindenburg disaster. These factors emphasize the need for reliable and accurate leak detection systems to detect and locate leaks in real time to ensure safe operation and minimize environmental and economic impacts.

Leak detection in pipeline systems is well-established, with various methods and technologies developed for typical fluids such as oil and natural gas. However, when it comes to blended hydrogen and natural gas pipelines, the landscape becomes more complex due to the unique properties of hydrogen and the need for comprehensive research in this area. Typically, leak detection systems (LDS) can be segregated into two broad categories: internally and externally based [17].

Externally based leak detection systems focus on monitoring the pipeline's exterior and

employ various technologies like fiber optic sensing, acoustic emission detection, and infrared imaging. Although some of these technologies may be suitable for hydrogen pipelines, they present considerable cost and practical implementation limitations for pipelines already in operation due to excavation costs. On the other hand, internally based systems use field instrumentation, such as flow, pressure, or temperature sensors, to monitor leaks from within the pipeline. While proven effective for oil and gas, the applicability of these internal systems for hydrogen transportation is yet to be confirmed. In both cases, the unique properties of hydrogen and blended hydrogen-natural gas mixtures present distinct challenges that require further study for effective leak detection.

Artificial Intelligence (AI) systems offer promising advancements in internally based leak detection systems (LDS), but they are often held back by transferability issues and explainability. Transferability refers to the ML model's capability to retain its accuracy when applied to different pipelines, particularly those transporting hydrogen blends with different properties like varying viscosity and density [17]. Explainability deals with the transparency of the ML algorithms, a critical factor for their adoption in high-stakes infrastructure settings. Existing models, especially those based on deep learning, are often criticized for their "black box" nature, which limits their practical use due to safety and compliance considerations.

1.2 Problem Statement

In the recent shift toward renewable energy sources, hydrogen has been considered a viable option, primarily due to its zero carbon emissions. However, this presents challenges, particularly regarding its efficient and cost-effective distribution. The potential of existing pipeline systems to transport hydrogen, especially in a blended state with natural gas, remains to be determined. A

comprehensive evaluation of these pipelines' compatibility with hydrogen and the advantages and potential drawbacks of this utilization needs to be investigated.

Data generation is another challenge in this upcoming field of study. Currently, there are only a limited number of hydrogen pipelines in operation, which hampers the availability of real-world data essential for testing and validating theoretical models. Predicting and preparing for potential issues that might surface in real-world operations is complicated without substantial data.

Adding to the complexity of the problem is the current leak detection systems, which often rely on hardware sensors. While these sensors can effectively detect leaks in pipeline systems, assuming widespread installation along existing pipelines is impractical. This approach would involve extensive and disruptive excavation, translating into prohibitive costs.

Artificial Intelligence (AI) systems present a promising alternative for leak detection in these blended pipelines. However, those systems introduce unique challenges in ensuring transferability between possible blend variability and lack of transparency in its classifications.

1.3 Objectives

This research aims to analyze and facilitate safe blended hydrogen transportation while using existing pipeline infrastructure from the point of view of infrastructure compatibility and leak detection.

For this, three primary goals are defined. The initial goal aims to solve the problem of limited data with hydrogen pipelines in operation by establishing accurate numerical models for pipeline gas flow during steady and transient states. With those models, datasets are generated for the latter two goals. Those focus on the study's tangible outcomes: Conducting a comprehensive feasibility analysis of using hydrogen blends with current pipelines and developing a transparent

and effective leak detection system (LDS) that relies only on existing field instrumentation. These interconnected objectives are visualized in Figure 1.1 and described in further detail in this chapter.

In tackling these aspects, the study aims to provide insights and solutions for safely transporting hydrogen via existing pipelines, aiding the transition to a more sustainable energy landscape.

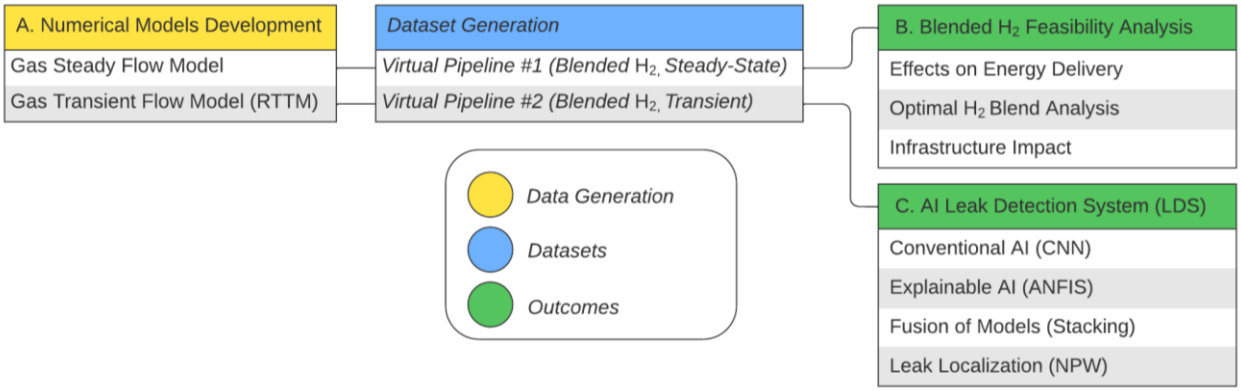


Figure 1.1: Diagram of Study Objectives.

A. Development of Numerical Models

The first objective is the development of two distinct numerical models for simulating the gas flow of blended hydrogen. The first model aims to capture the steady-state behaviour and the latter, transient flow phenomena. Those two models are used throughout the study to generate datasets.

The steady-state model leverages well-established equations in the gas hydraulic flow theory literature. This model evaluates the broader impacts of integrating hydrogen into the existing natural gas infrastructure. In this context, the steady flow model assists in evaluating a range of vital parameters, including but not limited to pipeline pressure, temperature, flow rates, and energy delivery.

On the other hand, the Real-Time Transient Model (RTTM) is developed based on the continuity and momentum equations and is designed to model the complex transient dynamics associated with leaks and other transient events in a hydrogen-natural gas blend. This model is fundamental for the leak detection aspects of the study since data from pipelines capable of transporting blended hydrogen in appreciable quantities are not yet widely available. The RTTM mitigates this gap by generating valuable data sets on a range of hydrogen blend scenarios.

The models serve as virtual proxies for the behaviour of real pipelines and allow for insights regarding the performance and safety of transporting blended hydrogen, even before dedicated pipelines become commonplace.

B. Analysis of the Feasibility of Blending Hydrogen through Pipelines

This study's second objective is to comprehensively analyze the feasibility of blending hydrogen into pre-existing natural gas infrastructure. This analysis is done from a hydraulic and engineering perspective.

Hydrogen has different physical and chemical properties than natural gas, affecting everything from the gas's behaviour under pressure to its interaction with pipeline materials. In terms of flow dynamics, for example, hydrogen's lower viscosity and density can influence the transport capacity of the pipeline system. Furthermore, the unique characteristics of hydrogen, including its wider flammability range and lower ignition energy, bring challenges that need to be addressed to ensure safety and operational reliability.

Consequently, one of the key aspects of this objective is to understand these challenges in depth. Such analysis includes studying the effects of hydrogen embrittlement on common pipeline materials, potential leakage rates, the behaviour of blended gas during combustion, the changes in

energy delivered, and pipeline capacities. This information can determine the necessary changes to existing infrastructure and guide the design of future pipelines.

The effects of switching from natural gas to a hydrogen blend are contextualized through a case study comprising transmission and distribution pipelines. The study simulates flow conditions within this virtual framework for a wide range of hydrogen blends, and this is used to illustrate the system's compatibility and identify required modifications for a comprehensive conversion to hydrogen. Finally, the analysis delves into this virtual pipeline system's optimal blend ratio.

In summary, this objective seeks to provide a well-rounded picture of the technical and environmental implications of hydrogen blending, providing invaluable insights into its practical feasibility and potential benefits.

C. Development of a Leak Detection System for Blended Hydrogen

The third objective of this research focuses on developing a leak detection system (LDS) for blended hydrogen and natural gas transportation. Instead of analyzing the compatibility, this tackles safe transportation by directly designing a system to handle potential leaks. The LDS is designed to identify and locate leaks – a crucial feature for informing response measures and minimizing the environmental impact.

Since most pipelines are buried and hard to access, installing additional external sensors can be impractical and expensive. Therefore, the proposed LDS relies only on pressure and flow rate sensors, which are present in most pipeline systems. This approach ensures the system is cost-effective and practical, avoiding significant hardware changes in existing pipeline infrastructures.

The Leak Detection System (LDS) combines the strengths of both conventional and

Explainable Artificial Intelligence (AI) by integrating a Convolutional Neural Network (CNN) with an Artificial Neural Fuzzy Inference System (ANFIS). This hybrid architecture employs ensemble learning, specifically model stacking with a logarithmic meta-model, to fuse the two models seamlessly. The CNN model performs this setup's computationally intensive leak detection task through frequency analysis. At the same time, the ANFIS model adds a layer of transparency, enabling better insight into the system's decision-making mechanics. Both models are combined with an ensemble learning technique (model stacking). Combining conventional and explainable AI addresses the "black box" issue commonly associated with AI systems, allowing for a more transparent and reliable LDS.

A virtual pipeline scenario is defined for training, validation, and demonstration. Mirroring a real-world pipeline situated in Alberta's geographical context, this virtual pipeline allows for the nuanced consideration of altitude effects on transient behaviour. The system is trained in this virtual pipeline by generating flow with the real-time transient model (RTTM) for scenarios with and without leak behaviour and with various hydrogen blends. A system that can handle multiple types of hydrogen blends directly addresses the transferability issue and offers a degree of future-proofing, as the blends used in the infrastructure may change over time.

In summary, developing an LDS specifically for blended hydrogen transportation sheds light on whether existing sensor technology can safely manage hydrogen, possibly introducing hydrogen into the current infrastructure with minimal changes. The ultimate aim is to ensure pipeline safety when transporting this blended gas in a potential economy that relies heavily on it.

1.4 Thesis Outline

The thesis organization is described in the following paragraphs. Chapter 2 presents the

literature review, diving into topics like hydrogen blending, pipeline compatibility with hydrogen, and diverse methods for leak detection. This chapter directly addresses technical aspects, such as the properties of hydrogen, the phenomenon of hydrogen embrittlement, specific hardware used for detecting leaks, and leak detection methods.

Chapter 3 consists of the study of the feasibility of introducing hydrogen blends into existing natural gas infrastructure. The chapter outlines the development of the steady-state gas hydraulic model and introduces a case study that serves as the empirical core of this feasibility analysis. Critical parameters such as energy transmission, flow rates, pressure effects, and infrastructural impacts are detailed. Finally, the chapter evaluates the trade-offs in selecting an optimal blend for the case study.

Chapter 4 details the development of the Gas Real-Time Transient Model (RTTM). It defines the transient flow equations and the numerical technique employed to solve them. In order to mirror real-world pipeline systems more closely, additional features such as intentional noise and randomized transient events are included in the model. Those features pose unique challenges for subsequent leak detection. The chapter also discusses the assumptions and limitations of the model and validates the results using commercial software benchmarks. Finally, it defines the second case study within the geometry of Alberta, laying the foundation for the leak detection system discussed in the subsequent chapter.

Chapter 5 describes the architecture of the proposed leak detection system (LDS) performance. First, it provides an overview of the system as a whole before diving into its core components. These include the conventional classifier (CNN), the explainable classifier (ANFIS), and the model fusion method via ensemble learning (Model Stacking). Leak localization is implemented through the Negative Pressure Wave (NPW) method. Finally, the LDS is tested

against the Alberta-based case study outlined in Chapter 4, and a comprehensive analysis of its effectiveness and limitations is presented.

Finally, Chapter 6 concludes the thesis, reflecting on the scientific outcomes resulting from this research and critically examining any limitations and assumptions made throughout the study. The discussion aims to provide a balanced perspective on the study's findings, considering its achievements and potential improvement areas.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the literature that forms the groundwork for this study. It begins by outlining the key physical properties of hydrogen and natural gas and the methodologies for blending and later separating these gases, providing context for properties referenced throughout the study. Next, the chapter delves into the direct impact of hydrogen on pipeline integrity, focusing mainly on hydrogen embrittlement (HE) and potential countermeasures. The chapter concludes by examining leak detection strategies discussing general methodologies and emerging techniques tailored explicitly for hydrogen.

2.1 Blending Hydrogen and Natural Gas

Hydrogen exhibits unique physical characteristics like lower volumetric density and a smaller molecular structure than natural gas. These differences can affect key operational parameters like flow rate and pressure when both gases are mixed and transported through pipelines initially designed for natural gas [18].

Additionally, achieving a uniform blend of hydrogen and natural gas along the pipeline is non-trivial [25]. If the two gases have significantly different densities, they may stratify, leading to different flow behaviour and leak characteristics. This can result in uneven energy distribution and operational issues with the pipeline.

Moreover, the process of blending the two gases is not trivial. It is important to ensure that the blended gas is homogeneous and uniform throughout the entire pipeline length [19]. If the two gases have significantly different densities, they may stratify, leading to different flow behaviour and leak characteristics. This can result in uneven energy distribution and operational issues with

the pipeline.

Depending on the specific needs of the end-user application, there might be a necessity to separate natural gas from hydrogen. Different applications often stipulate distinct purity standards. A range of separation technologies exists, including pressure swing adsorption (PSA), membrane separation, and cryogenic distillation. The selection of an appropriate method should not only match the required purity but also take into account factors such as cost, energy efficiency, and scalability.

2.1.1 Gas Properties

Natural gas typically consists of at least 95% methane, with other gases making up smaller fractions [9], as per Table 2.1. For this study, it is assumed that natural gas is entirely methane. Given the high methane content in typical natural gas, this assumption is reasonable and not expected to impact the study's findings substantially.

Table 2.1: Typical Natural Gas Composition. [9]

Gas Component	Methane (CH_4)	Ethane (C_2H_6)	Nitrogen (N)	Carbon Dioxide (CO_2)	Propane (C_3H_8)	Others
Volume (%)	95.3	2.16	1.86	0.44	0.19	0.05

Hydrogen and methane have different physical and chemical properties. These properties include molecular weight, density, viscosity, and specific gravity. Hydrogen has a molecular weight that is eight times lower than methane, which results in significantly lower values for most physical properties except viscosity; the viscosity of hydrogen is approximately 30% lower than methane at room temperature (25°C).

Table 2.2: Comparison Between Hydrogen and Methane Gas Properties at CNTP.

Properties	Hydrogen	Methane	Unit	Reference
Molecular Weight	2.016	16.043	kg/kmol	[15]
Density (Gas)	0.08	0.65	kg/m^3	[10]
Density (Liquid)	71	430-470	kg/m^3	[10]
Specific Gravity	0.0696	0.555	-	[20]
Viscosity (at 25°C)	$0.89 \cdot 10^{-5}$	$1.11 \cdot 10^{-5}$	Pa·s	[18]
Diffusion Coefficient in Air	0.61	0.16	cm^2/s	[21]
Energy Density	120	50	MJ/kg	[10]
Lower Heating Value (LHV)	10.2	34	MJ/m^3	[22]
Higher Heating Value (HHV)	12.5	37.8	MJ/m^3	[22]
Flammability Limits (in air)	4-75	5-15	Vol %	[10]
Detonation Limits (in air)	18.3-59	6.3-14	Vol. %	[18]
Minimum Spark Ignition Energy	0.02	0.29	MJ	[15]
Auto Ignition Temperature	858	810	K	[15]
Specific Heat Capacity	14.86	2.22	J/(g·K)	[23]
Specific Heat Ratio	1.383	1.308	-	[23]
Flame Temperature in Air	2318	2148	K	[23]
Solubility in Water	0.0016	0.025	kg/m^3	[23]

Hydrogen contains 2.4 times more energy per unit mass than methane. However, hydrogen per unit volume's lower heating value (LHV) is three times lower than methane due to its low density at standard conditions. This means that the energy content of the blended gas may be reduced with higher concentrations of hydrogen. From a safety perspective, higher concentrations of hydrogen tend to increase the risk of fire and explosion. Compared to methane, pure hydrogen has a much broader flammability range (5.3 times) and detonation limit range (7.1 times). It also has a significantly lower ignition energy (14.5 times lower), which makes it more prone to ignition and increases the risk of fire.

The low density of hydrogen also leads to higher diffusivity with other materials, which can result in more severe flammable and detonable events in confined spaces. Table 2.2 summarizes the properties of each gas at 20°C and 101.35 kPa in terms of physical properties, energy content, and risk.

2.1.2 Hydrogen Blending Methods

Several hydrogen blending systems are currently operational, as detailed in Table 2.3. Australia's HyP SA project [24] uses green hydrogen and ensures purification through specialized equipment between the electrolyzer and buffer tank. Another Australian project, ATCO's CEIH [25], blends up to 25% hydrogen using a sophisticated flow control system that employs gas chromatography and a controller for precise mixing. The HyDeploy project [26] employs a mixing loop and a gas control system to achieve optimal hydrogen concentrations in the natural gas pipeline. France's Jupiter 1000 project [27], commissioned in 2018, utilizes a 1.0 MW dual electrolyzer system for hydrogen generation and a methanation process for methane. The produced gases are then mixed to the desired concentrations before distribution. All projects employ varying technologies to ensure precise blending and monitoring of gas components.

Table 2.3: Examples of Hydrogen Blending Systems. [24–28]

Project	Country	Network	Electrolyser Capacity	Hydrogen Blend %	Others
HyP SA	Australia	Distribution	1.2 MW	5%	Uses green hydrogen
ATCO-CEIH	Australia	Distribution	0.15 MW	5-25%	GC and PID Control
HyDeploy	UK	Distribution	0.5 MW	20%	Measures Wobbe index
Jupiter 1000	France	Transmission	1.0 MW	6%	Injects hydrogen into transmission lines

Those blending systems aim to create well-mixed gas compositions to prevent stratification during transportation. Notably, the HyDeploy project [26] found that the gases do not stratify within the pipeline when a blended mixture of hydrogen and natural gas is transported under turbulent flow conditions. However, gas stratification can be a potential risk in laminar or stagnant flows due to density differences between the mixture components.

2.1.3 Blended Hydrogen Separation Methods

Separation of hydrogen from natural gas is not always necessary, but it is often required to ensure compatibility with end-use applications. The three methods commercially established to separate those gases is shown in Table 2.4. Those are the Pressure Swing Absorption (PSA), membrane separation and cryogenic distillation.

PSA is a technology with significant adoption, especially in refineries. This method consists of applying a pressure drop to segregate the hydrogen from its impurities and other components [31]. Because of the pressure drop required, the optimal location for implementing

PSA-based hydrogen separation is typically in pressure reduction stations. One of its main advantages is that it can output relatively high purity hydrogen [30], when compared to the other methods. On the other hand, it also for input gases with a relatively high hydrogen content (75-90%), which can be challenging for pipelines carrying blends with lower concentrations.

Table 2.4: Types of Separation Technologies. [29–31]

Category	Pressure Swing Absorption (PSA)	Membrane Separation	Cryogenic Distillation
Input Composition (H_2 mol%)	75-90	30-90	30-75
Output Purity (H_2 mol%)	>99	90-98	90-98
Processing Volume (Nm^3/h)	1000 – 10,000	<30,000	>10,000
Reliability (%)	95	100	Poor

Membrane separation exploits the pressure differential across a membrane to separate hydrogen. It uses specialized membranes with selective permeation for certain gas molecules over others, essentially allowing hydrogen to pass through while slowing other gases. This technology has the best range for different input compositions, from 30 to 90 mol% of hydrogen compositions. It also has the best reliability if unexpected shutdown happens. This can be, in part, attributed to its lack of mechanical components [30].

Finally, the cryogenic distillation separation process consists of using low temperatures to separate gases based on their volatility. This technology is particularly adept at handling very large volumes, of 10,000 Nm^3/h and above, but it also has a lower reliability [30]. Additionally, it has a relatively high energy consumption to maintain the low temperatures and to compress and cool

the gas mixture [31]. This is significant since it can even worsen further the problem of reduced energy content of hydrogen when compared to natural gas.

2.2 Hydrogen Influence on Pipelines

A significant concern in using hydrogen for pipeline transport is hydrogen embrittlement (HE). In this process, hydrogen permeates material lattices, weakening them and making them more susceptible to stress-induced cracking [32]. Additionally, hydrogen is more prone to leaks than methane and poses detection challenges due to its colorless and odorless nature [33]. These properties also make it hazardous for operators who lack specialized detection equipment.

This section explores the phenomenon of hydrogen embrittlement, the leak behaviour, and its consequences. Naturally, Leak detection should be in place no matter which type of gas is being transported by pipelines, and leak detection technologies will be discussed in the following section.

2.2.1 Hydrogen Embrittlement

Hydrogen embrittlement can reduce key mechanical properties like tensile ductility, fatigue resistance, and fracture toughness, rendering materials more susceptible to cracking [32]. According to a US Department of Energy report, most natural gas pipelines—99.7% of transmission lines and 50.4% of distribution lines—are composed of carbon steel materials [34]. In the context of repurposing these networks for hydrogen transport, assessing the compatibility of existing materials with hydrogen becomes imperative.

Specific carbon steels, including ASTM A 106 Grade B, ASTM A 53 Grade B, and API 5L Grades X42 and X52, have been deemed appropriate for hydrogen pipelines up to 14 MPa as per ASME B31.12 guidelines [32]. Table 2.5 offers an overview of the materials and operating

pressures in existing hydrogen pipelines. This data provides practical insight into the material and pressure considerations in the current hydrogen infrastructure.

Table 2.5: Operating Pressure for Hydrogen Pipelines Worldwide.

Location	Length (Km)	Diameter (mm)	Pressure (MPa)	Material	Operator	Reference
Texas, US	560	150 - 305	2.4 - 13.1	Steel Pipe	Air Products	[22]
Westliche, US	367	203 - 254	2.4 - 13.1	Steel Pipe	Air Products	[22]
Louisiana	-	150,203,305	2.4 - 13.1	Steel Pipe	Air Products	[22]
Los Angeles	19	150,254,305	2.4 - 13.1	Steel Pipe	Air Products	[22]
Texas, US	105	203	2.4	API 5L Grade B	Air Liquide	[22]
Texas, US	55	36	5.1	API 5L X60	Air Liquide	[22]
Netherlands	879	304.8	6.5 - 10	Seamless Carbon Steel	Air Liquide	[35]
UK	15	-	30	Carbon Steel	ICI Billingham	[35]
Germany (Leuna)	100	-	2 - 5	API 5L L290	Linde	[22]
France, Belgium	1040	100	9.7	API 5L X52	Air Liquide	[22]

The Hydrogen Environment Embrittlement (HEE) Index is a quantitative measure to assess a material's susceptibility to hydrogen embrittlement. This index compares a material's Notched Tensile Strength (NTS) when exposed to hydrogen against its NTS when exposed to air or helium. The calculation for the HEE Index is achieved through the following equation:

$$\text{HEE Index} = \text{NTS Ratio} = \text{NTS in Hydrogen}/\text{NTS in Air or Helium} \quad (2.1)$$

The HEE Index is a numerical value that falls from 0 to 1, where higher numbers signify a reduced likelihood of hydrogen embrittlement. According to NASA's report [36], Table 2.6 categorizes different levels of severity based on the HEE index.

Table 2.6: Categorization of Hydrogen Embrittlement Severity for Industrial Materials. [36]

HE Severity	HEE Index (NTS Ratio)	Description
Negligible	1.0 – 0.97	Materials can be utilized in the specified hydrogen pressure and temperature range with fracture and crack growth analysis in hydrogen
Small	0.96 – 0.90	
High	0.89 – 0.70	Limited application with fracture and crack growth analysis in hydrogen
Severe	0.69 – 0.50	Not recommended for usage
Extreme	0.49 – 0.0	

Table 2.7 outlines how common industrial materials fare in terms of their HEE Index, measured at specific hydrogen pressures of either 6.9 MPa or 68.9 MPa [32,36]. Notably, ferritic steels, standard in current gas pipelines, vary significantly in their susceptibility to hydrogen embrittlement when subjected to pressures up to 6.9 MPa at room temperature. In contrast, ductile substances like stainless steel varieties 304 and 316 and copper-based compounds show minimal HEE severity even when exposed to high-pressure conditions of 68.9 MPa. These materials also tend to resist hydrogen embrittlement in gaseous and liquid forms.

Table 2.7: HE Index for Commercialized Materials. [36]

Material		Lab-scale HE Evaluation			Material Compatibility with Hydrogen [37]		Reference
		Hydrogen Pressure	HE Severity	HEE Index (NTS[2])	Gas	Liquid	
Ferritic Steels (Carbon steels and low alloy steels)	A106-Gr. B	6.9 MPa	High	-	Acceptable	Not Acceptable	[32,36]
	A516	6.9 MPa	High	0.83	Acceptable	Not Acceptable	[32,36]
	API 5L X42	6.9 MPa	High	-	Acceptable	Not Acceptable	[32]
	API 5L X52	6.9 MPa	High	0.86	Acceptable	Not Acceptable	[32,36]
	API 5L X60	6.9 MPa	Small	0.92	Acceptable	Not Acceptable	[32,36]
	API 5L X65	6.9 MPa	Small	0.94	Acceptable	Not Acceptable	[32,36]
	API 5L X70	6.9 MPa	Small	0.90	Acceptable	Not Acceptable	[32,36]
Austenitic Steels	304L	68.9 MPa	High	0.87	Acceptable	Acceptable	[32,36]
	316	68.9 MPa	Negligible	1	Acceptable	Acceptable	[32,36]
Copper Based	Copper (OFHC[1])	68.9 MPa	Negligible	1	Acceptable	Acceptable	[32,36]
	70-30 Brass	68.9 MPa	Negligible	-	Acceptable	Acceptable	[32]

1: OFHC: Oxygen-free high thermal conductivity.

2: NTS Ratio is calculated in hydrogen and helium environments.

Hydrogen embrittlement is a complex issue governed by complex mechanisms. However, many materials have been evaluated for their resistance to embrittlement, and multiple guidelines exist to assist in selecting compatible materials. The best approach is to opt for materials with a known low vulnerability to embrittlement and to adhere to the internal pressure limitation. Internal coatings can be a viable alternative if the choice of material is not within the engineer's control, such as when repurposing existing natural gas pipelines.

2.2.2 Hydrogen Leak Behaviour

Hydrogen leakage in an open area is not typically a concern due to its highly diffusive characteristics. When burning, hydrogen flames release only 10% of the heat produced by hydrocarbon flames, resulting in minimal damage to the surrounding area [38]. Additionally, the extremely low density of hydrogen causes it to become buoyant when released into the air. Hydrogen is 14 times lighter than air and 57 times lighter than gasoline vapor [33], making it relatively safe to work in open-air spaces as it will dilute to non-flammable concentrations quickly [39].

The use of hydrogen in confined spaces can pose a higher risk due to its wide flammability range and significantly lower minimum spark energy when compared to natural gas (see Table 2.2). It can potentially cause damages associated with detonation and flammability [40]. If proper measures and design decisions are not implemented, hydrogen can accumulate quickly in confined spaces and lead to various hazards, such as initiating a fire, explosion, or even becoming an asphyxiant at high concentrations.

Ensuring proper ventilation when working with hydrogen in enclosed spaces is essential to address these risks. A report from the HySafe project recommends ventilation as one of the most effective safety measures for avoiding dangerous explosive conditions in hydrogen-treated facilities and guides the optimal design of a ventilation system in these facilities [41].

Another potential challenge with hydrogen is the potential stratification of a gas mixture containing hydrogen due to the significant difference in molecular weights between hydrogen and other gases. This could lead to increased safety risks, such as in the case of a blended gas leak event that ends up being primarily composed of hydrogen.

The nature of gas stratification in a mixture can vary depending on the flow regime.

Multiple studies note that blended gases mix well in turbulent flow [19,42]. A particular study noted that the mixture becomes homogenous within a short distance from the mixing point [26]. In contrast, laminar flow initially stratifies the mixture but becomes homogenous after a certain length. This critical distance can be roughly estimated by multiplying the nominal diameter by 4,000 [42]. Therefore, paying particular attention to operating pipelines with lower Reynolds numbers is important to minimize the risk of gas stratification.

Mejia et al. [13] conducted leak tests with blended hydrogen to assess the difference in leak rates of each gas component in the mixture. The tests were conducted with fully mixed 95% natural gas and 5% hydrogen at pressures up to 417 kPa (60.5 psi). The results showed that the difference in leak rates for each gas was minimal (around 1%). While these results are encouraging, it is of note that they were conducted at relatively low pressure and started from a fully mixed state.

Another study examined the relative size of gas leaks under different flow regimes (laminar, turbulent, and sonic), with leak rates estimated using the gases' dynamic viscosities, molar masses, and gas densities [43]. The results in Table 2.8 indicate that the estimated leak rate was significantly higher for hydrogen in all cases. For example, in the case of sonic flow (typical of leaks in high-pressure lines), hydrogen would leak volumetrically 2.8 times more than methane. However, considering energy density, the energy contained in such a hydrogen leak would be 0.88 times that of a similar methane leak.

Table 2.8: Leak Rate by Different Types of Flow Regimes. [43]

Properties	Hydrogen	Methane	Unit
Subsonic Flow (Laminar)	1.3	1	Relative leak rates (Volumetric)
Subsonic Flow (Turbulent)	2.8	1	
Sonic Flow	2.8	1	

2.2.3 Odorization

Another potential approach to hydrogen leak detection is gas odorization, which involves adding chemical odorants to the gas that the human olfactory sense can detect. This method is already widely used for natural gas, but the odorants used for natural gas are not suitable for hydrogen due to their significantly different densities and dispersion rates. While there are currently no known commercially usable odorants for hydrogen, several studies have evaluated odorants as a potential method for detecting hydrogen.

A research group in the Netherlands, consisting of Gasunie Transport Services (GTS) and Netbeheer, evaluated three odorants (THT, Spotleak 1001, Gasodor S-Free) at hydrogen concentrations ranging from 0% to 100%. These odorants could be used for a mixture of natural gas and hydrogen in distribution lines [28].

The Hy4Heat project in the United Kingdom also evaluated the hydrogen odorization performance of five chemical compounds across six sectors: Health and environment, olfactory, pipeline, flame boiler, fuel cell, and economic [28]. According to the evaluation results, the five odorants can be used for pure hydrogen except in fuel cell applications due to sulfur poisoning. The potential odorants for hydrogen gas are summarized in Table 2.9.

The odorants selected in this table were observed to be non-corrosive and metallurgically

intact. Their functionality was also proven through an olfactory test that evaluated their detectability at a minimum of one-fifth of the lower flammability limit of the gas composition and compatibility with hydrogen. In addition, the dilution effect by adding odorants is not foreseen for the listed odorants. All the odorants can be applied to hydrogen-serviced pipelines for general purposes. Spotleak 1001 is particularly suitable as a hydrogen odorant as it is broadly used in European gas networks and is low-cost with no harmful effects on pipelines or other appliances [28].

Table 2.9: Odorants for Use with Hydrogen Gas. [28]

Compound Name	Sulfur	Health	Olfactory Test	Corrosive after combustion	Material Degradation	Voltage Loss[3] [mv]	Fuel Cell Application
Spotleak 1001[1]	×	×	Pass	×	×	460 ±2	Purification Required
Standby Odorant 2[2]	×	×	Pass	×	×	40±2	O
THT[3]	×	×	Pass	×	×	225±2	Purification Required
Gasodor-S-Free[4]	×	×	Pass	×	×	10±2	O
Norbornene[5]	×	×	×[*6]	×	×	5±2	O

[1] Spotleak 1001: 2-Methyl-propanethiol 78%, Dimethyl Sulfide 22%

[2] Standby Odorant2: Odorant NB 34%, Hexane 64%

[3] Odorant THT: Tetrahydrothiophene – 100%

[4] Gasodor-S-Free: Methyl Acrylate 37.4%, Ethyl Acrylate 60.1%, 2-Ethyl-3-Methylpyrazine 2.5%

[5] Norbornene: 5-Ethylidene-2-Norbornene

[6] Not distinguishable from sulfur and oil

2.3 Leak Detection Methods

Leak detection in pipeline systems is well-established, with various methods and technologies developed for typical fluids such as oil and natural gas. However, when it comes to blended hydrogen and natural gas pipelines, the landscape becomes more complex due to the unique properties of hydrogen and the lack of comprehensive research in this area. The American Petroleum Institute Recommended Practice 1130 (API RP 1130) [17] provides guidelines for leak detection systems (LDS) and classifies those systems into two categories: internally based and externally based.

2.3.1 Externally Based Leak Detection

Externally based leak detection systems focus on monitoring the pipeline's exterior and can be divided into two primary categories: inspection-based and discrete hardware-based methods.

Inspection-based methods often rely on intermittent patrols, including direct human observation for liquid leaks or canine detection for gases and camera-based systems for above-ground pipelines [44]. However, these traditional methods encounter several challenges when applied to hydrogen. Because hydrogen is colorless [33], visual inspection techniques are ineffective unless they incorporate additional spectral technologies like LIDAR or IR. Similarly, olfactory-based inspections are unsuitable for hydrogen due to its odorless nature [33] unless odorants are added to the gas. Thus, these methods generally lack the capability for real-time monitoring and face particular limitations regarding hydrogen detection.

On the other hand, discrete hardware-based methods are more conducive to real-time leak monitoring and include technologies such as acoustic sensors, fiber-optic sensors, and gas

composition detectors [44]. While those are typically very effective, their installation can be prohibitively expensive, and this gets aggravated for pipelines that are already in operation and require extensive excavation for the installation.

Additionally, the compatibility of these detection methods with hydrogen remains an open question due to the gas's unique properties, such as high mobility, small molecular size, and high diffusivity with other materials. While comprehensive testing is required to verify their efficacy with hydrogen-blended gas, some studies indicate that minor sensor recalibrations may resolve these issues [29].

As for specialized hardware-based methods for hydrogen, they could serve as practical solutions for high-risk areas, albeit with the same concerns related to excavation costs and overall expense. Four types of sensors are employed for hydrogen detection: semiconductor metal oxide, electrochemistry, catalytic bead, and thermal conductivity [45–47]. These are shown in Table 2.10.

Table 2.10: Comparison of Commercially Available Hydrogen Detecting Sensors [45–47].

Sensor Type	Accuracy	Measuring Range	Response Time (t_{90})	Cost	Features
Semiconductor Metal Oxide	± 10-30 %	0 – 1000 ppm	<20 /s	$100 - $500	- Low Cost - Dependence on Humidity and Temperature
Electrochemical	< ± 4%	0 – 20000 ppm	<90 /s	$300 - $1200	- Good Selectivity to Hydrogen - Narrow Temperature Range
Catalytic Bead	< ± 5%	0 – 100 % H_2	<30 /s	$500 - $4,000	- Wide Temperature Range - No Hydrogen Selectivity
Thermal Conductivity	± 0.2%	0 - 100% H_2	<10 /s	<$25,000	- High Accuracy - Cross Sensitive to He

Semiconductor Metal Oxide (MOx) sensors are favored in industrial settings for their cost-effectiveness, though they suffer from limited accuracy and are influenced by environmental factors like humidity and temperature [45,48]. Electrochemical sensors provide better hydrogen selectivity and higher accuracy but are constrained by their narrow operational temperature range and slower detection speed. Catalytic sensors can detect a broad range of hydrogen concentrations and function well at different temperatures, but they come at a higher cost and lack hydrogen specificity.

On the other hand, thermal conductivity sensors boast the most comprehensive measurement range and highest accuracy but can yield false readings due to cross-sensitivity with helium (He). Given the need for both affordability and efficiency in industrial applications, semiconductor-based methods for detecting hydrogen leaks could offer a practical solution.

Due to the highly diffusive nature of hydrogen, the location of sensors is critically important to detect leaks accurately. Moreover, further study is needed to investigate how these sensors behave for blended hydrogen and natural gases. While some research has been conducted on hydrogen leak hardware development, studies specifically focusing on sensors for blended hydrogen and natural gas systems are comparatively sparse.

2.3.2 Internally Based Leak Detection

In contrast to externally based systems, internally based systems fall into two primary categories: Internal Mobile Sensing Systems and Computational Pipeline Monitoring (CPM) [44].

Internal Mobile Sensing Systems predominantly consist of Pipeline Inspection Gauges (PIGs). These sophisticated mobile devices traverse the pipeline, assessing its overall integrity, including factors like corrosion, right-of-way (ROW), and leaks [49]. Although PIGs provide invaluable data, they come at a significant cost and are typically deployed at infrequent intervals,

often spanning years.

Computational Pipeline Monitoring (CPM) leverages existing field instrumentation commonly found in most pipeline networks. As the main focus of CPM is on software and data analytics, it is more straightforward to implement, especially when the required field instrumentation is already in place. Given this advantage, developing CPM hydrogen-compatible methods could be especially valuable.

While these methods have demonstrated effectiveness with oil and gas, their compatibility with hydrogen remains uncertain. Further study is necessary to determine how these systems can be adapted for blended hydrogen and natural gas pipelines and to address potential challenges associated with hydrogen's unique properties.

Studies have been conducted on applying machine learning (ML) algorithms in CPM for traditional fluids such as oil and natural gas with significant potential for detecting leaks [50–52]. However, they are often held back by two critical limitations: transferability and explainability.

Transferability refers to the ability of an ML-based LDS model to maintain its accuracy and reliability when applied to a different pipeline network. Existing ML algorithms are typically trained on data from oil and natural gas pipelines, which may not accurately represent the unique properties of hydrogen. For instance, hydrogen's lower viscosity [18], higher diffusivity [24], and more rapid leak rate [43] can result in different leak signatures compared to traditional pipeline fluids. The varying hydrogen and natural gas blending ratios across different pipelines further highlight the need for robust ML models capable of adapting to diverse blend compositions.

Conversely, explainability concerns the transparency and interpretability of the ML algorithms' decision-making process. Deep learning models, in particular, are often criticized for their opacity, which can hinder their adoption in critical infrastructure like pipelines. Operators

and regulators require a clear understanding of the underlying reasoning behind the predictions made by LDS models for safety and compliance purposes. Moreover, comprehending the rationale behind a classification enables an operator to make informed decisions on whether an alert is false or legitimate, further improving the system's overall effectiveness.

2.4 Summary

Integrating hydrogen into pre-existing natural gas pipelines offers a cost-effective solution for transporting large volumes of hydrogen over long distances without building new infrastructure. Nevertheless, hydrogen's distinct physical attributes, including its lower volumetric density and unique molecular size, present challenges when transported blended with natural gas [18]. These differences may necessitate adjustments in operational parameters like flow rates and pose safety concerns for pipelines initially engineered for natural gas. Consequently, to ensure the secure and efficient transportation of hydrogen-natural gas blends, it is critical to evaluate and modify the original design specifications of pipelines and compressors to accommodate the unique requirements of these mixed gases.

Hydrogen's distinct characteristics, such as lower viscosity [18], higher diffusivity [24], and faster leak rate [43], present additional complexities in leak detection. Though effective, traditional hardware-based leak detection methods are often prohibitively expensive and impractical for retrofitting into existing pipelines. Computational Pipeline Monitoring (CPM) methods, on the other hand, offer a viable solution that could integrate well with current field instrumentation. While CPM systems have proven effective for conventional fluids like natural gas, their suitability for handling blended hydrogen is not yet fully understood. Hence, there is a desire to develop effective CPM-based leak detection systems that integrate seamlessly with existing instrumentation and are compatible with blended hydrogen and natural gas mixtures.

CHAPTER 3: Blended Hydrogen Feasibility Study

The transition from natural gas to hydrogen blends within existing pipeline networks presents a complex engineering challenge. A steady-state gas hydraulic model is developed to analyze this transition challenge. This mathematical model is based on well-established gas hydraulic flow theory and serves as a tool to explore the effects of different blends on factors such as the energy transmission of the pipeline, the flow rate, the maximum allowable operating pressure (MAOP), as well as derived effects such as temperature shifts due to joule-Thomson effect, and friction factors.

The model then simulates flow in a case study pipeline system composed of distribution and transmission pipelines. This system illustrates the feasibility of switching from natural gas to blended hydrogen in practice. The consequences of the transition are analyzed by identifying which hydrogen compositions can be accommodated without modifications and those that would necessitate changes—ranging from compressor selection to energy throughput adjustments. The study concludes by analyzing the trade-offs in selecting an optimal hydrogen blend for transportation.

3.1 Steady-State Gas Hydraulic Model

The steady-state gas hydraulic model simulates the characteristics of gas flow with the universal gas flow equation, the thermal effects are modeled with the Joule-Thompson effect, and the friction factor is calculated with the Colebrook-White equation. This section describes the theoretical equations used to develop the model.

Since one of the biggest concerns with transporting hydrogen is economical since it has less energy per volume compared to methane, the model is developed with two main inputs as design factors: The hydrogen blend (The ratio of methane and hydrogen) and the energy transmission. The energy transmission is defined relative to the transmission achieved by pure methane transportation.

3.1.1 Gas Hydraulics

Many factors, such as pressure, temperature, and the inherent physical properties of the pipeline, like length and diameter, determine the characteristics of gas flow within pipelines. The universal gas flow equation can encapsulate the fundamental physics governing this flow under steady, isothermal conditions [53], represented in equation (3.1).

$$P_1^2 = P_2^2 + \left(\frac{Q \cdot P_b}{1.1494 \cdot 10^{-3} T_b \cdot D^{2.5}}\right)^2 G \cdot T_f \cdot L \cdot Z \cdot f \tag{3.1}$$

where Q represents the volumetric flow rate (m^3/day), f stands for the Darcy friction factor (dimensionless), P_b signifies the atmospheric pressure (kPa), T_b corresponds to the atmospheric temperature (K), P_1 and P_2 are the absolute pressure (kPa) at upstream and downstream points respectively, G denotes gas gravity (dimensionless), T_f is the mean gas flowing temperature (K), L is the pipe length (km), Z stands for the gas compressibility factor at the flowing temperature (dimensionless), and D is the pipe's inner diameter (mm).

3.1.2 Energy Transmission

Energy transmission, or energy delivery, is one of the most important factors when considering the transition from natural gas to blended hydrogen. As mentioned in Table 2.2, even

though pure hydrogen has 2.4 times more energy per unit mass if compared to methane, it also has a 3 times smaller lower heating value (LHV). In other words, the higher the concentration of hydrogen in the blend, the lower the amount of energy that will be delivered if compared to methane due to its lower volumetric energy content. On the other hand, a blend of hydrogen and natural gas will be a trade-off and represent a less drastic decrease in energy transmission when compared to pure methane.

$$E_{deliv} = Q \cdot LHV_X \tag{3.2}$$

where Q represents the volumetric flow rate (m^3/day) and LHV_X represents the LHV (MJ/m^3) of the gas blend.

3.1.3 Flow Rate

The desired energy transmission and the hydrogen concentration in the blend influence the flow rate. In other words, higher energy transmissions require higher flow rates. The flow rate required to sustain a certain level of energy transmission can be derived using equation (3.3).

$$Q_r = \frac{Q_i \cdot LHV_{CH_4}}{LHV_B}, (i = 1, 2) \tag{3.3}$$

where Q_r refers to the targeted volumetric flow rate, Q_i stands for the initial volumetric flow rate (defined in Table *3.1*), i indicates the category of the pipeline (transmission or distribution), and LHV_{CH_4} and LHV_B denotes the lower heating values for methane and the blended gas (comprising methane and hydrogen), respectively.

3.1.4 Friction Factor

The friction factor will directly depend on the flow regime. The friction factor is solely a function of the Reynolds number for the laminar flow regime, as demonstrated in equation (3.4).

The turbulent flow regime, on the other hand, can be characterized by the Colebrook-White equation [53], illustrated in equation (3.5). This equation considers the pipeline roughness, pipeline diameter, and the Reynolds number.

$$f = \frac{64}{Re} \tag{3.4}$$

$$\frac{1}{\sqrt{f}} = -2Log_{10}\left(\frac{e_d}{3.7D} + \frac{2.51}{Re\sqrt{f}}\right) \tag{3.5}$$

where e_d represents the absolute pipe roughness (mm) and Re is the Reynolds number (dimensionless).

The calculation of the friction factor (f) is directly used in the model to establish the upstream and downstream pressures together with Eq. (3.1). Generally, high friction factors are undesirable as they can increase pressure drops and a decrease in flow rates, which could necessitate adjustments to the pipeline design.

3.1.5 Thermal Effects

Two primary sources contribute to temperature changes during fluid transportation via pipelines. The first is the heat transfer between the pipeline's interior and external environment. This source can be significantly reduced with proper pipeline insulation [54]. The second source is the Joule-Thompson effect, a temperature shift when the gas expands due to pressure disparities between the inlet and outlet [55]. This phenomenon is shown in equation (4.5).

$$J_t = \left(\frac{\partial T}{\partial P}\right)_H \tag{3.6}$$

where J_t refers to the Joule-Thompson coefficient (°C/Pa), T signifies temperature (°C), and P corresponds to pressure (Pa). The term H implies that the partial derivative is evaluated at constant Enthalpy.

The Joule-Thompson equation shows that its coefficient is influenced by the gas's properties and the gas's pre-expansion temperature and pressure. All gases have an inversion point where the Joule-Thompson coefficient changes its sign, switching from warming to cooling.

Most gases, including methane, possess a high inversion point temperature; thus, pressure drops incur cooling effects. Hydrogen, however, is among the rare gases having a very low inversion point, with an inversion point of 202 K at atmospheric pressure [55]. Because of this, hydrogen usually undergoes a warming effect in usual pipeline transportation conditions when subjected to a pressure drop. For the model, the Joule-Thompson effect is assumed to have constant values for methane (-0.5°C/bar) and hydrogen (0.035°C/bar), derived from prior literature [56].

3.1.6 Maximum Allowable Operating Pressure (MAOP)

The operational pressure is directly related to material stress. This is especially important in the scenario of repurposing existing methane pipelines for hydrogen transport, as the existing pipeline's material characteristics must be accounted for. Standard design guidelines for pipelines generally suggest that the primary stress is the hoop stress (in the circumferential direction), which can be mathematically represented by the subsequent equation [53].

$$S_h = \frac{P \cdot D_o}{2t_p} \tag{3.7}$$

where S_h denotes the hoop stress of the pipe (Pa), t_p refers to the wall thickness of the pipe (mm), D_o stands for the outer diameter (mm), and P represents the internal operating pressure (Pa).

This equation shows how the hoop stress is influenced by the operating pressure, diameter of the pipeline, and design thickness. The equation highlights that the operational pressure directly impacts the stress level within the pipe. Industrial standards [32,57] provide the practical relationship between these variables, which usually include design and safety factors. For hydrogen pipelines, the maximum allowable operating pressure ($MAOP$) can be expressed by the following equation [32].

$$MAOP = \frac{2S \cdot t_p}{D_o} D_f \cdot E_f \cdot T_f \cdot H_f \tag{3.8}$$

where $MAOP$ is the already mentioned maximum allowable operating pressure (Pa), S signifies the minimum yield strength (Pa), t_p refers to the wall thickness (mm), D_o stands for the external diameter (mm), D_f is the design factor (dimensionless), T_f represents the temperature derating factor (dimensionless), E_f indicates the pipe seam joint factor (dimensionless). H_f denotes the material performance factor (dimensionless).

3.1.7 Compressor Selection

The previous sections described the multiple effects that the operational pressure and flow rate would have on the different parameters. However, they did not consider how those flow

conditions would be established. This section approaches this topic by studying the selection of compressors necessary to achieve the flow conditions and assessing whether or not it is feasible.

The first relevant parameter is the compression ratio, which indicates how much the compressor can increase the suction pressure and can be computed with the subsequent equation.

$$C_r = \frac{P_d}{P_s} \tag{3.9}$$

where C_r denotes the compression ratio (dimensionless), while P_d and P_s are the discharge and suction pressures (Pa).

The model developed in this study assumes a compression ratio (C_r) of 1.5, which is a typical value for centrifugal compressors [53]. This constraint is used to estimate the number of required compressors. For instance, if a compression ratio of 2.8 is required for a particular target of operational pressure and flow rate, two compressors would be necessary since this exceeds the ratio of 1.5.

The other two parameters to consider for the compressor selection are compressor head (ΔH_c) and compressor power ($Power$), which are related to the flow rate and contribute to determining the number of compressor stages required by the system. The standard formulas for compressor head can be found in equations (3.10) and (3.11) [53].

$$Power = 4.0639 \left(\frac{\gamma}{\gamma - 1}\right) Q \cdot T_1 \left(\frac{Z_1 + Z_2}{2}\right) \left(\frac{1}{\eta_a}\right) \left[\left(\frac{P_2}{P_1}\right)^{\frac{\gamma-1}{\gamma}} - 1\right] \tag{3.10}$$

$$Power = \frac{M_f \times \Delta H_c}{\eta_c} \tag{3.11}$$

where *Power* denotes the compression power (kW), γ is the ratio of specific heats of the gas (dimensionless), Q indicates the volumetric flow rate (m^3/day), T_1 stands for the suction temperature of the gas (K), P_1 and P_2 represent the suction and discharge pressures of the gas (kPa), Z_1 and Z_2 refer to the compressibility of gas at suction and discharge (dimensionless), η_a is the compressor adiabatic efficiency (dimensionless), M_f refers to the mass flow rate of gas (kg/s), ΔH_c denotes the compressor head (m), and η_c stands for the compressor efficiency (dimensionless). The compressor efficiency (η_a) is referred to as the compressor adiabatic efficiency.

3.2 Case Study 1 – Transmission and Distribution Pipelines

The case study is introduced to understand the effects of the different blends and the energy transmission in sample pipeline systems. This study examines the properties of two pipelines, a 322 km transmission line, and a 10 km distribution line. Those two were chosen to represent the extremes in pipeline transportation.

Both pipelines are illustrated in Figure 3.1. Transmission lines are typically designed to operate at higher pressures, transporting substantial volumes of fluids over long distances. Conversely, Distribution lines are typically shorter and operate at lower pressures. The specifications of these pipelines were intentionally varied to assess the potential impacts of a switch from natural gas to blended hydrogen under different pipeline conditions.

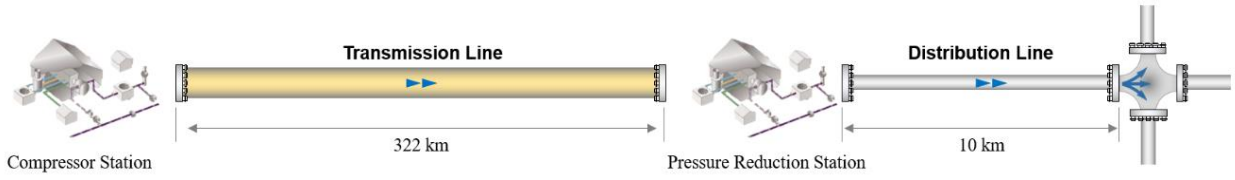


Figure 3.1: Transmission and Distribution Pipelines in Case Study 1.

The selected pipelines' design specifications were primarily derived from Case Study 6 in the "Transmission Pipelines Calculations and Simulations Manual" [58]. Alterations were made concerning the pipe material and weld type that were originally used to make it compatible with those commonly used in current pipelines. The full specification, including details such as fluid composition, location, material, tensile strength, diameter, weld type, wall thickness, surface roughness, delivered pressure, operating flow, and length, can be found in Table 3.1 below.

Table 3.1: Design Specifications for the Case Study 1 Pipelines.

Specification	Transmission Line	Distribution Line
Fluid	H_2: 0 ~ 100%, CH_4: 100% ~ 0%	
Inlet Temperature	20 °C	
Location	Underground (Fully Insulated)	
Altitude Range [m]	0	
Pipeline Material	API 5L X70	ASTM A53 Gr. B
Tensile Strength [MPa]	565.37	330.95
Inner Diameter [mm]	609.6 (NPS 24)	102.26 (NPS 4)
Weld Type	ERW	Seamless
Wall Thickness [mm]	12.7	6.02
Surface Roughness [mm]	0.01778	
Delivered Pressure [MPa]	3.45	0.41
Operating Flow [m3/day]	14,158,423 (500 MMSCFD)	15,000
Length [km]	322	10

Finally, it should be noted that this case introduces no altitude range (i.e., the pipeline is built on a constant elevation). This is done because altitude changes are particular to each pipeline

network, and adding an arbitrary elevation to this example could impact the interpretation of the results.

3.3 Feasibility of Blended Hydrogen Transportation

The model is applied to both pipelines described in case study 1. The simulation cases are blends ranging from 0% hydrogen and 100% natural gas to all the alternatives in between, as well as for different relative energy transmissions. Then, the effects of those different blends and energy transmissions are analyzed regarding flow rate, maximum pressure, friction factor, temperature, and compressor selection.

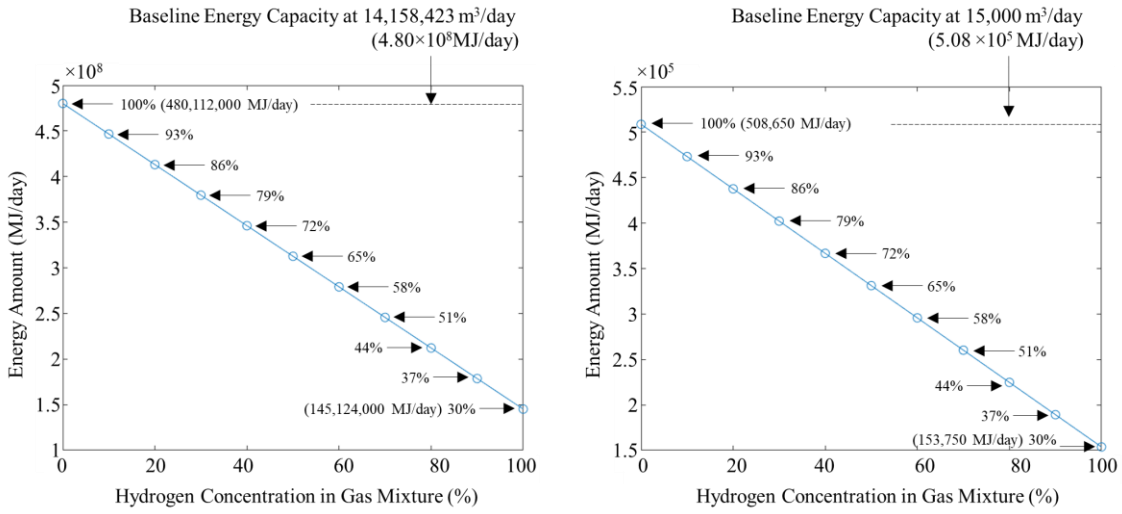
Finally, considering all those factors, an optimal blend is suggested for this case study. This blend is selected to ensure efficient transportation while minimizing infrastructure changes and maintaining a reasonable energy delivery. Thus, this serves as a practical tool that could be used to analyze any other pipeline network, aiding in analyzing the specific impact of that pipeline.

3.3.1 Energy Transmission

As mentioned, pure hydrogen innately has less energy per volume than pure methane. Therefore, blends with higher concentrations of hydrogen will also have less energy than blends with lower concentrations.

Figure 3.2 illustrates this effect for case study 1 pipelines, (a) represents the Transmission pipeline, whereas (b) shows the Distribution pipeline. For this figure, it is assumed that no changes are made to pressure or flow rate.

For example, when the transmission pipeline transports pure natural gas, it has a baseline energy transmission of 80×10^8 MJ/day, but with pure hydrogen, this would drop to 1.44×10^8 MJ/day. The figure shows the full range of variation for different hydrogen concentrations.



(a) Transmission Pipeline (b) Distribution Pipeline

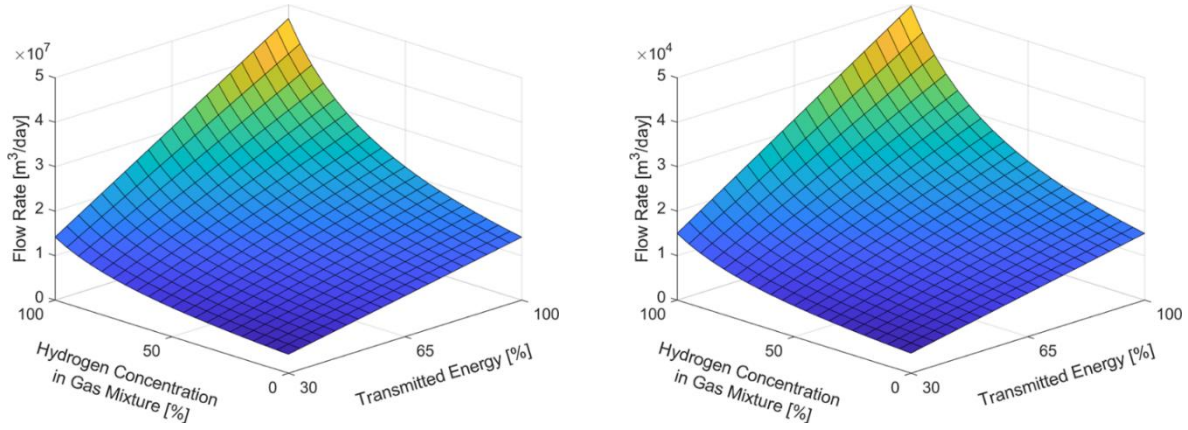
Figure 3.2: Energy Transmission by Hydrogen Concentration in the Gas Mixture.

Thus, if it is desired to maintain the energy transmission the same in a pipeline conversion, the flow rate or operating pressure must be increased accordingly. These changes may not be trivial, as increasing the operating pressure may be constrained by the pipe material used, or it may be limited by the number of compressors that can be installed.

3.3.2 Flow Rate

The variation in flow rate is shown in Figure 3.3 for both transmission and distribution pipelines as a function of both hydrogen concentration and amount of transmitted energy. The energy transmission is on a scale from 30% to 100%, with values corresponding to those shown in the previous section in Figure 3.2.

In other words, 30% corresponds to 1.44×10^8 MJ/day (Pure hydrogen case) and 100% to 4.80x108 MJ/day (Pure methane case). As expected by intuition, a significant increase in flow rate would be needed to maintain high energy transmission with high hydrogen concentrations.

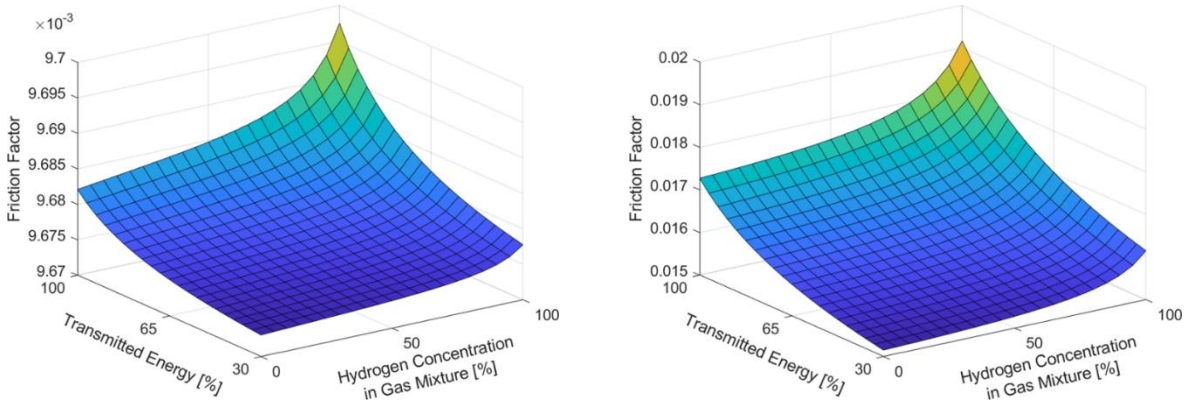


(a) Transmission Pipeline (b) Distribution Pipeline

Figure 3.3: Flow Rate with Hydrogen Concentration and Energy Transmission.

3.3.3 Friction Factor and Temperature

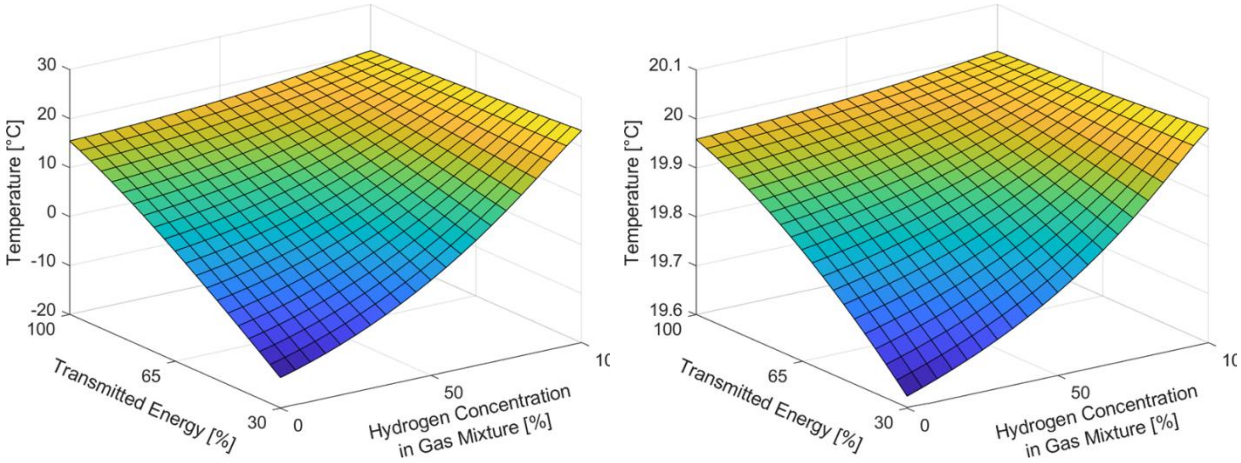
The effect on the friction factor is shown in Figure 3.4 for both pipelines as a function of hydrogen concentration and energy transmission. Higher energy transmissions with high hydrogen concentrations increase the friction factor, largely due to a higher Reynolds number and a more turbulent flow regime. However, it is interesting to note that pure hydrogen transportation has a lower friction factor than pure methane. This makes sense since the viscosity of hydrogen is lower than methane, as shown in Table 2.2.



(a) Transmission Pipeline (b) Distribution Pipeline

Figure 3.4: Friction Factor Variation with Hydrogen Concentration and Energy Transmission.

As mentioned in the decisions and assumptions for the model, it is assumed that the pipeline is mostly buried underground and with proper insulation, and thus, all temperature variation will be due to the Joule-Thompson Effect. While this is sometimes a concern for hydrogen due to its low inversion point, Figure 3.5 quantifies how much of a temperature change this impact is.



(a) Transmission Pipeline (b) Distribution Pipeline

Figure 3.5: Temperature Variation due to Joule-Thompson Effect with Transmitted Energy and Hydrogen Concentration.

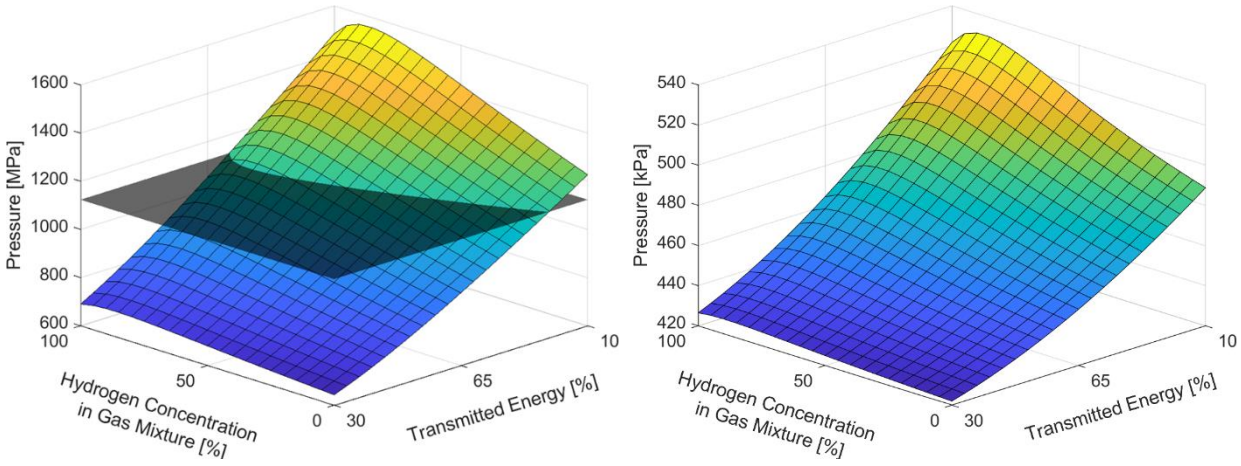
The range of temperature variation in the transmission pipeline is much larger than that of the distribution pipeline. This is expected due to its higher volumetric flow rate. Nevertheless, even in the transmission pipeline with a greater temperature change, it is still not significant enough to cause infrastructural concern.

3.3.4 Maximum Allowable Operating Pressure (MAOP)

Equation (3.8) shows that for the same pipe design, the MAOP for hydrogen pipelines will

be lower than the ones used for natural gas service. Once the MAOP is determined, it can be used with the general flow equation - equation (3.1), to obtain feasible inlet pressures for different hydrogen concentrations and transmitted energy. Figure 3.6 shows the result of that.

Note that a black–shaded square area is included to represent areas where transportation is not feasible because the inlet pressure would be higher than the MAOP. In other words, any scenario above that region would not be reachable with the current pipeline of the case study and end up posing a structural risk. Moreover, the square is only shown on (a) the transmission pipeline because all scenarios are below the MAOP for (b) the distribution pipeline. This is interesting to note because it shows that for pipelines that operate well below the MAOP, increasing the pressure and flow rate significantly to maintain energy delivery might be possible.



(a) Transmission Pipeline (b) Distribution Pipeline

Figure 3.6: Inlet Pressure Variation with Hydrogen Concentration and Energy Transmission.

3.3.5 Compressor Selection

Besides looking at the pipeline itself, it is also important to look at the effects on compressor selection. As mentioned in the methodology, the model will assume a compression

ratio of 1.5 for the pipelines of this case study. With this set as the limit, Figure 3.7 illustrates the number of compressors necessary to achieve different hydrogen concentrations and energy transmissions. Similar to the concept introduced in the MAOP section, a black shaded square is added here to illustrate regions that – if exceeded – would require an additional compressor.

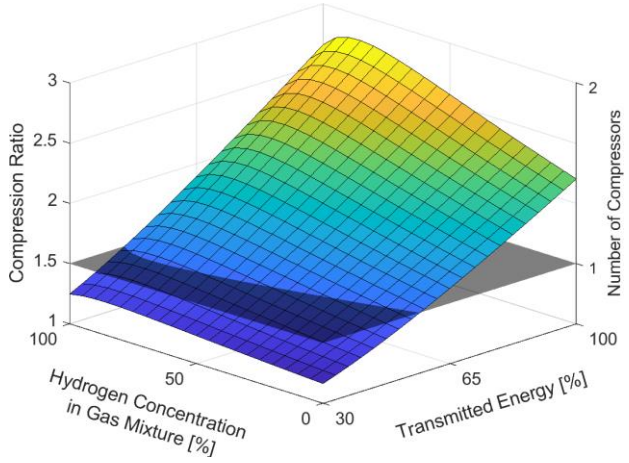
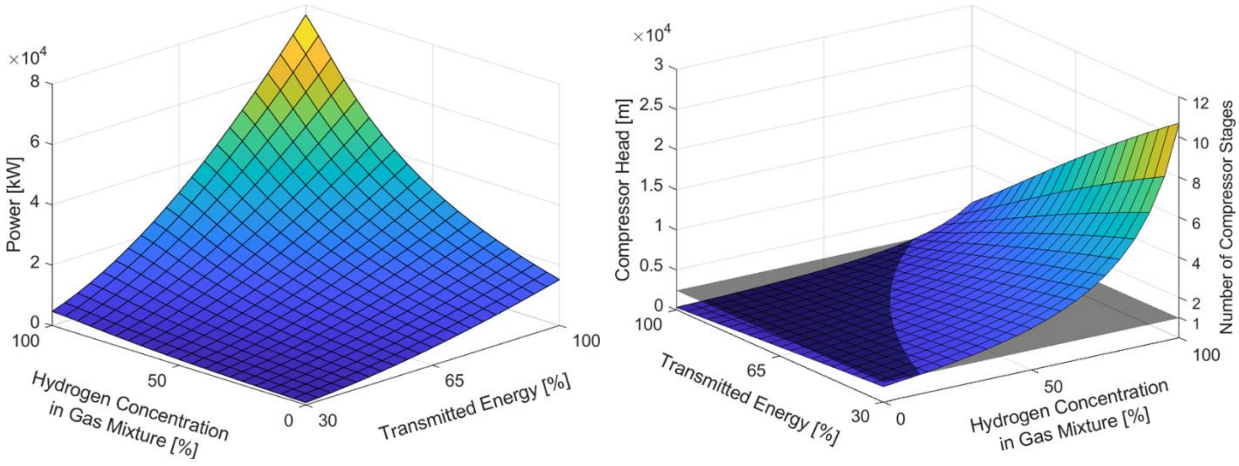


Figure 3.7: Compression Ratio Variation with Hydrogen Concentration and Transmitted Energy.

Besides establishing the number of compressors, defining the required compressor power and head is important. This is illustrated in Figure 3.8, (a) for the compressor power and (b) for the compressor head and – consequently – the number of compressor stages. As expected, higher hydrogen concentration with high energy transmission will require a significant increase in compressor power. It is also interesting to note how many more compressor stages would be required for high-energy deliveries. For example, to achieve 100% transmitted energy with pure hydrogen, eleven compressor stages would be required.



(a) Compressor Power (b) Compressor Head and Compressor Stages

Figure 3.8: Compressor Power, Head, and Number of Compressor Stages with Hydrogen Concentration and Transmitted Energy.

3.3.6 Optimal Hydrogen Blend for Pipeline Transportation

An optimal blend for any pipeline will always depend on the specific goals and the available budget. The model results provide a clear understanding of the consequences and trade-offs involved. While considering every possible scenario would be unrealistic, a selection of interesting examples has been computed and is available in Appendix B. This table presents various scenarios and indicates the extent of modifications required for each pipeline, whether it involves no modification, changes to pipe material, adjustments to compressors, or both.

While the table is comprehensive, some notable examples should be discussed. For example, transitioning from natural gas to pure hydrogen without affecting the energy transmission would require extensive modifications to everything, from the compressor selection to the pipeline material. This would most likely be economically and logistically impractical.

Conversely, a trade-off can be achieved with 65% of the original energy transmission and a blend with 60% hydrogen and 40% methane. This selection would only require the addition of

a single compressor. Even further, transmitting 51% of the energy with a 30% hydrogen blend would be achievable without any infrastructure alteration.

3.4 Summary

The hydraulic feasibility of blended hydrogen transportation is analyzed by developing a steady-state gas hydraulic model and a case study. It was shown that transitioning to blended hydrogen will always result in a loss of energy transmission if compared to pure natural gas for similar pressures and flow rates. However, in some cases, the pipeline might not be operating near its maximum operational pressure (MAOP), and for those cases, the loss of energy transmission might be mitigated.

Finally, the study shows how the optimal hydrogen blend for a given pipeline is contingent on specific objectives and budget. A detailed evaluation of all potential options is crucial for determining each unique scenario's most compatible hydrogen-natural gas mixture. Projects that permit fewer infrastructure modifications may have less flexibility regarding blends with elevated hydrogen concentrations, yet they still present an opportunity for incremental hydrogen integration. The investigation outlined in this section can be a tool for making more informed decisions, allowing for a balanced consideration of trade-offs within project constraints.

CHAPTER 4: Blended Gas Real-Time Transient Model

A real time transient model (RTTM) model is developed. This model serves as the basis for data generation for the AI-based leak detection system introduced in the following chapter, and more than that, it specifically addresses the lack of existing data for blended hydrogen pipelines. There are relatively few hydrogen pipelines and fewer documented leaks for such pipelines.

The RTTM is developed based on the transient flow equations that govern flow in pipelines, grounded in classical principles such as mass conservation and Newton's second law. A methodology specifically to model leaks is also implemented. Since mixed gases present an additional challenge, a specific combination of an equation of state and thermodynamic property tables is used to accurately simulate the properties of arbitrary mixtures of hydrogen and methane.

Recognizing that no model can completely replicate the real world, specific additions are also made to mimic real-world applications. Some additions include randomized transients in the pipeline, proportional integral derivative (PID) response in the boundaries simulating compressors with variable-frequency drive (VFD) and adding artificial white noise to simulate possible instrumentation error and malfunction.

Then, the model is validated against the commercial software Pipe Flow Expert. Finally, a case study inspired by a real-world pipeline is presented to showcase the model's application to such a pipeline and the challenges involved with that. This pipeline is later used as the training and testing data for the leak detection system (LDS) developed in the following chapter.

4.1 Transient Flow Equations

The fundamental equations that describe transient flow are the continuity and the momentum equations. Those are well established in literature [44,59]. The continuity equation arises from the conservation of mass. For an infinitesimal control volume, the net mass accumulation rate inside the volume must equal the differences between the mass entering and leaving the volume.

It should be noted that the continuity equation is often derived assuming a constant density for incompressible fluids (such as liquids). However, the RTTM of this study is developed for gases, and thus, it considers the compressibility of the fluid and variable density. The continuity is expressed as [59]:

$$\frac{\partial P}{\partial t} + V\frac{\partial P}{\partial x} + \rho a^2 \frac{\partial V}{\partial x} = 0 \tag{4.1}$$

where P is the pressure (Pa), V is the fluid velocity (m/s), ρ is the fluid density (kg/m^3), and a is the wave speed (m/s).

The momentum equation is based on Newton's second law of motion. It stipulates that the rate of change of momentum of each fluid element must equal the resultant force acting on it. For fluids in pipelines, this resultant force arises from pressure gradients, gravitational effects (i.e., elevation differences), and viscous friction. The equation is given as [59]:

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + \frac{1}{\rho}\frac{\partial P}{\partial x} + g\,sin\theta + \frac{fV|V|}{2D} = 0 \tag{4.2}$$

where g is the acceleration due to gravity (m/s^2), θ represents the angle with respect to the horizontal for inclined pipes, f is the Darcy friction factor, and the D is the inner diameter of the pipeline.

In the momentum equation, the first term ($\partial V/\partial t$) represents the temporal change in velocity. The second term, ($V\ \partial V/\partial x$), represents the convective acceleration due to changes in velocity along the pipeline. The third term ($\frac{\frac{1}{\rho}\partial p}{\partial x}$) represents the acceleration due to the pressure gradient. The fourth term ($g\sin\theta$) is the acceleration due to gravity. Finally, the last term represents the acceleration due to frictional losses. Notably, the Darcy friction factor calculation, as introduced in section 3.1.4, remains consistent here.

Moreover, through the rest of the model development and derivation, the flow rate is related to the fluid velocity, and the hydraulic head is related to pressure through the fluid density, gravitational acceleration, and elevation. Both of those relations are shown in the following equations.

$$Q = V \cdot A \tag{4.3}$$

$$H = \frac{P}{\rho \cdot g} + z \tag{4.4}$$

where Q is the flow rate (m^3/s), V is the fluid velocity (m/s), A is the cross-sectional area (m^2), H is the hydraulic head (m), ρ is the fluid density (m/kg^3), g is the acceleration due to gravity (m/s^2), and z is the elevation in meters.

To model leaks, they are portrayed as orifices. The orifice size and pressure gradient (inside and outside the pipeline) determine the leak discharge rate. Naturally, such leaks introduce a discontinuity due to mass loss. The leak flow rate can be calculated as follows [60]:

$$Q_L = C_d A_L \sqrt{2gH} \tag{4.5}$$

where Q_L is the steady leak discharge (m^3/s), C_d is the discharge coefficient (unitless).

It should be noted that the discharge coefficient accounts for energy loss due to friction and compressibility factors. C_d is typically obtained empirically, but in this study, it is adopted as a constant of 0.67, a typical value for developed turbulent flow [61].

4.2 Numerical Solution to Transient Flow

The transient flow equations introduced in the previous chapter are partial differential equations (PDEs). They describe the time and spatial variation of pressure (or head) and velocity (or flow rate). Three predominant numerical techniques exist to solve those partial differential equations: explicit integration, implicit integration, and method of characteristics (MoC) [44]. Each one of those has advantages and disadvantages.

The explicit integration method computes the state of a system at a future time based on the current state. It directly approximates the solution using the existing values without solving additional equations. While explicit methods are simple and can be faster for specific applications, they are conditionally stable. The stability is typically governed by the Courant-Friedrichs-Lewy (CFL) condition, which demands specific time-step sizes related to spatial discretization.

Implicit integration methods, on the other hand, determine the future state of a system by solving an equation that includes both the present and future states. These methods are generally unconditionally stable, implying they are not constrained by the same time-step restrictions as explicit methods. However, the trade-off is that they often require more computational effort per time step because they entail the solution of algebraic systems of equations.

Finally, The MoC stands out as it transforms the PDEs into ordinary differential equations (ODEs) along certain curves called characteristics. This method is especially proficient in conserving the wave behaviour of transient flow equations, which is invaluable for surge analysis

tasks, such as leak detection [62]. However, like the explicit method, this method is also constrained by the Courant condition, requiring a carefully designed and stable grid to ensure accurate results. For its adeptness at being applied for leak detection, the MoC will be employed in this research. Specific techniques will be introduced later on to deal with its drawbacks.

The application and derivation of the MoC to the transient flow conditions are well documented in the literature [44,59,62]. However, this study will also include its derivation for added clarity. The characteristic equations can be defined by linearly combining the momentum and continuity equations and introducing an unknown multiplier (λ). By multiplying Eq. 4.2 by this unknown multiplier, adding Eq. 4.1, and representing it in terms of flow rate (Q) and head (H) instead of pressure (P) and velocity (V), the following relation can be obtained [59]:

$$\left(\frac{\partial Q}{\partial t} + \lambda \cdot a^2 \frac{\partial Q}{\partial x}\right) + \lambda \cdot g \cdot A \left(\frac{\partial H}{\partial t} + \frac{1}{\lambda}\frac{\partial H}{\partial x}\right) + \frac{f \cdot Q \cdot |Q|}{2 \cdot D \cdot A} = 0 \tag{4.6}$$

Given that both the hydraulic head and flow rate vary with time and distance ($Q(x,t)$, $H(x,t)$), a solution for the unknown multiplier exists such that:

$$\lambda \cdot a^2 = \frac{dx}{dt} \tag{4.7}$$

By using the solution of the multiplier (λ) and rewriting equation 4.6, we can obtain the compatibility equations [59]:

$$\frac{\partial Q}{\partial t} + \frac{gA}{a}\frac{\partial H}{\partial t} + \frac{f}{2DA}Q|Q| = 0 \tag{4.8}$$

And similarly,

$$\frac{\partial Q}{\partial t} - \frac{gA}{a}\frac{\partial H}{\partial t} + \frac{f}{2DA}Q|Q| = 0 \tag{4.9}$$

Because of the multiplier introduced, it should be noted that equation 4.8 is only valid as long as ($dx/dt = a$), and similarly, equation 4.9 is only valid if $\frac{dx}{dt} = -a$. In other words, those PDEs behave like ODEs along the characteristic lines in the x-t plane.

The x-t plane is illustrated in Figure 4.1. Both lines represent the characteristic lines (each one in a direction). From this point forward, every discrete step in the x-t plane in the x-direction (Δx) is referred to as a spatial node, whereas the t-direction is a temporal node (Δt). The figure represents how, by having two known points (A and B) in the current temporal node (t_1), it is possible to calculate an unknown point (P) at time t_2. These known points are referred to as boundary conditions and will be further detailed later.

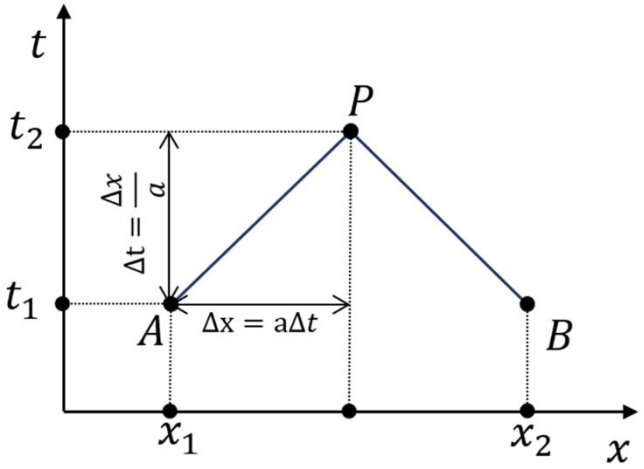


Figure 4.1: Characteristic Lines in the x-t Plane.

By integrating the characteristic lines over known points, it is possible to obtain the positive characteristic equation (C_p) and the negative characteristic equation (C_n), as shown in the following equations on known points A and B:

$$C_p = Q_A + \frac{g \cdot A}{a} H_A - \frac{f}{2DA} \cdot \Delta t \cdot Q_A |Q_A| \tag{4.10}$$

$$C_n = Q_B - \frac{g \cdot A}{a} H_B - \frac{f}{2DA} \cdot \Delta t \cdot Q_B |Q_B| \tag{4.11}$$

Physically, the positive characteristic equation represents the wavefront traveling in one direction, whereas the negative characteristic equation represents the wavefront in the opposite direction. Finally, the flow rate and hydraulic head at an unknown point P can be obtained with:

$$\mathrm{Q_p} = C_P - \frac{gA}{a} \cdot H_p \tag{4.12}$$

$$\mathrm{Q_p} = C_n + \frac{gA}{a} \cdot H_p \tag{4.13}$$

As mentioned before, the most significant drawback of the MoC stems from its stringent adherence to the CFL condition. This condition dictates that there must be a direct correlation between grid spacing in both spatial and temporal steps and the wave speed, which was evident during the derivation of the method. Mathematically, this is stated as [62]:

$$a = \frac{\Delta x}{\Delta t} \tag{4.14}$$

This limit is considered a drawback because this means very fine uniform distance data must be available for a pipeline to obtain accordingly small time steps, which might not always be possible or economical, to be done. This limitation will be addressed directly later with interpolation techniques to ensure that accurate data can be obtained.

4.3 Initial and Boundary Conditions

The previous section emphasized that boundary conditions are essential to solving the transient flow conditions in the x-t grid. This section will discuss both the boundary conditions and the initial conditions.

The initial conditions are responsible for starting the RTTM with known values and representing the state of the pipeline in the first temporal step. The initial state condition can be defined by assuming the pipeline starts in steady-state flow [59]. The flow conditions for steady-state are calculated using the same equations already discussed in great detail in Chapter 3.1 during the development of the steady-state gas hydraulic model.

Conversely, Boundary conditions describe the relationship between the flow rate or hydraulic head at the inlet and outlet boundaries of the system. A frequently employed boundary condition is the fixed boundary, also known as the Dirichlet boundary type [63]. In this boundary condition, one variable is pre-defined and fixed (e.g., flow rate), and the second variable is calculated based on the former (e.g., hydraulic head). For instance, in the case of a fixed inlet flow rate and fixed outlet hydraulic head boundary conditions, the state at both boundaries could be fully determined with:

$$\mathrm{H}_{in} = C_{n,in} + \frac{a}{g \cdot A} \cdot Q_{in,fix} \tag{4.15}$$

$$\mathrm{Q}_{out} = C_{p,out} + \frac{a}{g \cdot A} \cdot H_{out,fix} \tag{4.16}$$

where $Q_{in,fix}$and $H_{out,fix}$ are the fixed inlet flow rate and outlet hydraulic head, respectively. $C_{n,in}$ is the negative characteristic equation at the inlet, $C_{p,out}$ is the positive characteristic line in the outlet, and finally, Q_{out} and H_{in} are the derived flow rate and hydraulic heads.

While the fixed boundary type is a good characterization of certain scenarios, it does not accurately represent a real-world pipeline system. This is because pipelines often use pumps (for liquids) and compressors (for gases) coupled with variable frequency drives (VFDs) and a control mechanism to regulate the flow and respond to unexpected conditions during operations, such as a leak.

Because of this, the RTTM model developed in this study uses instead controlled boundary types. These are analogous to the Dirichlet boundary type, but it defines a setpoint instead of fixing an inlet or outlet variable. Then, a control mechanism adjusts the variable based on the setpoint. These boundaries are defined as such:

$$H_{in} = H_{prev} + \frac{a}{g \cdot A} \cdot Control\_Resp(Q_{in,set}, Q_{prev}) \tag{4.17}$$

$$Q_{out} = Q_{prev} + \frac{a}{g \cdot A} \cdot Control\_Resp\left(H_{out,set}, H_{prev}\right) \tag{4.18}$$

where H_{prev} and Q_{prev} represents the value of the hydraulic head or flow rate in that location for the previous temporal node, $H_{in,set}$ and Q_{set} represents the setpoint in the control system for those two variables.

Within this context, $Control\_Resp$ represents the control mechanism's response to adjustments in the variable. In this study, the chosen control mechanism is a standard proportional-integral-derivative (PID) controller, a common choice for such applications [59]. This control mechanism was developed based on the principles outlined in [64] and is modeled by:

$$Control\_Resp = K_p \cdot error + K_i \cdot \int error\, dt + K_d \cdot \frac{\partial(\text{error})}{\partial \text{t}} \tag{4.19}$$

where K_p, K_i, and K_d are, respectively, the proportional, integral, and derivative gains, and error is the difference between the setpoint and the current value. The errors are calculated between the

current variable value and the previous value.

The specific gain values and setpoints are defined on a per-pipeline system basis and will be defined later for the specific pipeline simulated. Finally, it should be noted that in real-life scenarios, the VFD's response is delayed allowing the system to respond before adjusting the flow. A sampling interval for the control mechanism also simulates this.

4.4 Simulating Different Fluids

The calculations previously introduced depend on the fluid state at each temporal step. For instance, the Reynolds number is influenced by the fluid's current density and viscosity. Similarly, an accurate determination of the wave speed is essential for defining the relationship between the temporal and spatial nodes that constitute the x-t grid.

RTTMs tailored for liquid transportation often utilize fixed values for density and viscosity. This approximation is justified when dealing with incompressible fluids, as those values typically do not vary significantly. Conversely, gases are much more sensitive to those variations. For example, natural gas density is 300 times more sensitive to pressure variations than crude oil [44]. Because of that, the model must consider the instantaneous values of these properties. Such properties are typically derived from thermodynamic tables for pure, standard gases.

This task is more challenging when dealing with blended gases, such as the blend of natural gas and hydrogen. The RTTM supports continuous blends in this research, encompassing any ratio from 0% to 100% of these gases. As such, relying solely on thermodynamic tables is insufficient. Therefore, a combined approach is adopted: The density and wave speed are derived from an EOS for mixtures. At the same time, viscosity is obtained using a thermodynamic library and an EOS

specific to mixture viscosity.

The chosen EOS for this study is GERG-2008 [65]. This internationally recognized EOS provides a comprehensive method for determining the thermodynamic and transport properties of various gases and their mixtures. Accurately simulating continuous ranges of mixtures is important for ensuring a diverse range of flow conditions to train and test the leak detection system later.

The methodology to compute the density and wave speed is meticulously detailed in GER-G2008's publication [65]. For the sake of replicability, the pertinent details are reiterated here. First, the density of a gas mixture can be calculated with:

$$\delta = \frac{\rho(P,T)}{\rho_r(\bar{x})} \tag{4.20}$$

where ρ(P, T) is the density at given pressure (P) and temperature (T), $\rho_{\mathrm{r}}(\bar{x})$ is the composition-dependent reducing functions for the mixture at the current state and $\bar{x}$ is simply a variable that defines the mixture composition. Finally, δ is the reduced mixture density ($\delta = \frac{\rho}{\rho_r}$).

Then, the composition-dependent reducing function ($\rho_{\mathrm{r}}(\bar{x})$) can be calculated with the equation below:

$$\frac{1}{\rho_r(\bar{x})} = \sum_{i=1}^{N} x_i^2 \frac{1}{\rho_{c,i}} + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} 2x_i x_j \beta_{v,ij} \gamma_{v,ij} \cdot \frac{x_i + x_j}{\beta_{v,ij}^2 x_i + x_j} \cdot \frac{1}{8}\left(\frac{1}{\rho_{c,i}^{1/3}} + \frac{1}{\rho_{c,j}^{1/3}}\right)^3 \tag{4.21}$$

where $\beta_{v,ij}$ and $\gamma_{v,ij}$ are binary parameters fitted to the data for binary mixtures, provided in Table A8 in the publication. $\rho_{c,i}$ are the critical parameters for the individual gases in the mixture, provided in Table A5 of the publication (mol.dm^{-3}).

The wave speed can be calculated with the following relation:

$$\frac{a^2(\delta,\tau,\bar{\bar{x}})M}{RT} = 1 + 2\delta\alpha_\delta^r + \delta^2\alpha_{\delta\delta}^r - \frac{(1+\delta\alpha_\delta^r - \delta\tau\alpha_{\delta\tau}^r)^2}{\tau^2(\alpha_{\tau\tau}^o + \alpha_{\tau\tau}^r)} \tag{4.22}$$

where a is the wave speed (m/s), R is the molar gas constant (8.314472 J.mol^{-1}K^{-1}), M is the molar mass of the mixture (g/mol), and α is the Helmholtz free energy (Dimensionless), and α_δ^r is its derivation with respect to the reduced mixture density δ. This calculation is described in Table B5 of the publication.

Finally, viscosity is obtained by a combination of methods. First, pure hydrogen and methane viscosity are obtained for the current state (temperature and pressure) using CoolProp™ [66], an open-source thermophysical property library. Finally, the viscosity of the mixed gas is derived using the Wilke mixing rule, which is especially suited for hydrocarbons and can be applied to mixtures containing up to seven components [67]. The viscosity of the blend can be calculated as follows:

$$\mu_{mix} = \sum_{i=1}^{N} \frac{x_i\mu_i}{x_i + \sum_{\substack{i=1 \\ j\neq 1}}^{v} \phi_{ij}x_j} \tag{4.23}$$

where μ_{mix} is the mixture viscosity (Pa.s), μ_i is the viscosity of component i, x_i and x_j are mixture`s mole fractions. The dimensionless constant ϕ_{ij} is defined by Equation 4.24.

$$\phi_{ij} = \frac{\left(1 + \sqrt{\frac{\mu_i}{\mu_j}}\left(\frac{M_j}{M_i}\right)^{\frac{1}{4}}\right)^2}{\sqrt{8\left(1+\frac{M_i}{M_j}\right)}} \tag{4.24}$$

where μ_i and μ_j are the viscosities of pure components i and j, respectively. Similarly, M_i and M_j are their respective molecular weights.

Using the previously described techniques, figures were generated to demonstrate the blended hydrogen and natural gas densities, wave speed, and viscosities for pressures ranging from 0.1 MPa to 10 MPa in 0.1 MPa steps and a fixed temperature of 20°C. The blends selected ranged from 100% hydrogen and 0% methane to 0% hydrogen and 100% methane in 5% steps. Figure 4.2 shows the density and wave speed obtained using GERG's EOS, and Figure 4.3 shows the viscosity obtained using CoolProp's thermodynamic library and Wilke's equation. In both figures, the blends are represented on a color scale where the darker colors represent higher amounts of hydrogen, and lighter colors have higher amounts of methane.

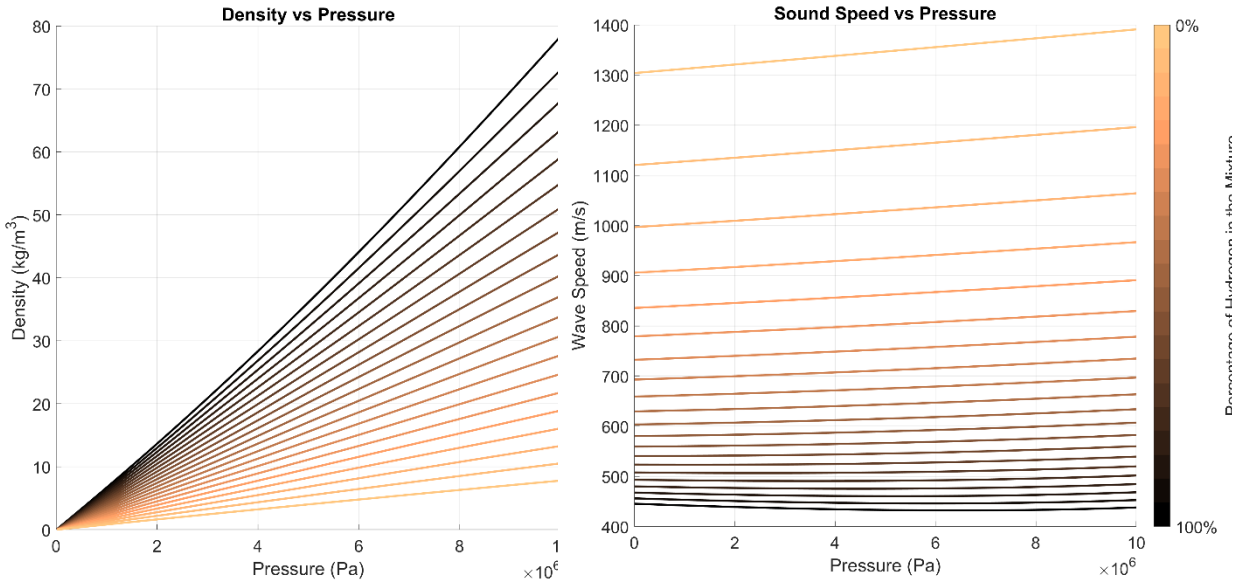


Figure 4.2: Density and Wave Speed Across Hydrogen Blends at Varying Pressures.

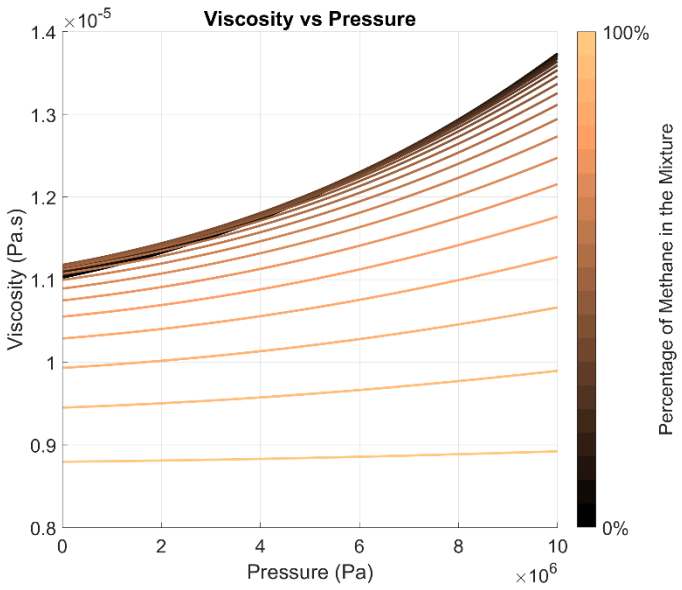


Figure 4.3: Viscosity Across Hydrogen Blends at Varying Pressures.

4.5 Transient and Noise Modeling

In order to simulate real-world transient conditions that pose additional challenges to the leak detection system developed subsequently, there are options to intentionally add a randomized transient event and noise to the system.

The transient behaviour represents a condition that could occur due to unplanned events, such as a valve closing upstream or a malfunction in a pump or compressor. Those events take the form of a perturbation in the inlet pressure and velocity for a randomized value up to 10% of the inlet values, representing a transient event traveling downstream. The transient also has a randomized start time as well as duration. Figure 4.2 illustrates a random example of an inlet sensor reading of pressure with a leak but no transient added on the left and transient added on the right. In this example, the transient event starts significantly before the leak, leading to a larger pressure drop than without the event.

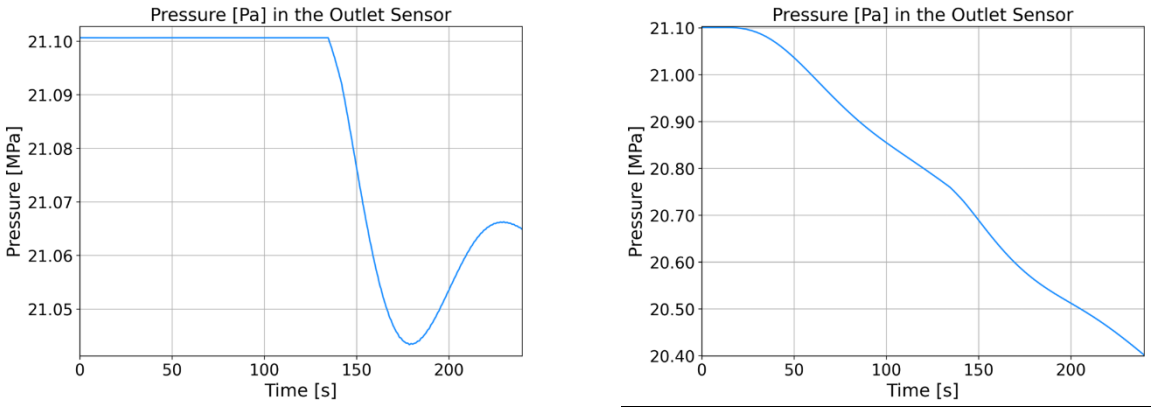


Figure 4.2: Leak Event Comparison - With and Without Transient Events.

Furthermore, the computational model's pressure and flow rate data are clean and devoid of the noise and other uncertainties commonly encountered in real-world measurements. In order to bridge this gap between the idealized model and practical applications, Gaussian white noise is artificially introduced into the data, constituting 1% of the signal's magnitude. This inclusion emulates the potential measurement errors and random fluctuations typically observed in field applications.

The Gaussian white noise is generated based on a zero-mean Gaussian (or Normal) distribution, characterized by its bell-shaped probability density function. This distribution ensures that each data point in the time series is perturbed by a statistically independent and identically distributed random variable, adhering to the principles of a Gaussian random process. Such a noise model is often employed in analyzing real-world sensor data, as it closely approximates the kinds of random disturbances commonly observed [68]. Consequently, introducing this noise enhances the realism of the simulated data and offers a more robust framework for evaluating the performance and reliability of sensor-based systems. The two examples of leak cases that illustrate this are shown in Figure 4.3. On the left, the same reading for the outlet sensor is shown for the

pressure as before, and on the right, the sensor with the noise added.

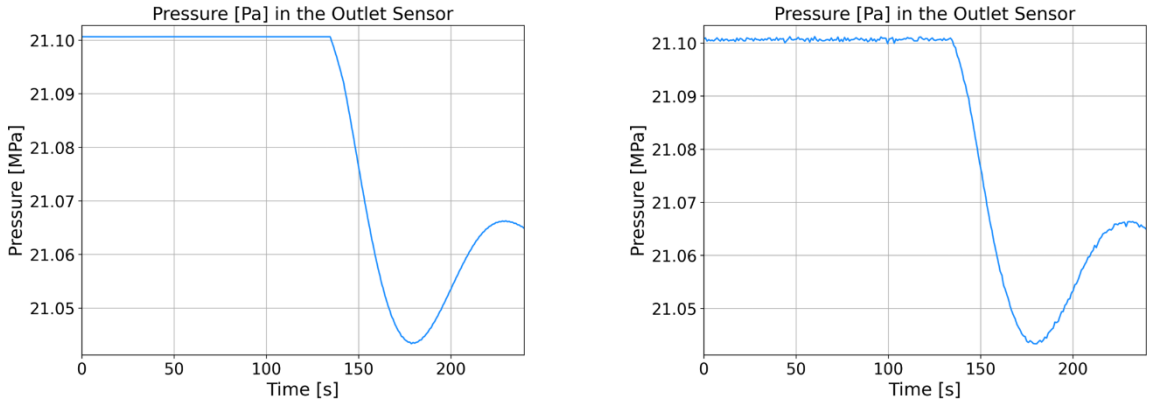


Figure 4.3: Leak Event Comparison - With and Without Artificial Noise.

By implementing transients and noises, the model strives to align more with the complexities and unpredictability of real-world pipeline systems, ensuring that the LDS is tested under more challenging and authentic scenarios.

4.6 Assumptions and Limitations

While the RTTM developed so far has taken steps to capture realistic behaviour as much as possible, it has assumptions and limitations that should be addressed. In finite element modeling, assumptions that simplify the model and simplify a model without hampering its accuracy are considered good practice, as a simpler model often offers better computational speed and has less room for unaccounted errors.

One of the assumptions taken is that the flow is one-dimensional, meaning the fluid properties must be constant throughout the pipe cross-sectional area. This is a common assumption for commercial models [62]. Second, it is also assumed that any mixture of gases is homogeneous.

The homogeneity of a hydrogen mixture was already discussed in the literature review, and it was found that for turbulent flow conditions, hydrogen and natural gas do not stratify in the pipeline [26].

The flow is assumed to be isothermal, a consideration based on the presumption that most pipelines are insulated and buried. This assumption serves as a balance between model complexity and computational efficiency. Incorporating temperature variations would intensify the computational demands of the equations of state. Furthermore, the feasibility study with the steady-state model already explored the joule-Thomson effects.

The model was mainly designed for a blend of hydrogen and natural gas in the study's objective, but it can also be used for liquids. As long as the equations of state (EOSes) used accurately represent the properties of the liquid, the transient flow equations and the boundary conditions remain applicable.

Regarding limitations, the x-t grid spacing is set according to the Courant-Friedrichs-Lewy (CFL) condition. The fixed grid means the simulation runs at a set wave speed defined in the initial time step. That being said, Figure 4.2 shows the variation of wave speed for different pressures, and it is shown that – for a specific blend – the variation is less than an order of magnitude and should not significantly impact the interpretation of the results.

In summary, while the developed RTTM is a comprehensive framework for transient gas flow in pipelines, it is important to consider the assumptions and limitations when discussing the results.

4.7 Validation – Comparison with Commercial Software

The RTTM developed in this research is validated against a commercial software, Pipe

Flow Expert™ [69]. The author acknowledges and thanks the company for providing a license for this study. This software is designed for pipe loop systems and accommodates both liquids and gases using non-compressible and compressible equations.

For validation, a pipeline spanning 20 km (19866 m) is virtually modeled in the RTTM and Pipe Flow Expert with 44 nodes each. Consistent inlet and outlet boundary conditions were established for both. In RTTM, the boundary conditions were as described in the previous chapter. In Pipe Flow Expert™, a flow demand at the inlet node (equivalent to a fixed flow rate) and a Tank node at the outlet (equivalent to a fixed pressure) were configured. Figure 4.4 illustrates the pipeline as modeled in Pipe Flow Expert.

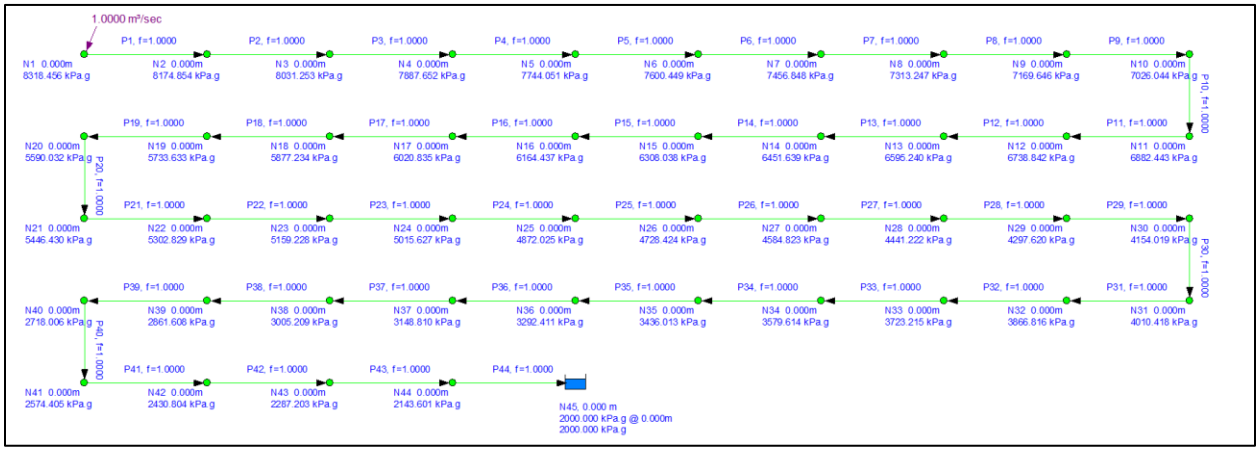


Figure 4.4: Graphical Definition of a 44-Node Pipeline in Pipe Flow Expert

There are two functional differences between the RTTM's approach and Pipe Flow Expert regarding modeling fluids and leaks. Leaks are not supported natively for Pipe Flow Expert, but flow discharges are supported. Therefore, to simulate a leak, the rate is first calculated in the RTTM using the orifice model approach [61], and then it is used as a flow discharge at the leak node in Pipe Flow Expert.

The second difference is that Pipe Flow does not directly define gas mixtures. It instead allows the user to directly define new fluids (or mixtures) with their properties. To work around this, the density and viscosity for each mixed fluid are calculated beforehand using the approach mentioned in the previous section of GERG-2008 [65] and CoolProp [66] and used to define the mixtures in Pipe Flow.

A total of ten tests were performed for the validation. Those have their gas mixture composition varying from 100% hydrogen and 0% methane to 0% hydrogen and 100% methane in 25% steps. Each mixture is compared for leak and no-leak scenarios using the same boundary conditions. Table 4.1 summarizes the input conditions.

Table 4.1: Input Conditions for Comparison Between RTTM and Pipe Flow Expert.

Fluid Type	Temperature	Outlet BC Pressure	Inlet BC Flow Rate	Density [kg/m³]	Viscosity [Pa.s]
100% H_2 0% CH_4	20°C	20 MPa	2 m³/s	14.7065	9.16E-06
75% H_2 25% CH_4				41.426	1.47E-05
50% H_2 50% CH_4				71.721	1.73E-05
25% H_2 75% CH_4				122.793	1.86E-05
0% H_2 100% CH_4				131.636	1.93E-05

For example, the comparison results for the case with pure hydrogen can be visualized in Figure 4.5. The left plot shows the comparison of the pressure gradient. Blue lines represent Pipe

Flow results, whereas orange results represent the RTTM results. The right plot also shows the flow rate throughout the pipeline for both methods.

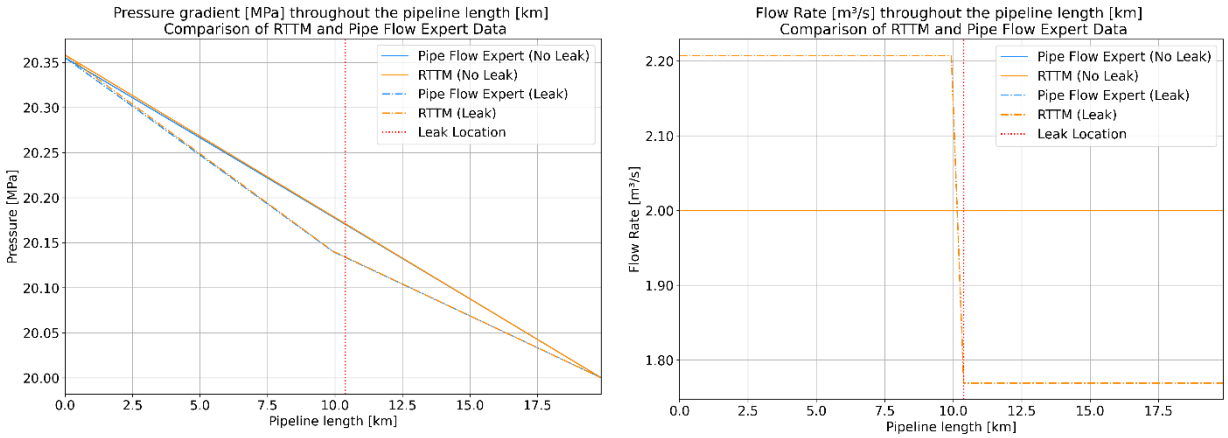


Figure 4.5: Comparison of Results for 100% Hydrogen for RTTM and Pipe Flow.

In the 100% hydrogen example, the inlet pressures calculated for both the leak and no leak cases were 20.358 MPa (RTTM) and 20.355 MPa (Pipe Flow). The maximum relative difference in the entire pressure gradient was 0.02%. For the flow rate, the outlet calculated was consistently 2 m^3/s for both methodologies in the no-leak case and 1.769 m^3/s in the leak cases. The relative difference was, at most, 0.0034%.

This comparison process was repeated for every fluid type, and the results are shown in Table 4.2. The values shown are for the pressure at the inlet and flow rate at the outlet, as showing the entire line would be impractical. Nevertheless, all the figures are shown in Appendix *A*. Finally, the table shows the maximum relative error for pressure and flow rates.

Table 4.2: Results of the Validation Between RTTM and Pipe Flow Expert.

Inputs		Pressure Gradient			Flow Rate Profile		
Fluid Type	**Flow Condition**	**RTTM (Inlet) [MPa]**	**Pipe Flow (Inlet) [MPa]**	**Maximum Rel. Error (Line)**	**RTTM (Outlet) [m³/s]**	**Pipe Flow (Outlet) [m³/s]**	**Maximum Rel. Error (Line)**
100% H2 0% CH4	**No Leak**	20.36	20.35	0.02%	2.00	2.00	0.00%
	Leak	20.36	20.35	0.02%	1.77	1.77	0.00%
75% H2 25% CH4	**No Leak**	21.00	20.98	0.12%	2.00	2.00	0.00%
	Leak	21.00	20.98	0.11%	1.86	1.86	0.05%
50% H2 50% CH4	**No Leak**	21.73	21.66	0.33%	2.00	2.00	0.00%
	Leak	21.73	21.66	0.32%	1.90	1.90	0.05%
25% H2 75% CH4	**No Leak**	22.95	22.76	0.85%	2.00	2.00	0.00%
	Leak	22.91	22.72	0.87%	1.90	1.90	0.11%
0% H2 100% CH4	**No Leak**	23.16	22.95	0.96%	2.00	2.00	0.00%
	Leak	23.16	22.95	0.95%	1.93	1.93	0.19%

Since the fluid properties were defined beforehand, this comparison is a direct look at the implementation of the gas flow model implementations. The results show that the error was consistently below 1% for all cases, indicating a good correlation between RTTM and Pipe Flow Expert. Moreover, as expected by intuition, mixtures with higher viscosities and densities also had larger pressure drops and, thus, required a higher inlet pressure.

4.8 Case Study 2 – Alberta Products Pipeline

The second case study uses geographical data to simulate a real-world application. Specifically, this simulation draws inspiration from an actual pipeline in Alberta: the Alberta

Products Pipeline (APPL), regulated by the Alberta Energy Regulator (AER) and operated by Trans-Northern Pipelines [70]. The decision to emulate a real-world pipeline allows for a detailed discussion of the challenges of adhering to the CFL condition and studying a pipeline rooted in reality.

Stretching over an expanse of roughly 300 kilometers, the APPL connects the cities of Edmonton and Calgary. The precise pathway of this pipeline is proprietary to Trans Northern Pipelines and remains confidential. Therefore, in order to determine a similar path for this study, the right of way (ROW) was derived by manually selecting geographical points via Google Earth [71] and cross-referencing with a publicly available map from Trans Northern Pipelines [70]. The primary objective is not a perfect emulation of APPL's right of way but to mimic flow behaviour in a similar pipeline topology, considering comparable distances and elevation profiles. Figure *4.6* graphically showcases this simulated pipeline's geographical route and elevation, spanning from Edmonton to Calgary.

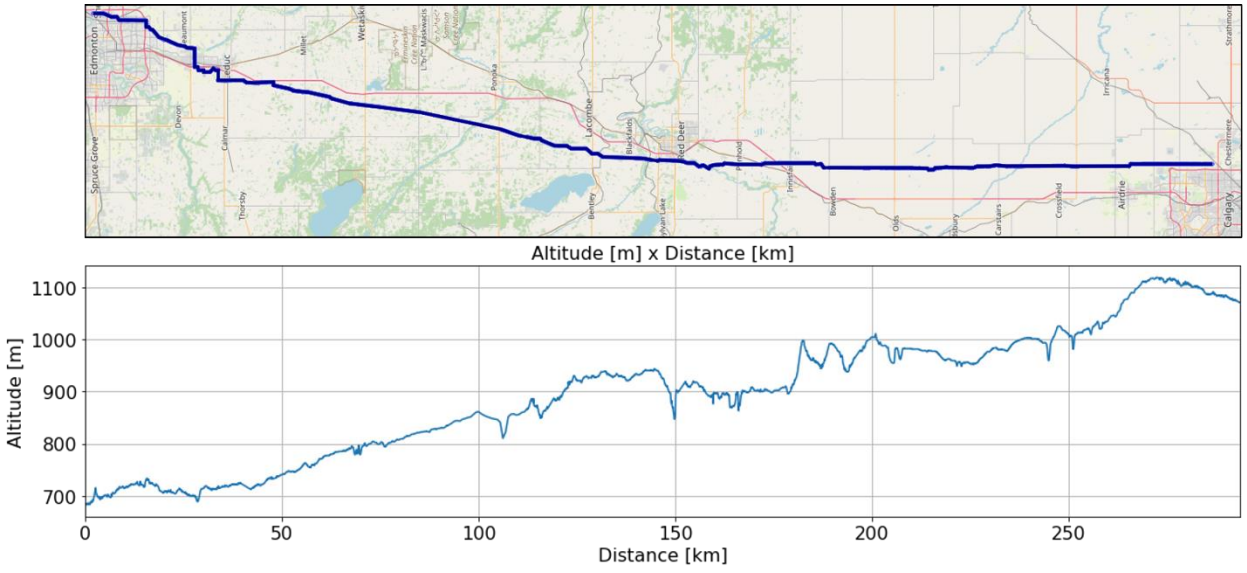


Figure 4.6: Topographical Route of Edmonton-Calgary Pipeline in Case Study 2.

The main difficulty in employing the RTTM on such a path is that it is challenging to ensure a homogeneous distance between hand-picked points in a map. This issue is solved by implementing an algorithm that interfaces with the Google Maps™ API [72] and interpolates those points until a desired resolution is reached, allowing for potentially arbitrarily small and precise resolutions. Accordingly, the elevation was obtained using the Elevation module of Google Maps™ API and combined with the depth of cover to determine the pipeline altitude.

Moreover, besides defining the pipeline ROW, the compressor stations were also identified. It is assumed that the only points of the pipeline with instrumentations are those compressor stations; therefore, the RTTM will be used to simulate flow between those sites. Each distance between two compressor stations or the compressor and the start or end of the pipeline is designed as a segment. Those eight segments are overlayed on the map, with their elevation shown in Figure 4.7.

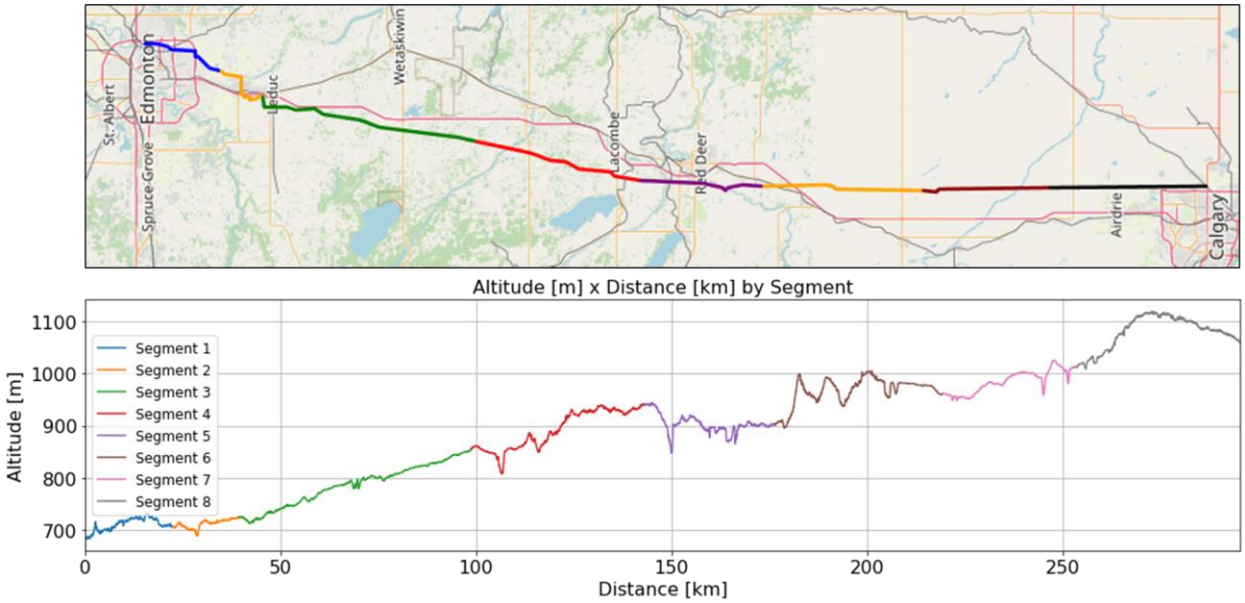


Figure 4.7: Segment-Stratified Topographical Route of Case Study 2 Pipeline.

The exact design specifications of the APPL are also not in the public domain. Thus, this case study postulates a design specific to this research. It was designed as a large-scale transmission pipeline, potentially ferrying a hydrogen blend from Edmonton to Calgary in a hydrogen economy. Key specifications are shown in Table 4.3.

Table 4.3: Design Specifications for the Case Study 2 Pipeline.

Contents	Alberta Pipeline
Fluid	H_2: 50 ~ 100%, CH_4: 100% ~ 50%
Flow Temperature	20 °C
Location	Underground (Fully Insulated)
Altitude Range [m]	682 to 1120
Pipeline Material	Steel (ANSI)
Inner Diameter [mm]	501.65 (NPS 22)
Wall Thickness [mm]	53.98
Surface Roughness [mm]	0.046
Inlet Boundary Condition	Flow Rate (2 m^3/s)
Outlet Boundary Condition	Pressure (20 MPa)
Control Response	PID (K_p = 0.01, K_i = 0.01, K_d = 0.001, 0.2 Hz)
Length [km]	295.4

Furthermore, as mentioned beforehand, each segment is simulated with both the inlet and

outlet of the boundary control types. The inlet is simulated with a flow rate setpoint and the outlet with a pressure setpoint. The specifics of the setpoints are shown in the table as well. It is also assumed that the pipeline can transport blends varying continually between 50% to 100% hydrogen and 100 to 50% methane to account for potential variations in different batches. Finally, the PID control gains were defined by trial and error and shown in the table.

Table 4.4: Segment Lengths and Spatial Steps for Case Study 2 Pipeline.

Segment	Number of Discretization Steps	Total Length [km]	Spatial Step (Δx) [m]	Minimum Elevation [m]	Maximum Elevation [m]
1	45	22.5	500.8	682.2	733.3
2	45	16.9	375.0	688.8	724.7
3	45	59.3	1318.6	712.4	855.0
4	45	44.1	980.7	808.2	941.9
5	45	33.4	742.7	847.3	944.3
6	45	43.1	957.2	895.6	1005.3
7	45	34.0	756.1	948.0	1025.3
8	45	42.0	933.7	1009.0	1120.0

Each modeled segment has its x-t grid and is discretized with a resolution of 45 interpolated spatial points (Δx) using the technique mentioned before. While enforcing a consistent grid size between segments would be possible, the approach of having varying grid sizes was made to pose an additional challenge to the leak detection system and represent better scenarios where limited data might be available for different regions of a pipeline. In other words, with this approach, longer segments will have coarser grids, whereas smaller segments will have finer grids. Table 4.4

shows each segment's length, spatial steps, and elevation range. Finally, it should be noted that this information alone is not sufficient to determine the spatial step size (Δt), as this also depends on the wave speed for the blend in transport.

4.9 Summary

The RTTM is developed using the continuity and momentum partial different equations (PDE). Leaks are modeled as orifices, with their sizes and internal pressure gradient determining the leak discharge rate. Those PDEs are solved using the Method of Characteristics (MoC). This numerical solution method is typically suitable to simulate leaks and is used in commercial RTTM simulators but is only stable if the Courant-Friedrichs-Lewy (CFL) condition is respected.

The model's boundary conditions are control-type on both inlet and outlet, and their response is dictated by a classic PID response model, with the specific gains tweaked per pipeline. Additionally, the pipeline can generate arbitrary transients through perturbations in the inlet conditions, simulating scenarios such as valves closing upstream or compressor failures.

The properties of blended hydrogen and methane are obtained with a two-fold strategy. First, the density and wave speed are obtained from a wide-range equation of state, GERG-2008. Then, the viscosity is obtained through an open-source thermodynamic property library, CoolProp, and Wilke's mixing rule.

The RTTM is validated against the commercial software Pipe Flow Expert with ten different test cases, contemplating leaks, and no leak cases for different mixtures. The relative errors are smaller than 1% for all cases, indicating a good correlation between the model's results and Pipe Flow's.

The model is then applied to a case study virtual pipeline based on the Alberta Products Pipeline (APPL), which spans approximately 300 kilometers between Edmonton and Calgary. In order to respect the CFL condition, an algorithm is implemented that interpolates hand-picked points along the pipeline right-of-way and ensures a consistent grid size per pipeline segment.

CHAPTER 5: AI-Based Blended Hydrogen Leak Detection System

The leak detection system (LDS) is developed, detecting leaks with a combination of a conventional AI classifier (CNN), an explainable AI classifier (ANFIS), and an ensemble learning technique (model stacking) to fuse both models into a single output. Leak localization is achieved through the negative pressure wave (NPW) method.

The system is trained and tested on a substantial synthetic dataset generated based on the specifications of Case Study 2, a virtual pipeline between the geography of Edmonton and Calgary. This dataset covers various leak scenarios, sizes, and other influencing factors to simulate realistic conditions as closely as possible. Standard data splitting techniques, such as train-test splitting and cross-validation, are used to evaluate the system rigorously. This methodology ensures that the system is not just accurate but also transferable.

After testing, the performance metrics of each classifier and the ensemble model are examined in detail. Various metrics such as accuracy, precision, recall, and F1 Score are used to understand each model's strengths and limitations comprehensively. This analysis reveals areas where each model excels and where improvements are necessary.

5.1 Leak Detection System (LDS) Overview

The leak detection system (LDS) consists of two interconnected subsystems: leak detection (LD) and leak localization (LL). A schematic representation of the system decision flowchart is shown in Figure 5.1. The system receives pressure and flow rate data through existing instrumentation, such as Supervisory Control and Data Acquisition (SCADA) systems.

This raw data is then processed, classified by the LD subsystem, and determined to fall

into the regular operation category (no leak was detected) or irregular operation (leak was detected). In the case of regular operation, the processing loop ends. On the other hand, if a leak is detected, the data proceeds to the leak localization step, which utilizes the same data to identify the leak's location.

The LDS is engineered to be compatible with pipelines initially designed for natural gas but repurposed for hydrogen blends. Given this premise, the system operates under the assumption of using only the most basic, existing field instrumentation for pressure and flow rate measurements. While incorporating additional sensors, such as mass flow rate meters, could substantially enhance the LDS's performance, their high cost has led to their exclusion in the system's design.

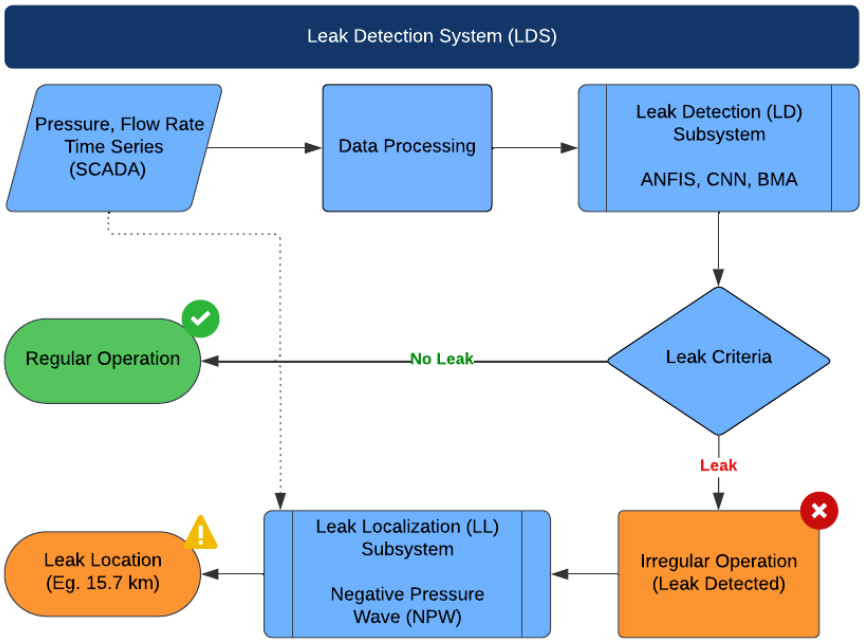


Figure 5.1: Leak Detection System Decision Flowchart.

The leak detection follows the processing steps that are shown in Figure 5.2. Pressure and

flow rate data first undergo a common pre-processing step, which involves segregating real-time data into windowed segments of 240 seconds each, with a 12-second overlap, and normalizing the data. Then, the data is divided into two separate streams.

The first stream undergoes a frequency analysis step and is classified by the conventional AI classifier, the convolution neural network (CNN). The second stream has nondimensionalized features extracted from it and is classified by the explainable classifier, the Adaptive Neuro-Fuzzy Inference System (ANFIS). Both models are fused using an ensemble learning technique called model stacking. Finally, the system has a binary output for either a leak or not a leak.

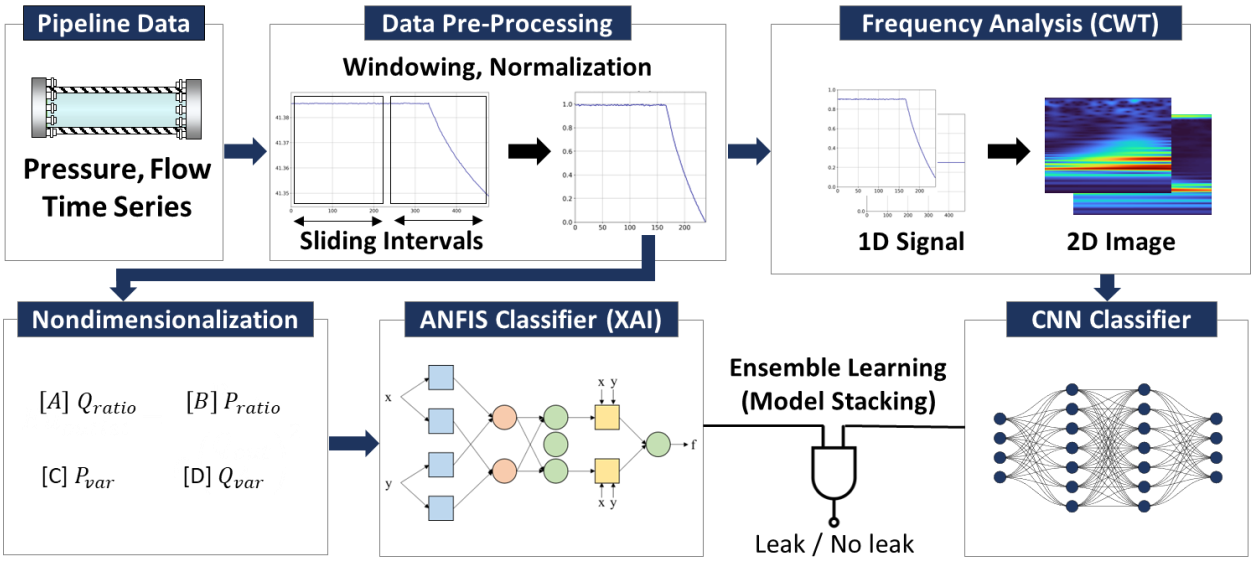


Figure 5.2: Overview of the Leak Detection (LD) Subsystem.

Once a leak is detected, the procedure progresses to the leak localization phase, as depicted in Figure 5.3. This phase employs the raw SCADA data before any pre-processing for its analysis. Specifically, the Negative Pressure Wave (NPW) method is applied to determine the leak's location [73].

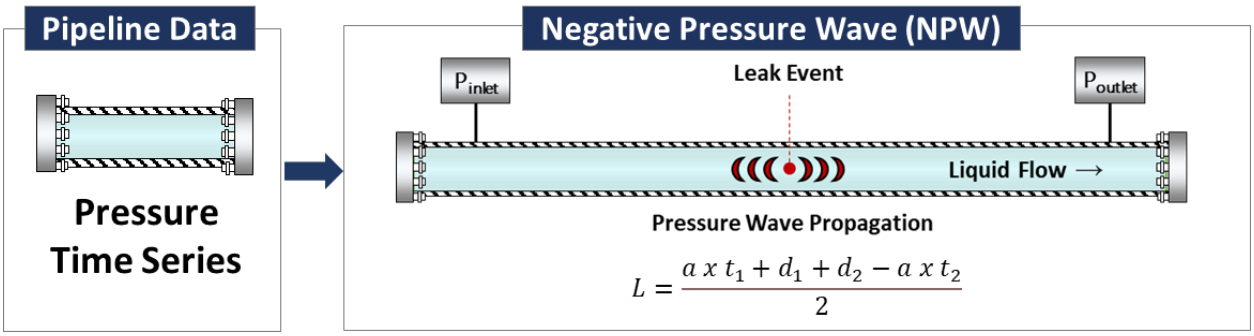


Figure 5.3: Overview of the Leak Localization (LL) Subsystem.

In subsequent sections, the data pre-processing techniques are thoroughly examined. This is followed by an extensive discussion on conventional, explainable, and ensemble learning methods. The mechanics of leak localization are then outlined, and the final part provides a comprehensive overview of the system's implementation and operational procedures.

5.1.1 Data Pre-Processing

The data pre-processing step is essential for any system that relies on machine learning models, as it ensures consistency with the input data, reducing potential biases and inaccuracies. This system employs two primary techniques: windowing and normalization.

The raw pipeline data is windowed using sliding intervals, each encompassing 240 seconds and a 12-second overlap between successive windows. Those intervals allow for the effective segmentation of continuous data, providing a manageable and structured format for further processing. The window length was chosen to balance the granularity of the data and computational efficiency. This standard length enables the system to detect and respond to changes within a reasonable timeframe while maintaining the complexity of the model at an acceptable

level.

The overlap between windows is designed to ensure continuity and completeness of the information. In the context of leak detection, this overlap could be the difference between capturing a leak event and missing it entirely, mainly if it occurs near the boundary of two windows. The overlap adds an additional level of certainty to the detection process, ensuring no significant event is overlooked due to unfortunate timing.

However, the implementation of this windowing method does impose a limit on the system's response time. Given the window length and the overlap, the minimum time for the system to reliably detect a leak event is roughly 12 seconds. Given the scale and potential consequences of pipeline leaks, this response time is acceptable.

Normalization is the second pre-processing step applied to the data. In the context of machine learning, normalization is a scaling technique that adjusts the values measured on different scales to a standardized scale. The system is normalized to a range between 0 and 1, as shown in equation (5.2).

$$y_{norm} = \frac{(y - y_{min})}{(y_{max} - y_{min})} \tag{5.1}$$

where y is the original value, y_{norm} is the normalized value, y_{min} is the minimum value within the dataset for that feature, and y_{max} is the maximum value within the dataset. y_{min} is scaled to 0 after the normalization, and y_{max} to 1.

This step is important to ensure that no particular feature dominates others due to its scale, which can lead to a bias in the trained model. Moreover, it helps with transferability by training the system on the relative changes in measurements rather than their absolute values.

5.1.2 AI Classifier (CNN)

The conventional AI receives the pre-processed pressure and flow rate time series data and transforms it into scalograms using the continuous wavelet transform (CWT). As a time-frequency analysis technique, CWT represents the signal's frequency content variation over time, revealing patterns and features that can be used for leak detection.

The Morlet wavelet is chosen for the transformation due to its well-balanced localization properties in both the time and frequency domains [74]. This choice ensures effective detection of transient events, such as leaks, in the time series data. Figure 5.4 illustrates that process and compares it with the original pressure time series with their respective scalograms for both the leak and no leak scenarios.

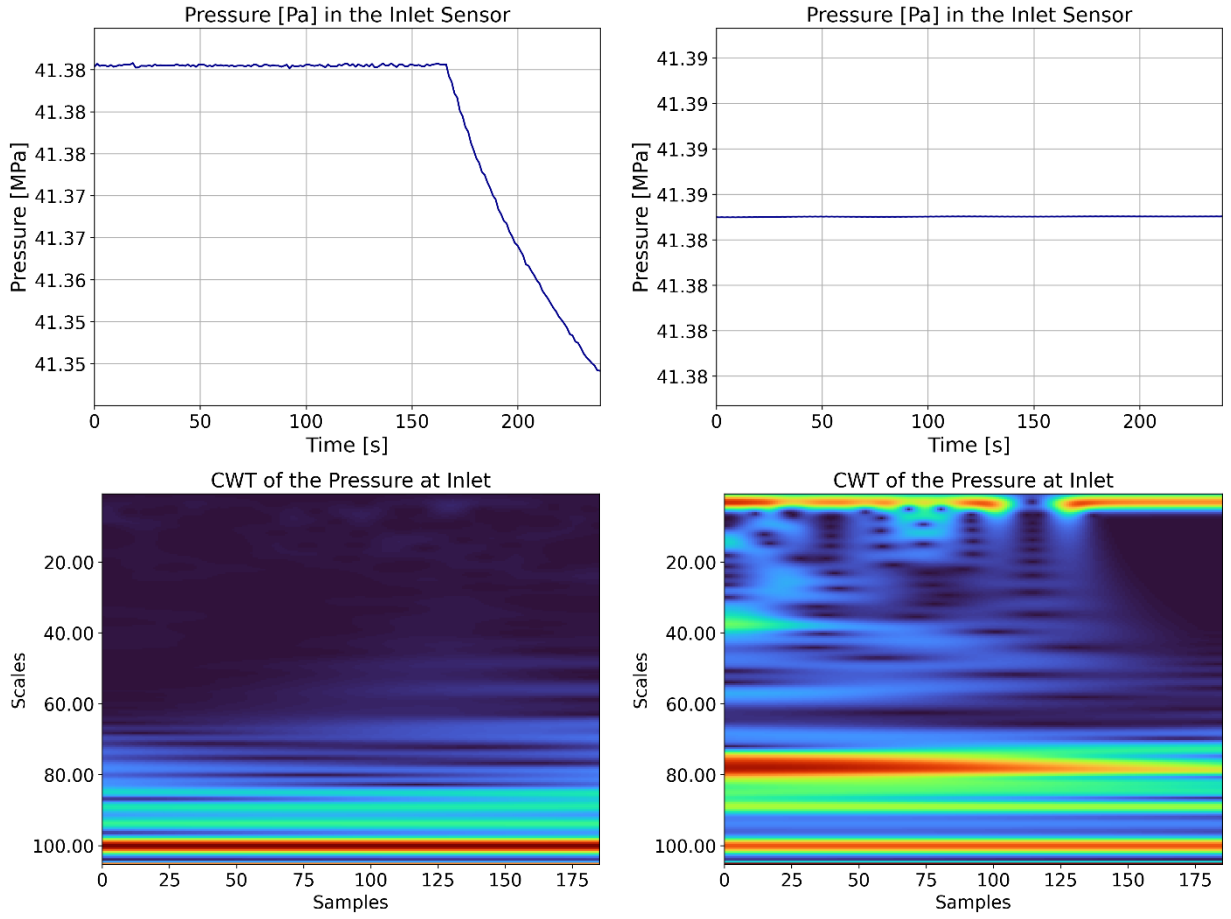


Figure 5.4: Comparison of CWT - Leak and No Leak Events.

The scalograms in the figure are shown in scales, which can be related to pseudo-frequencies using the relation in equation 5.2 [75]. However, it is important to note that the figure is stripped of any axis's information in the classification process, and only the plot is used for classification.

$$f_s = \frac{f_c}{a_s \cdot T_s} \tag{5.2}$$

where f_s is the pseudo-frequency (Hz) at a given scale a_s, f_c is the center frequency of the mother wavelet (Hz), Morlet in this case, and T_s represents the sampling period (s).

A high-pass filter is applied to remove unwanted low-frequency components in the resulting CWT, thereby avoiding training the classifier on noise, which could lead to overfitting unrelated features. High-pass filters, in general, are applied to allow frequencies above a specific cutoff to pass through while attenuating frequencies below that threshold [68]. The high-pass filter with a cutoff of 0.01 Hz is selected for this system. This cutoff removes the signal's low-frequency noise and any Direct Current (DC) components. It is aimed to preserve the high-frequency transient events, such as leaks, and improves the classifier's performance and robustness against noise. Figure 5.5 illustrates the impact of this high pass filtering on the CWT using the same pressure time series presented in Figure 5.4.

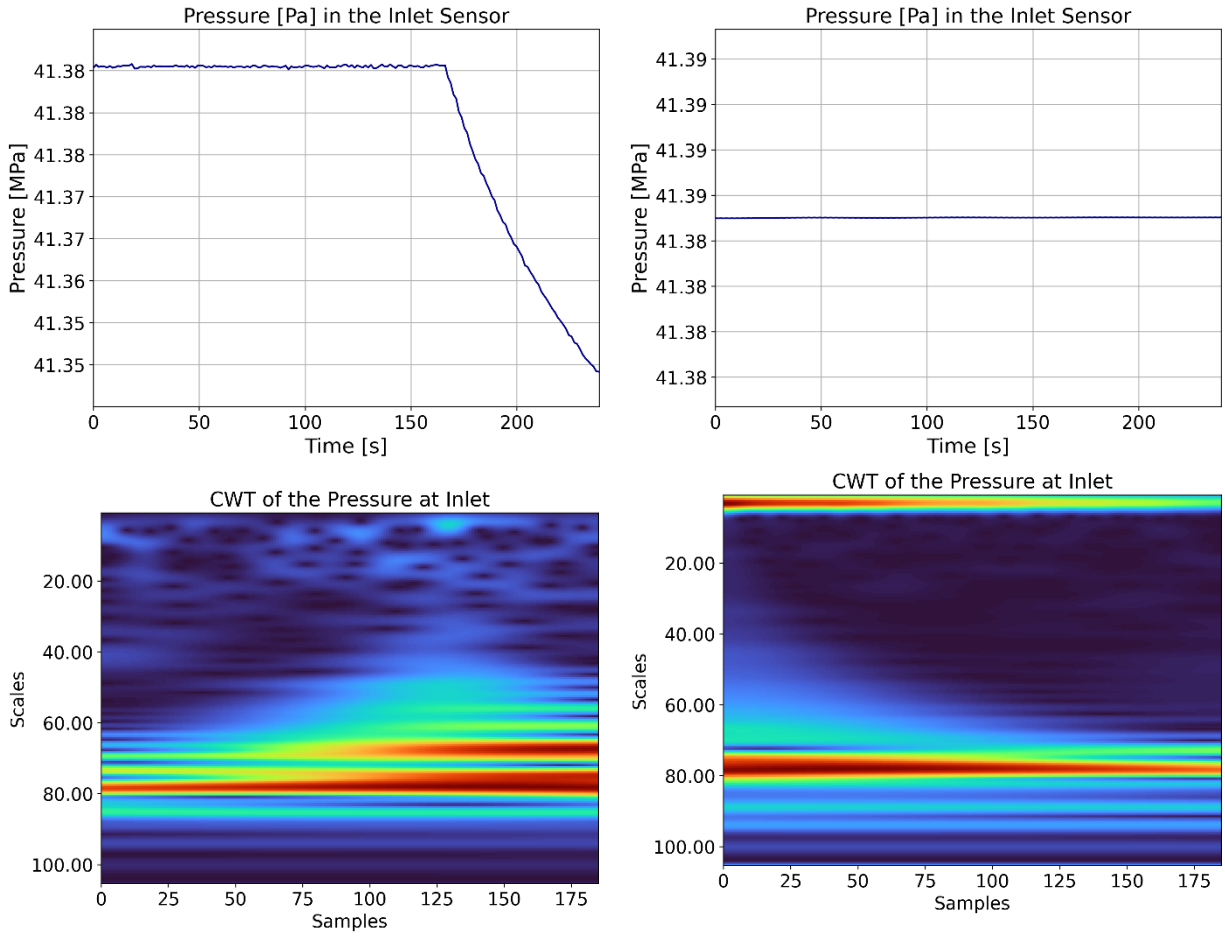


Figure 5.5: Comparison of Filtered CWT - Leak and No Leak Events.

The scalograms then get classified by the convolutional neural network (CNN). This type of classifier was chosen because of its high effectiveness in image recognition tasks, particularly in identifying and extracting patterns and features from complex data [76].

The CNN architecture is shown in Figure 5.6. It receives inputs of shape (224, 224, 3), representing the 224 x 224-pixel scalogram with three color channels (RGB). The input undergoes processing through five blocks of convolutional layers, each followed by a max-pooling layer, with an increasing number of filters from 32 to 512. After the final max-pooling layer, the model is flattened, and two fully connected layers are added. The output layer has a single node with a

sigmoid activation function, outputting a probability between 0 ('no leak') and 1 ('leak'). The model is compiled with the Adam optimizer [77] and binary cross-entropy loss function, suitable for binary classification problems [78], with accuracy and precision as the performance metrics.

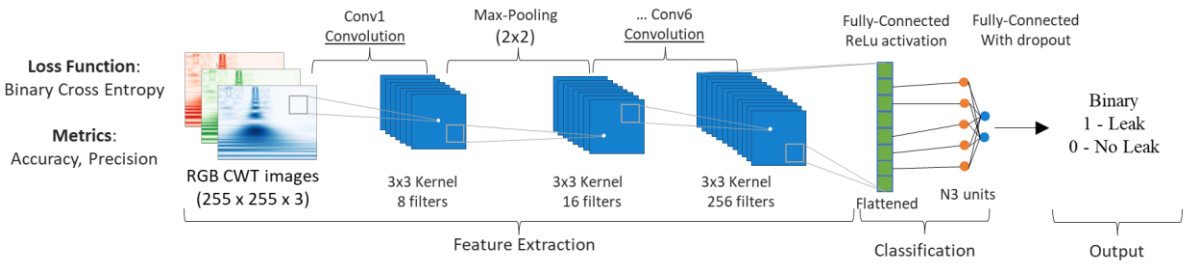


Figure 5.6: Architecture of the Conventional Classifier (CNN).

It should be noted that this custom CNN layout was developed for this leak detection system instead of using a pre-existing architecture to meet the requirements of processing continuous incoming data with limited computational power. A tailor-made architecture balances the model's complexity and performance while minimizing resource consumption, ensuring efficient and reliable leak detection operation in real-time.

5.1.3 Explainable AI Classifier (ANFIS)

The explainable AI classifier is an Adaptive Neuro-Fuzzy Inference System (ANFIS), an artificial neural network based on the Takagi-Sugeno fuzzy inference system. Jang officially published the system in 1993 [79] and has since been widely used in fields requiring pattern recognition, prediction, and control.

ANFIS operates by applying a set of predefined rules modeled using fuzzy logic. One of its most significant attributes is its use of membership functions, which transform crisp input

values into degrees of membership in fuzzy sets. In the context of this leak detection system, the membership function would represent the degree to which a particular data pattern belongs to a 'leak' or 'no leak' category.

ANFIS receives the same pre-processed pressure and flow rate time series data as the conventional AI classifier. However, unlike the latter, ANFIS does rely on frequency analysis processing. Instead, it utilizes different nondimensional features extracted from the data. This step aims to identify parameters in the sensor data that signify a leak event and represent physical changes within the pipeline. These features were carefully selected for their connection to physical measurements, enabling operators to comprehend the characteristics of a leak event better when an alarm is triggered.

The selected parameters, as shown in Table 5.1, are analyzed statistically within the time window of the pre-processing step, with each parameter evaluated based on its respective statistical property. A combination of domain expertise knowledge and empirical testing selected the features.

The first pair of features, namely the Flow Ratio (feature A) and the Pressure Ratio (feature B), are computed as the time-averaged ratios of the outlet to the inlet values for flow rate and pressure, respectively. These averages are calculated over the entirety of the observation window.

The latter two features - Pressure Outlet Variance (feature C) and Flow Rate Outlet Variance (feature D) - measure the variance of changes in the outlet pressure and flow rate relative to their initial values at the onset of the observation window. These variances offer a statistical measure of the fluctuation in these parameters, which can highly indicate a leak.

Table 5.1: Nondimensional Features, Definition, and Statistical Properties.

Feature	Definition
A. Flow Ratio (Inlet/Outlet)	$Q_{ratio} = \text{avg}\left(\frac{Q_{out}}{Q_{in}}\right)$
B. Pressure Ratio (Inlet/Outlet)	$P_{ratio} = \text{avg}\left(\frac{P_{out}}{P_{in}}\right)$
C. Pressure Outlet Variance	$P_{out,var} = \frac{\text{var}\left(\text{P}_{\text{out,t}} - \text{P}_{out,init}\right)}{P_{out,init}}$
D. Flow Rate Outlet Variance	$Q_{out,var} = \frac{\text{var}\left(\text{Q}_{\text{out,t}} - \text{Q}_{out,init}\right)}{Q_{out,init}}$

The system maps the input space through Gaussian combination membership functions (gauss2mf [80]). Each input variable is associated with four membership functions. This function type combines two Gaussian distributions, which can provide a more nuanced fit to the data than a single Gaussian distribution. The membership functions were chosen by trial-and-error optimization, where the one with the best performance was ultimately chosen.

Initially, a grid partitioning approach is employed to configure the Fuzzy Inference System (FIS). This technique divides the input space into multiple grids, each with fuzzy rules, allowing for a more detailed interpretation of the input data and serves as a foundational structure upon which the Gaussian combination membership functions can be applied for more nuanced analysis.

In summary, the ANFIS model employs multi-dimensional inputs, leverages the complexity of Gaussian combination membership functions, and utilizes a grid partitioning technique for the initial creation of the Fuzzy Inference System (FIS). Specifically, the model operates on four nondimensionalized features, translating their relationships into interpretable outputs.

5.1.4 Ensemble Learning (Model Stacking)

The conventional AI (CNN) and explainable AI classifier (ANFIS) integration is achieved using an ensemble learning strategy called model stacking. This approach combines individual model predictions by applying a meta-model to generate the final classification [81]. In the context of this system, a logistic regression model serves as the chosen meta-model.

Mathematically, each instance x in the dataset is first predicted by each classifier, the CNN ($P_{CNN}(x)$) and ANFIS ($P_{ANFIS}(x)$). Those two probabilities are then stacked into a new feature vector. The meta-model (logistic regression) is then trained on these stacked features to predict the ground truth. This relationship can be represented as:

$$y = \frac{1}{1 + \exp(-z)} \tag{5.3}$$

$$z = \beta_0 + \beta_1 \cdot P_{CNN}(x) + \beta_2 \cdot P_{ANFIS}(x) \tag{5.4}$$

where β_0, β_1, and β_2 are the parameters learned during the training of the logistic regression model.

This fusion of methods through model stacking allows the system to benefit from the CNN's high accuracy and pattern recognition capabilities while also incorporating the transparency and interpretability of the ANFIS model. As a result, the final output provides both an accurate leak detection and a transparent explanation of the decision-making process, enhancing the overall reliability and trustworthiness of the leak detection system.

5.1.5 Leak Localization

The leak localization process is only executed after the leak detection classifiers have identified a leak. This sequential arrangement is necessary because once the system moves to the leak localization stage, it operates assuming a leak exists. This precedent enables the negative pressure wave (NPW) method to focus exclusively on locating the leak.

The method is illustrated in Figure 5.7. When a leak occurs in a pipe, it generates a pressure wave that propagates in both directions: towards the inlet and outlet of the pipe [73]. This wave can be detected by pressure sensors placed along the pipeline as perturbations in the pressure signal, proportional to the leak size.

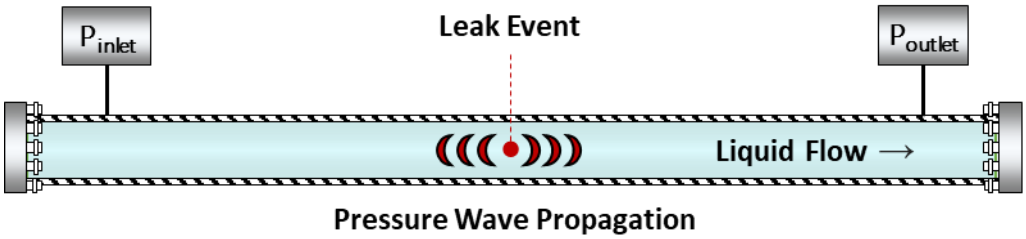


Figure 5.7: Illustration of the Negative Pressure Wave (NPW) Method.

Since the pipeline has a sensor upstream and downstream from the leak, those two sensors detect the wave at different times. The difference in this leak detection time and the known wave speed allows the leak's location to be estimated. This method is well established in the literature [73] and mathematically can be defined as follows:

$$L - d_1 = a \; x \; t_1 \tag{5.5}$$

$$d_2 - L = a \; x \; t_2 \tag{5.6}$$

where d_1 is the distance to the inlet sensor and d_2 to the outlet sensor. L is the unknown leak location, t_1 and t_2 are the time when the inlet and outlet sensors detect the wave, and finally, a is the wave speed of the fluid in the pipeline.

By solving this system of linear equations, the leak location can be determined with:

$$L = \frac{a\ x\ t_1 + d_1 + d_2 - a\ x\ t_2}{2} \tag{5.7}$$

The equation for leak localization, while being precise and mathematically straightforward, presents a considerable challenge regarding the practical aspects of identifying the precise times t_1 and t_2 at which the negative pressure wave arrives at the upstream and downstream sensors. Standard thresholding methods for signal anomaly detection are insufficient, as they do not adapt well to different operating conditions, like variations in inlet/outlet pressures and flow rates, which significantly impact the magnitude of pressure drops due to leaks.

To address this challenge, a specialized algorithm was designed to detect the pressure wave's arrival based on the relative changes in the pressure signal. The algorithm first calculates the first-order difference of the entire pressure signal, which is mathematically represented as:

$$\Delta P[i] = P[i+1] - P[i] \text{ for i} = 1, 2, \ldots \text{n} - 1 \tag{5.8}$$

where $\Delta P[i]$ is the first-order difference for an array of n pressure values $P[i]$.

The algorithm then identifies the maximum negative drop within these first-order differences, a proxy for the most significant pressure drop in the given time window. Subsequently, it flags the first-time index where the pressure drop is at least 10% of this maximum negative drop.

The main advantage of this methodology lies in its adaptability. By utilizing a relative threshold based on the maximum negative drop, the algorithm is robust to various operational conditions and pipeline configurations.

5.1.6 System Implementation and Operation

The system is implemented using two major software platforms: Python™ [82] and MATLAB™ [83]. Python version 3.10 is utilized to implement the Convolutional Neural Network (CNN) classifier, model fusion, and leak localization components of the system. Python's robust library support, particularly the Keras [84] and PyTorch [85] libraries, was a factor in its selection. These libraries offer comprehensive tools for deep learning, making them crucial in creating and training neural networks, like the CNN employed in this system.

Conversely, MATLAB™ is employed for its capabilities in high-performance technical computing. It is specifically used to develop the ANFIS and to generate batch data in conjunction with the RTTM. MATLAB excels in handling matrix operations and offers an expansive toolbox that aids in crafting and refining the ANFIS model. Its support for sparse matrix calculations and parallel computing proved particularly beneficial in expediting the training process.

Throughout this study, the system's training and testing phases utilized RTTM data that simulated the flow of blended natural gas and hydrogen through a hypothetical pipeline set in Alberta's geography. The primary reason for this approach stems from the scarcity of actual pipelines transporting such gas blends. Despite this, examining the potential limitations and application in real-world pipeline systems remains crucial.

The first important consideration is with regards to field instrumentation. The system's design hinges on a minimal set of anticipated field instrumentations, notably pressure and flow rate sensors. However, if a pipeline is equipped with additional sensors like those for density or

even direct mass flow rate measurements—such as Coriolis flow meters—incorporating these into the LDS's feature engineering process can enhance its functionality and boost its accuracy.

During the data generation phase, pipeline segments were intentionally chosen to reflect diverse spatial and temporal steps, leading to varying sampling rates. This intentional variation aims to demonstrate the system's versatility across non-uniform sampling rates. The sampling rate for each test, determined by the CFL condition, fluctuated based on initial parameters and the length of the pipeline segment, typically ranging between 1 to 2 Hz. However, it's crucial to understand that in real-world implementations, any significant decrease from these sampling rates (i.e., slower rates) might require system adjustments. Such modifications could include expanding the window size, altering the CWT wavelet selection, or fine-tuning the feature engineering process.

5.2 Data Generation, System Training and Validation

The training and validation dataset for the system is generated using the RTTM described in Chapter 4. Simulation is generated in batches by varying parameters defined within the design constraints outlined in Table 4.3. The parameter selection aims to emulate realistic operating conditions like fluctuating operational demands, changes in gas mixture compositions, and transient events.

Each simulation runs for the previously mentioned window size of 240 seconds per pipeline segment. This simulates different snapshots of measurements in a real-world scenario. Three parameters—gas mixture composition, inlet boundary conditions, and outlet boundary conditions—are randomized within specified limits. The gas mixture is randomized between 50% hydrogen and 50% methane to 100% hydrogen and 0% methane. Inlet and outlet boundaries are

randomized by up to ±7.5% of their nominal values.

In addition, transient events are incorporated into 50% of the simulation cases. These events are characterized by random durations and magnitudes, varying up to ±10% of the nominal boundary condition values. Those transients allow the model to simulate extreme conditions, like sudden valve openings or anomalous compressor operations, creating rigorous testing scenarios for the Leak Detection System (LDS).

Table 5.2: Training Dataset Conditions for Case Study 2 Pipeline.

Parameter	Value
Gas Mixture	[50% H_2, 50% CH_4] to [100% H_2, 0% CH_4]
Flow Rate BC [m^3/s]	Nominal Flow Rate ± 15%
Pressure BC [Pa]	Nominal Pressure ± 15%
Transient Condition	50% Yes / 50% No
Transient Magnitude	Random (Up to 10% BC value)
Leak Case	50% Yes / 50% No
Leak Effective Area [mm^2]	86.1 to 795.5 (step size = 25.3)
Leak Location [%]	30 to 70% (step size = 5%)
Leak Start Time [%]	40 to 60% (step size = 5%)

Finally, the leak characteristics are defined. 50% of the cases generated have leaks, and 50% do not, allowing the LDS to be trained in normal and abnormal operations. The remaining parameters concern the leaks' location, duration, and sizing. The leak sizing starts at an effective area of 86.1 mm^2, (Equivalent diameter of 10.4 mm) and increases to 795.5 mm^2 (Equivalent

diameter of 31.8 mm) in 29 steps. To provide some perspective, the smallest leak corresponds to approximately 0.04% of the pipeline cross-sectional area, whereas the largest one corresponds to 0.40% of the cross-sectional area. All those parameters are summarized in Table 5.2.

This data generation process is structured to account for the unique flow behaviour across different pipeline segments with their own topographical and properties. The varying segment lengths also impact the data with different pressure and velocity profiles. In total, 41,760 simulations are generated, 5,220 per pipeline segment.

It is important to note that using randomized parameters in data generation ensures the system is exposed to a broad range of scenarios, preventing overfitting, and promoting the model's ability to generalize well to new and unseen conditions.

Following common practice in the field, the dataset was split 50% for training, 25% for validation, and 25% for testing. During the training phase, the models were only trained and fine-tuned using the training and validation sets, respectively, without any exposure to the testing dataset. The results reported in this study strictly pertain to the testing set. Adopting this rigorous data partitioning strategy validates the models' generalization capabilities and mitigates the risk of overly optimistic performance estimates [86].

On average, each simulation required about 96 seconds of computational time and ran on 16 parallel threads. In practice, the number of simulations can be adapted to cater to specific pipeline risk profiles by, for example, dynamically adjusting specific segments to have a finer mesh. Additionally, computational efficiency could be improved by leveraging more robust computing resources. This is shown in Table 5.3.

Table 5.3: Simulation Parameters and Computation Time for Each Pipeline Segment.

Pipeline Segment	Number of Simulations	Data Split	Spatial Nodes	Pipeline Length	Computation Time
Segment 1	5,220	50% Training 25% Validating 25% Testing	45	22 km	140 hours
Segment 2				17 km	148 hours
Segment 3				59 km	143 hours
Segment 4				44 km	144 hours
Segment 5				33 km	143 hours
Segment 6				43 km	141 hours
Segment 7				34 km	148 hours
Segment 8				42 km	146 hours
Total	41,760	-	360	294 km	866 hours

5.3 Leak Detection System (LDS) Performance

The leak detection system is evaluated with the data described in Table 5.3, generated using the RTTM to simulate flow in the virtual pipeline in case study 2. The main goal is to assess its effectiveness across various conditions, verify its adaptability, and establish the viability of leak detection for blended hydrogen transportation.

Initially, an analysis is conducted on the performance of individual classifiers, the conventional classifier (CNN), the explainable classifier (ANFIS), and the ensemble learning technique (model stacking). The leak localization process is then examined using the negative pressure wave (NPW) method. Ultimately, the system is evaluated as a unified entity to assess its capabilities and limitations comprehensively.

5.3.1 AI Classifier (CNN)

The conventional AI classifier is a Convolutional Neural Network (CNN), which works by classifying images generated by the Continuous Wavelet Transform (CWT) of the pressure signal. Those images comprehensively represent the pressure variations within the pipeline during the windowing signal and allow CNN to discern patterns indicative of leaks. This is why, during the training process, it is crucial to include a variety of scenarios that the system could encounter in practice, such as leaks in multiple locations with multiple durations and sizes, as well as unplanned events (e.g., transients) that could harm leak interpretation.

While it is impractical to display all the images generated by this study, Figure 5.8 shows a hand-picked selection of those images to help illustrate the differences between some scenarios. The figure shows four examples for leak cases and four for no leak cases. The cases in the top row represent cases without any transient added, whereas the bottom row represents cases with transients added. Note that the images here are shown without an axis, as this is how the neural network classifies them.

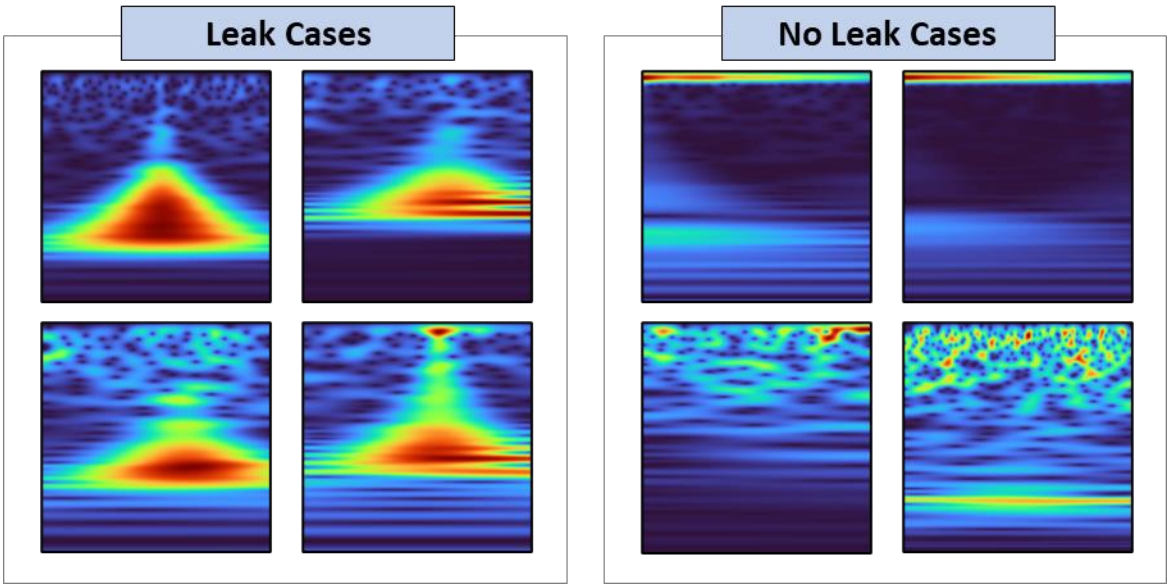


Figure 5.8: Examples of Continuous Wavelet Transform (CWT) for Leak and No Leak Events.

Even by simple observation, there is a clear contrast between signals with and without transient events added to them, showcasing the additional challenge that those events pose to the classification process. Nevertheless, it is still possible to visually detect the difference between leak and lo leak cases even to the untrained eye, suggesting the strength of the CWT processed images in making the leak signals more explicit and detectable. This visually recognizable differentiation between leak and no leak cases provides a solid basis for the CNN model to perform effective classification. It shows how choosing a proper wavelet function and data processing is important.

The performance of the conventional classifier on the pipeline for Case Study 2 is summarized in the confusion matrix displayed in Figure 5.9. As mentioned in the training section, the evaluation of the models is done exclusively on the testing data. In other words, 25% of the entire dataset, or 10440 cases.

The confusion matrix shows that, out of 5184 true negatives (instances correctly identified as not being leaks), there were zero false positives, meaning the model did not incorrectly predict any leaks, avoiding false alarms. On the other hand, the model correctly identified 5211 leaks as true positives, with only 45 false negatives. This is particularly noteworthy given the high cost of false negatives in leak detection, demonstrating that the classifier has an impressive recall rate.

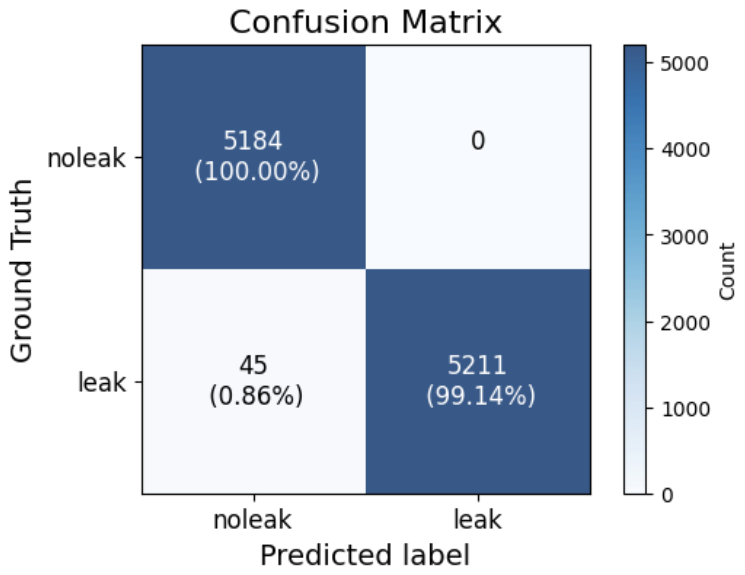


Figure 5.9: Confusion Matrix for the Conventional Classifier (CNN) on Case Study 2 Pipeline.

5.3.2 Explainable AI Classifier (ANFIS)

The Explainable AI Classifier employed is an Adaptive Neuro-Fuzzy Inference System (ANFIS). ANFIS is a hybrid neural network that leverages the power of both fuzzy systems and neural networks, combining the adaptive capabilities of the latter with the interpretability and structured knowledge representation of the former. While the convolutional neural network (CNN) uses frequency analysis of the pressure signal for classification, ANFIS uses nondimensionalized features for leak detection.

The feature selection for the nondimensionalization was discussed in the previous chapter and aims to transform the data so that the resulting features become independent of the units of measurement, making them scalable and comparable across various scenarios. The result of this process can be seen in Figure 5.10.

The figure comprises four hexbin plots, each corresponding to a unique feature. In a hexbin plot, the 2D data space is tessellated with hexagons, each serving as a 'bin' that encapsulates the

density of data points. The color of each hexagon reflects the average condition - either 'leak' (in red) or 'no leak' (in blue). This visualization method offers valuable insight into the efficacy of the feature engineering process by illustrating how each feature correlates with the predictive variable.

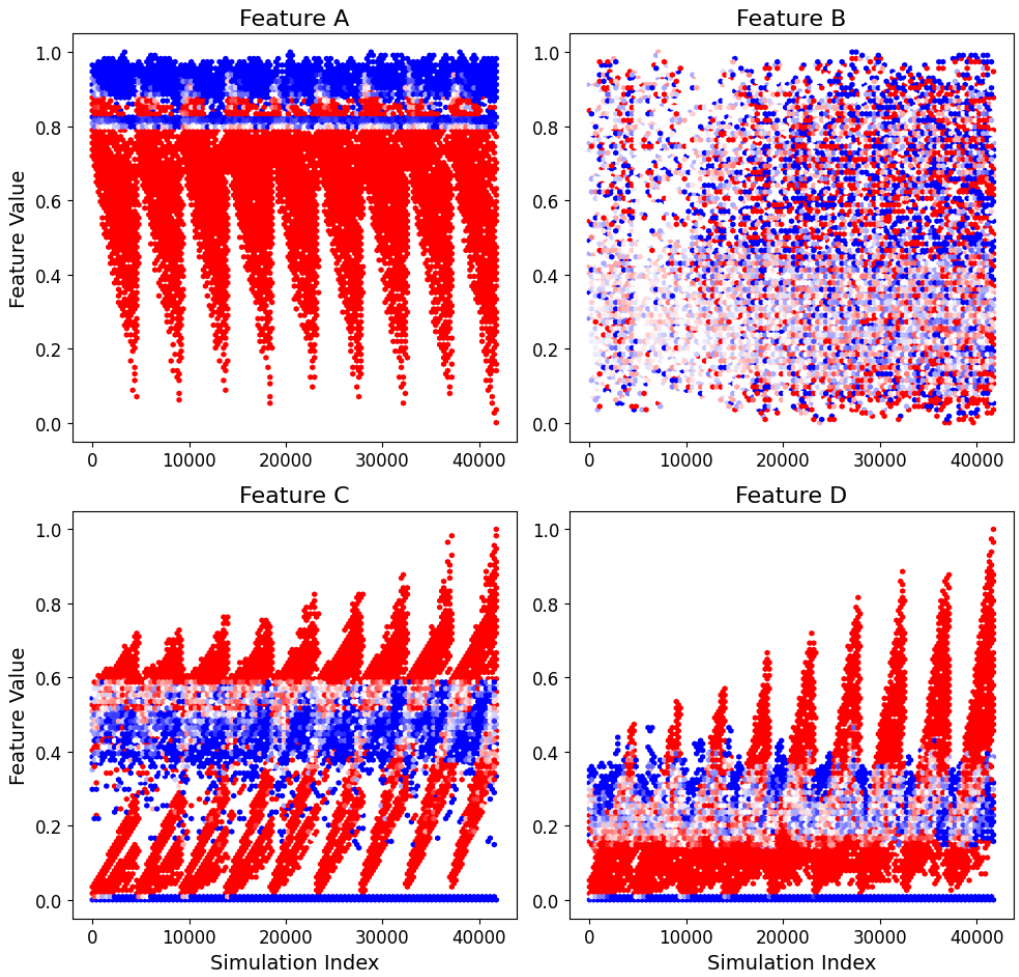


Figure 5.10: Nondimensional Features for Case Study 2 Pipeline, Highlighting Leak Events.

Notably, features A, C, and D show clear distinctions between leak and no leak cases, emphasizing the effectiveness of the feature selection process, with clear zones in certain magnitudes. For example, for feature A, magnitudes smaller or equal to 0.7 are exclusively represented by leak cases. Feature B exhibits a more mixed behaviour, mainly due to the influence

of the additional randomness included in the tests, reiterating the added challenge of those conditions for leak detection systems.

One of the most crucial elements of the ANFIS is its capacity to formulate membership functions (MFs) throughout the training phase. Membership functions act as mappings that translate each point in the input domain to a membership score, which falls between 0 and 1. These functions are vital in evaluating how much an input value aligns with a particular fuzzy set or class [87].

Each input variable is associated with four membership functions in the Leak Detection System (LDS) context. This choice represents a balancing act between model complexity and predictive accuracy. Employing a more significant number of membership functions may increase the model's precision but at the risk of overfitting. Conversely, utilizing fewer functions could compromise the model's ability to represent the data's intrinsic characteristics faithfully.

The membership functions and corresponding membership degrees are depicted in Figure 5.11. These plots serve as visual guides to understanding the extent to which specific feature values belong to defined categories. The x-axis signifies the magnitude of feature values, while the y-axis illustrates the corresponding membership degrees. A higher membership degree suggests a stronger association of the given feature value with the category specified by the MF. Membership function plots offer insights into how the ANFIS system interprets and classifies the data.

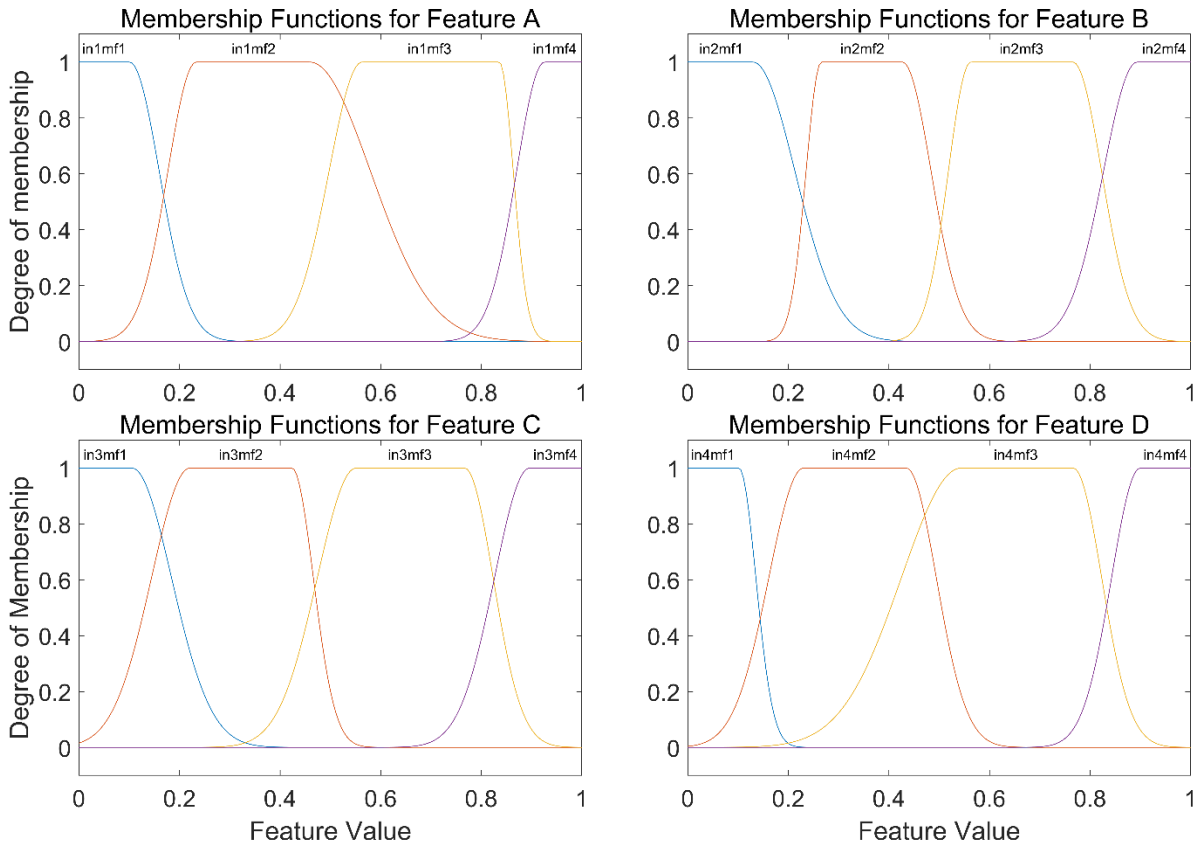


Figure 5.11: ANFIS Membership Functions for the trained Nondimensional Features.

In an effective fuzzy model, it is generally advantageous for the membership functions to have some degree of overlap, ensuring smoother transitions between different categories. Additionally, the membership functions should span the entire range of the input domain to avoid gaps, which could imply poor representation of specific input ranges [87]. An inspection of Figure 5.11 reveals that the features are well-distributed across their respective domains, and there is reasonable overlap between the membership functions, suggesting that the features have been appropriately selected.

The performance of the ANFIS classifier for the pipeline of Case Study 2 is summarized in the confusion matrix in Figure 5.12. The ANFIS model correctly identified 4825 instances as

not being leaks. However, there were 359 false positives where the model incorrectly predicted leaks, potentially leading to unnecessary interventions or alerts.

On the other hand, the classifier accurately identified 5212 leaks while yielding only 44 false negatives. Like the CNN model, the ANFIS model exhibits a strong recall rate, an important metric given the high cost and risk associated with false negatives in leak detection scenarios. Nevertheless, it is slightly inferior regarding false positives.

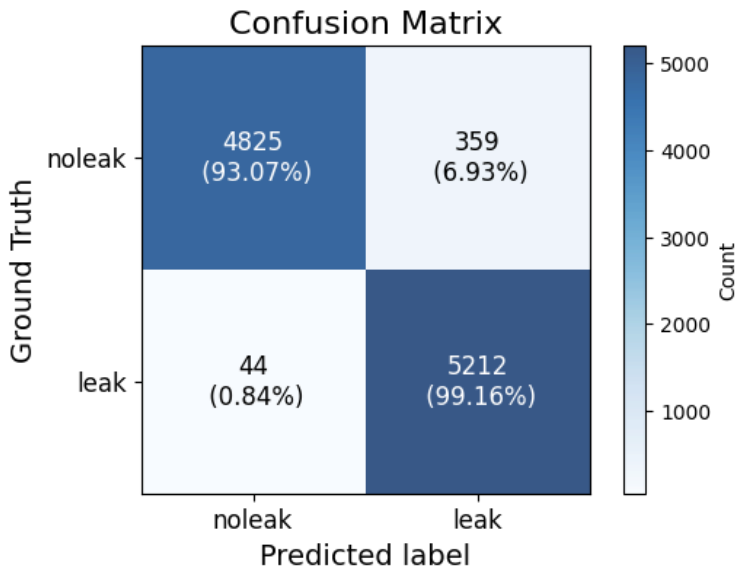


Figure 5.12: Confusion Matrix for the Explainable Classifier (ANFIS) on Case Study 2 Pipeline.

5.3.3 Ensemble Learning (Model Stacking)

After the individual prediction of both classifiers, the models are fused with an ensemble learning technique called Model Stacking that uses a logistic regression model. The output of this fused model is the actual, binary output of the system. It is important to recognize that the

explicability and transparency of the system are closely tied to the ANFIS output, but the overall performance is tied to the stacked model.

The decision boundary for the ensemble model is depicted in Figure 5.13. The decision boundary is a geometric representation in the feature space where the model switches its decision from one class to another. In this case, it marks the threshold between the "leak" (probability = 1) and "no leak" (probability = 0) conditions based on the predicted probabilities from both CNN and ANFIS classifiers.

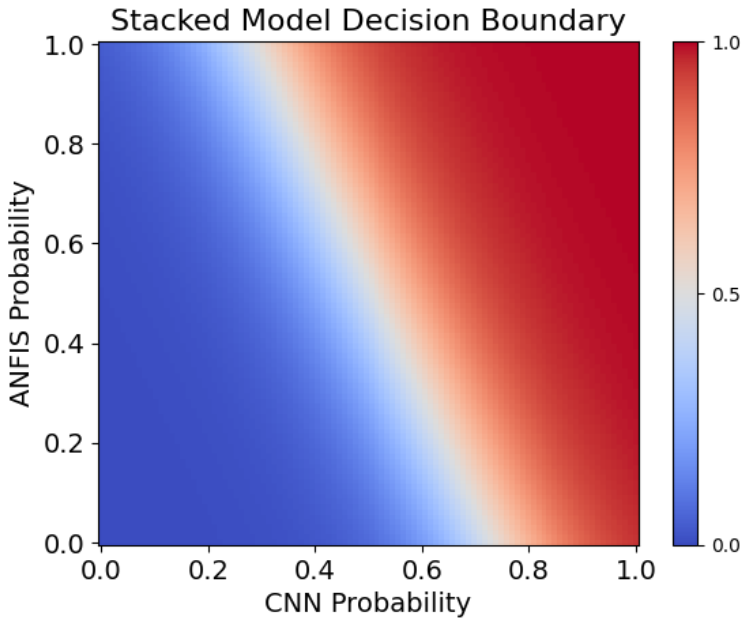


Figure 5.13: Decision Boundary of the Stacked Model.

The decision boundary is helpful to gain insight into the ensemble model's performance and decision-making process. For the CNN model's predicted probabilities less than 0.2, the ensemble model predominantly classifies instances as "no leak" regardless of ANFIS probabilities.

Conversely, when the CNN model's predicted probabilities exceed 0.8, the ensemble model largely classifies instances as "leak," irrespective of what the ANFIS model suggests.

In other words, The CNN model has a higher influence on the ensemble model's decision in scenarios where it is either very confident of a "leak" or very confident there is "no leak." This could be attributed to CNN's robust performance metrics. For all other scenarios in the middle, both classifiers have a decisive influence, where the ensemble model leverages the explainable features of ANFIS to make a more nuanced determination.

Finally, regarding the actual performance of the overall stacked model, the results are summarized in the confusion matrix shown in Figure 5.14. This stacked model identified all 5184 instances as true negatives, correctly classifying every non-leak event with zero false positives, matching the CNN model’s strong performance in minimizing false alarms.

The ensemble model identified 5218 true positives when accurately detecting leaks, surpassing individual classifiers. More notably, it reduced the number of false negatives to 38, lower than the two individual classifiers. This improvement is significant, given that failing to identify a leak can have serious consequences.

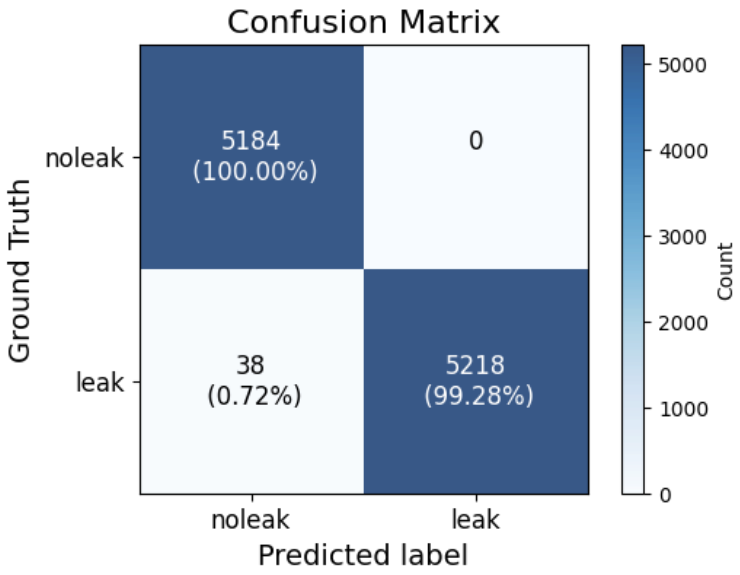


Figure 5.14: Confusion Matrix for the Ensembled Model on Case Study 2 Pipeline.

5.3.4 System Transparency

One of the most common challenges in deploying AI-based leak detection systems is their "black-box" characteristic. This lack of transparency hinders operator trust and poses compliance issues with regulatory standards. Addressing this challenge has been central to the architecture of the leak detection system developed in this study.

This architecture incorporates an Adaptive Neuro-Fuzzy Inference System (ANFIS) classifier as an Explainable Artificial Intelligence (Explainable AI). This design decision allows for greater transparency in the system's decision-making process. It enables the operators to analyze the nondimensionalized features outlined before and whether they are deviating from their normal operating ranges.

Nevertheless, it should be noted that the Convolutional Neural Network (CNN) generally outperforms ANFIS regarding raw performance metrics. This superior performance comes at the cost of explainability, introducing a trade-off scenario. Specifically, there could be instances where

the system's decision-making may lean toward the black box behaviour. This is particularly true when discordant outputs among the CNN, ANFIS, and the ensemble model are formed by stacking them.

Despite its limitations in raw performance metrics compared to the Convolutional Neural Network (CNN), the ANFIS classifier demonstrated specific strengths in the case study so far. Notably, ANFIS misclassified only 44 leak cases, representing a mere 0.84% of the total. However, the system experienced a higher rate of false alarms, misclassifying 359 no-leak cases, or 6.93% of the total, implying that most of the disagreements between the classifiers tend to result in false alarms rather than missed leaks. While false alarms are not entirely inconsequential, they are considerably less critical than failing to detect an actual leak, indicating a favorable error distribution for the system.

5.3.5 Leak Localization

The leak localization subsystem employs the Negative Pressure Wave (NPW) method. This physics-based approach does not necessitate data division into training and testing subsets like the other machine-learning methods. Instead, it is applied universally across all available data sets. The performance metrics selected to evaluate the method's efficacy are the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE).

The Mean Absolute Error (MAE) is a straightforward metric that calculates the average of the absolute differences between the actual and predicted leak locations, expressed in meters. Meanwhile, the Mean Absolute Percentage Error (MAPE) measures the same discrepancies but scales them according to the true values and then averages the absolute percentages.

The results are shown in Table 5.4. The MAE across all cases averaged 3711 m, with a MAPE of 30.4%. The pipeline segments, for reference, ranged between 22 to 60 kilometers in

length. Moreover, the table also shows the results stratified by pipeline flow conditions, which shows how imbalanced the bad performers are – cases with transients have a significantly larger MAE and MAPE than cases without those conditions.

Table 5.4: Leak Localization Performance for Case Study 2 Pipeline.

Pipeline Flow Condition	Mean Absolute Error (MAE)	Mean Absolute Percentage Error (MAPE)
All Cases	3711 m	30.4 %
Regular Operation	83 m	1.4 %
Transient Events	7505 m	59.4 %

Stratifying the results by pipeline flow conditions provides insights into the disparities in performance. The method achieved an MAE of 83 meters and a MAPE of just 1.4% for regular operating conditions, indicating high accuracy. cases involving transient flow conditions fared substantially worse, with an MAE and MAPE of 7505 meters and 59.4%, respectively. This discrepancy highlights the limitations of the NPW method when dealing with transient events, which often present as significant pressure drops that the algorithm might incorrectly attribute to leaks.

5.3.6 Case Study Pipeline Performance

The results of the individual classifiers and methods so far were discussed regarding their metrics. However, to analyze the combined model's performance more holistically, four key metrics will be used: Accuracy, Precision, Recall, and the F1 Score.

Accuracy is the ratio of correct predictions (both true positives and true negatives) to the total number of predictions [88]. It is a straightforward measure of the model's overall performance. Precision is the ratio of correctly predicted positive instances (true positives) to the total number of predicted positives (true positives and false positives) [88]. It reflects the system's ability to avoid false positives or alarms.

Recall, on the other hand, is the ratio of correctly predicted positive instances (true positives) to the total actual positives (true positives and false negatives) [88]. It quantifies the model's ability to identify all relevant instances and detect all real leaks. A high recall rate is critical in a leak detection system, as the cost of missing a leak (false negative) is significantly higher than that of a false alarm (false positive). Lastly, the F1 Score is the harmonic mean of precision and recall [88] and attempts to strike a balance between these two metrics. It is advantageous when the class distribution is unbalanced.

The results for AI-based classifiers (CNN, ANFIS, Stacked Model) are analyzed through their Accuracy, Precision, Recall, and F1 Score. The results for the leak localization method (NPW) are analyzed through the mean absolute error (MAE) and mean absolute percentage error (MAPE), introduced previously. The results for the performance of the overall system can be seen in Table 5.5.

Table 5.5: Overall System Performance for Case Study 2 Pipeline.

Subsystem	Method	Accuracy	Precision	Recall	F1 Score
Leak Detection	**Conventional (CNN)**	99.57%	100%	99.14%	99.57%
	Explainable (ANFIS)	96.14%	93.56%	99.16%	96.28%
	Ensemble Learning (Model Stacking)	99.64%	100%	99.28%	99.64%

Subsystem	Method	MAE	MAPE
Leak Localization	**Negative Pressure Wave (NPW)**	Overall: 3711 m Regular Operation: 82 m Transient Events: 7505 m	Overall: 30.4 % Regular Operation: 1.4% Transient Events: 59.4%

As expected by the result of the confusion matrixes, the ensemble learning approach stands out as the most effective, demonstrating the highest overall metrics. It marries the predictive solid capabilities of the conventional CNN model with the transparency and near-equivalent recall rate of the explainable ANFIS model. While the CNN model excels in accuracy and precision, the ANFIS model adds value with its strong recall rate and explainability.

In a leak detection system like this, the greatest confidence in predictive accuracy is best when both the CNN and ANFIS models concur. If there is a discrepancy between the two models, the responsibility shifts to the operator to assess whether additional investigation is necessary. The system further enhances its predictive transparency by incorporating ANFIS, an explainable AI model. This feature allows the operator to scrutinize the decision features and make an informed judgment.

This highlights the importance of system reliability; a false positive may result in

unnecessary expenditure to locate a non-existent leak, but the consequences of a false negative are much graver, potentially leading to environmental damage, loss of resources, or safety risks.

While evaluating the system's overall performance is essential, it is equally critical to identify its limitations and blind spots. Figure 5.15 shows the four performance metrics - accuracy, precision, recall, and F1 score - for each individual and stacked model. The performance metrics are stratified per leak effective area to visualize whether specific leak sizes are more challenging for the models.

A closer examination reveals that the CNN and the stacked model consistently exhibit strong performance across the entire spectrum of leak sizes. They maintain all performance metrics above the 98% threshold, irrespective of whether the leaks are small or large.

The ANFIS model, on the other hand, presents a different picture. Its performance metrics fluctuate between 87% and 95% for leaks smaller than 200 mm^2, indicating a particular challenge in detecting those smaller leaks. Interestingly, ANFIS compensates for this limitation when it comes to larger leak sizes. The recall rate, a metric pivotal for minimizing false negatives or undetected leaks, shows significant improvement, nearly reaching a 100% rate for larger leaks.

In other words, smaller leaks pose a more significant detection challenge for a complete output with explainable indicators due to ANFIS's slightly lower performance. However, the CNN and stacked models maintain robust reliability across all leak sizes, offering a more stable and reliable solution for practical applications.

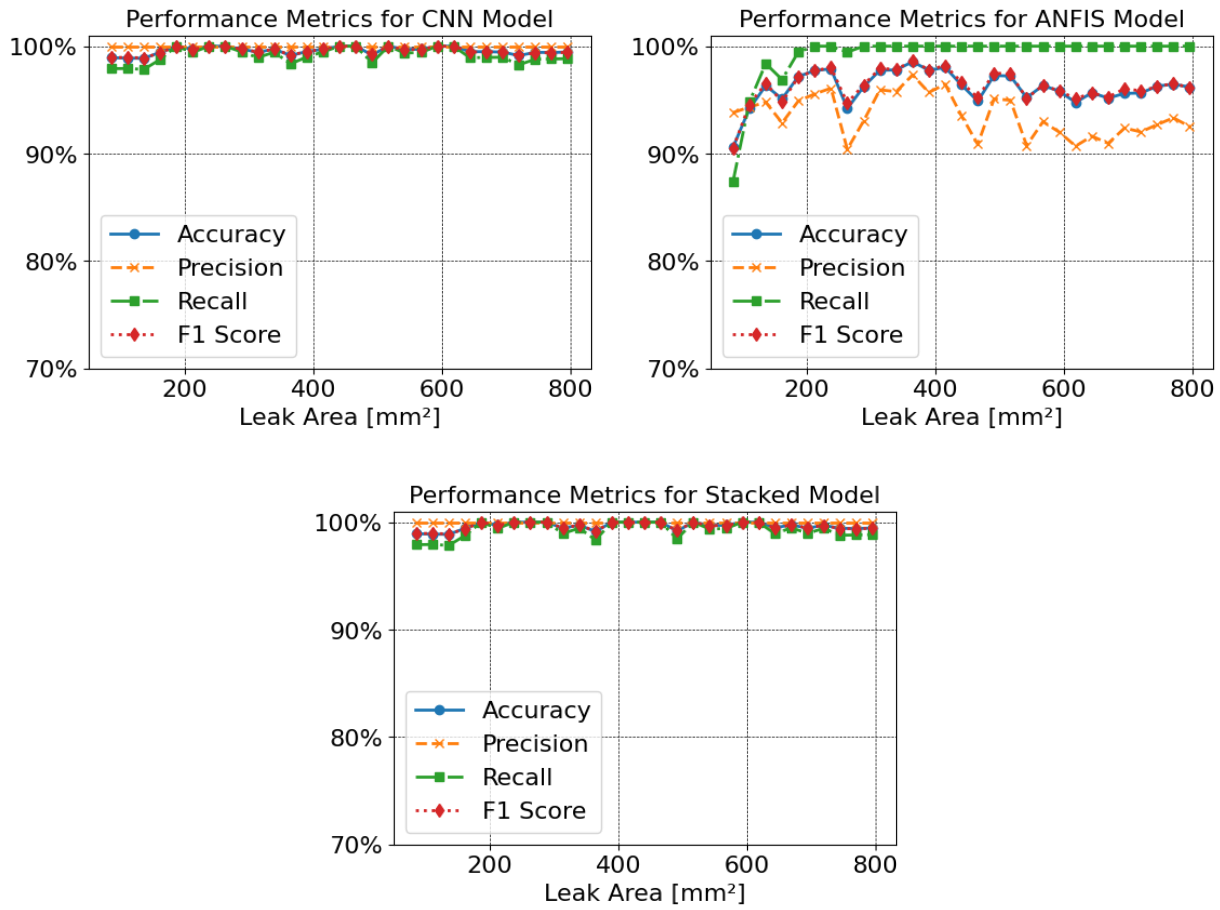


Figure 5.15: Performance Metrics by Pipeline Segment and Model.

5.4 Summary

The performance for the (LDS) was overall positive, as evidenced by the consistently high recall rates across all employed models. The conventional AI classifier (CNN) had a recall of 99.14%, the explainable AI classifier (ANFIS) of 99.16%, and finally, the ensemble learning (model stacking) of 99.28%. This is encouraging, as reducing the number of false negatives should be the goal of any leak detection system.

Additionally, the system demonstrated robust performance in handling data transients and synthetic noise, key factors that often introduce variability and uncertainty in real-world pipeline monitoring scenarios. This resilience to noise confirms the models' utility and robustness, even in

less-than-ideal conditions. However, a more nuanced analysis revealed that the explainable AI classifier (ANFIS) has some limitations with detecting leaks smaller than 200 mm^2. On the other hand, the CNN and the stacked model continue to demonstrate consistent performance, even when dealing with these more challenging, smaller leaks.

Regarding leak localization, the Negative Pressure Wave (NPW) method showcased promising results under standard testing conditions devoid of transient events. It achieved a Mean Absolute Error (MAE) of just 83 meters and a Mean Absolute Percentage Error (MAPE) of 1.4%. However, the performance took a noticeable hit when transients were introduced into the test scenarios. The MAE increased substantially to 203 meters, and the MAPE soared to 40.0%. This significant divergence indicates that while the NPW-based leak localization is reasonably accurate and reliable under standard operating conditions, its performance is compromised in more dynamic, transient-laden environments.

CHAPTER 6: CONCLUSIONS

As the world accelerates its transition towards clean and renewable energy sources to confront climate change, blended hydrogen and natural gas emerges as promising energy carrier alternative to fossil fuels. However, it still faces numerous challenges for safe transportation, with many aspects of its adoption yet to be fully understood and optimized. This study delves into these challenges, focusing on the issues of pipeline infrastructure compatibility and leak detection associated with hydrogen transportation.

This chapter represents the conclusions of this research, summarizing and pointing out its key findings. It describes an overview of the study and its scientific outcomes, acknowledges the assumptions and limitations, and offers suggestions for future investigations in the field.

6.1 Overview

The motivation behind this study was the growing demand for renewable energy sources and the subsequent shift towards incorporating hydrogen as an energy carrier. Despite its potential to significantly lower carbon emissions, hydrogen's safe and efficient distribution presents an array of challenges directly addressed in this study.

The research started with developing two distinct models for pipeline gas flow during steady and transient states. These models were crucial due to the limited availability of real-world data from the early state of hydrogen pipelines.

A comprehensive analysis of the feasibility of blending hydrogen into existing natural gas infrastructure was then conducted, focusing on the compatibility of blended hydrogen with existing pipeline materials. This analysis contemplated the influence of the blended gas on

pressure, flow, energy delivery, and compressor selection. It was shown to what extent the impacts would be of selecting higher concentrations of hydrogen for transportation in relation to the energy delivered and which scenarios would be completely impractical. Then, it was also detailed which modifications would be required in the pipeline for different scenarios.

Then, the issue of leak detection was tackled with the development of a comprehensive leak detection system specific to blended hydrogen. This system used Artificial Intelligence (AI) techniques. It was designed to use existing field instrumentations, overcoming the issue of requiring current pipelines (often underground) to deploy new sensors – something that could be cost-prohibitive and technically challenging.

However, the use of AI also has its challenges. One of those is the 'black box' problem. This problem happens when it is unclear how a particular decision or output is reached. This can raise operational and safety concerns, particularly in detecting and responding to leaks in a pipeline system. The second challenge is the problem of transferability. An AI system used for blended hydrogen would have to support a wide range of blends and pipe systems, and thus, the AI would have to be easily transferrable.

The LDS developed solved this problem by combining a conventional AI classifier with an Explainable AI classifier. The conventional classifier was chosen in this system due to its high effectiveness in recognition tasks, particularly in identifying and extracting patterns and features from complex data. In contrast, the explainable one was chosen to provide the necessary transparency into the system's decision-making processes.

Finally, the LDS addresses the transferability issue by relying on data generated by the blended hydrogen real-time transient model (RTTM). The system is trained on a broad data set spanning a wide range of hydrogen blends. This training method is one of the system's advantages,

which allows it to be tailored to different pipeline systems, providing a degree of resilience to future changes in the hydrogen infrastructure.

6.2 Scientific Outcomes

This study took a comprehensive and technical approach to examine the primary obstacles in blending hydrogen with existing pipelines. Its main goals were to create numerical methods for pipelines, analyze the feasibility of hydrogen blending, and develop a leak detection system. As a result, the study has made the following scientific contributions:

A. Gas Hydraulics Model Based Pipeline Feasibility Study

A gas hydraulic steady-state model was developed capable of analyzing the impacts of using existing pipelines for blended hydrogen transportation has been developed. The model can be used to delve into the specifics of a particular pipeline system and understand the inherent trade-offs. In the study, this was illustrated through the definition of a case study and the subsequent analysis of the practical impacts of using different hydrogen blends in this case study.

The model simulates the gas flow characteristics with the universal gas flow equation, thermal effects are modeled with the Joule-Thompson effect, and the friction factor is calculated with the Colebrook-White equation. The scenarios are calculated for different hydrogen blends and energy transmission values. The study suggested an optimal hydrogen concentration that minimally impacts the pipeline's performance and a comprehensive table of viable options with more nuanced trade-offs.

In the example of the case study given, a 322-kilometer transmission and 10 km distribution pipeline could be converted to transport a 60% hydrogen 40% methane blend and maintain 65% of its original energy delivery with the addition of a single compressor. Conversely,

if no modifications are possible due to budgetary constraints, it would be possible to transport a 30% hydrogen 70% methane blend and maintain 51% of its original energy delivery without any infrastructure alteration.

B. Blended Gas Real-Time Transient Model (RTTM)

The blended gas RTTM developed in this study is a tool that simulates transient events, such as leak scenarios, for blended hydrogen and natural gas pipelines. While this directly addresses the lack of data for this type of mixed gas, it also serves as a reliable data generation tool for the AI-based leak detection system developed in the study.

The RTTM is developed using the continuity and momentum partial differential equations, which are solved using the Method of Characteristics (MoC). Leaks are modeled as orifices, with the leak discharge rate being determined by a combination of the leak size and the internal pressure gradient. Different blends of hydrogen and methane are simulated with a combination of a wide-range equation of state (GERG 2008) and a thermo-physical property database (CoolProp).

Recognizing that no model can perfectly replicate the real world, several features are developed to replicate real-world conditions as much as possible. First, the inlet and outlet boundary conditions are implemented as a proportional integral derivative (PID) response, which simulates the behaviour of compressors with variable-frequency drives (VFDs). Second, the simulation supports the addition of randomized transients, which represent situations such as valves opening or closing and compressors operating incorrectly. Finally, the data intentionally has noise added to it to simulate typical field instrumentation noise that can exist in real-world systems.

The model is benchmarked against the software Pipe Flow Expert with ten test cases,

contemplating leak, and no leak cases for different mixtures. The relative errors were smaller than 1% for the cases analyzed, indicating a good correlation between both.

Finally, a case study virtual pipeline inspired by the Alberta Products Pipeline (APPL) is developed. This virtual pipeline spans approximately 300 km and is defined geographically using Google Earth^{TM} software. The pipeline was defined in eight segments, each designating the distance between two compressor stations. Then, a specialized algorithm was developed to interpret geographical coordinates, interpolate when necessary, and translate those into a homogeneous mesh to respect the CFL condition for the MoC and allow for RTTM simulations within this virtual pipeline.

C. AI-Based Blended Hydrogen Leak Detection System (LDS)

Another scientific contribution is the development of an AI-based Leak Detection System (LDS) that, while potentially applicable to any pipeline system, is uniquely designed to cater to blended hydrogen and natural gas properties. The blend's distinct characteristics pose specific challenges and requirements.

One of the fundamental strengths of the LDS lies in its transferability, stemming from its data-driven nature. The system relies on simulated data for training, which allows it to overcome a significant barrier in the field: the lack of real-world blended hydrogen data due to the emerging nature of the technology. This simulated-data approach ensures that the LDS can be efficiently applied and optimized across different pipeline systems and scenarios.

Different than other studies of leak detection systems, this system relies exclusively on commonly available field instrumentations - pressure and flow rate sensors. This distinction is important because installing new sensors to old pipelines is costly when accounting for the digging

costs and, thus, impractical in many cases. This focus on practicality and applicability ensures that the system seamlessly integrates into existing pipeline infrastructures, making the conversion process more cost-effective and manageable.

Overall, the LDS is composed of two subsystems. The first one is leak detection (LD), which is comprised of a conventional AI model (CNN) and explainable AI (ANFIS), with their classifications being fused using an ensemble learning technique (Model Stacking). The second subsystem comprises leak localization (LL) through the negative pressure wave (NPW) method.

The model fusion step allows the conventional AI to provide the core prediction and estimation capabilities. At the same time, the explainable AI component ensures transparency, directly addressing the "black box" problem that plagues similar systems and offering users a clear understanding of how the LDS is making its predictions. It is important to note that the explainability facet primarily pertains to the leak detection phase. Once a leak is confirmed, the focus shifts to quantitative measurements of leak localization, reducing the need for explainability.

The LDS was extensively tested in the case study based on the 300-kilometer Alberta Products Pipeline (APPL), defined between Edmonton and Calgary. In total, 41760 simulations were conducted for the eight pipeline segments. Each simulation had a randomized hydrogen blend, varying between 50% hydrogen and 50% natural gas and 100% hydrogen and 0% natural gas. Moreover, a wide range of parameters was varied, such as inlet and outlet parameters, leak sizes and locations, presence of random transients, and noise addition, to pose a realistic representation of potential operational conditions.

Overall, the LDS showed promising results. For leak detection (LD), the recall rate was high across all employed models. The conventional AI classifier (CNN) had a recall rate of 99.14%, the explainable AI classifier (ANFIS) had a recall of 99.16%, and finally, the model fusion

(stacking) had a recall of 99.28%. Maximizing the recall is extremely important in any leak detection system, as false negatives can represent unnoticed leaks.

Additionally, the system showcased robustness in handling data transients and synthetic noise, factors that commonly introduce variability and uncertainty in real-world applications. The models also demonstrated adaptability across pipeline segments with varying altitudes, lengths, and meshes, highlighting the system's potential for broad applicability. Nevertheless, it is worth noting that the ANFIS model did exhibit a noticeable decrease in performance for detecting leaks smaller than 200 mm^2, an area where both the CNN and the stacking models remained consistently effective.

Regarding the leak localization capabilities, the use of the Negative Pressure Wave (NPW) technique demonstrated strong performance during tests without transient disturbances, yielding a Mean Absolute Error (MAE) of only 83 meters and a Mean Absolute Percentage Error (MAPE) of 1.4%. Nevertheless, introducing transient events in the test settings led to a marked degradation in its performance. Specifically, the MAE spiked to 7503 meters, and the MAPE escalated to 59.4%. These considerable shifts underscore that, while the NPW method performs well in stable conditions, its reliability wanes when facing the complexities introduced by transient events.

This combination of transferability, practicality, comprehensive functionality, and transparency positions the LDS developed in this study as a significant scientific advancement. It enhances the safety and efficiency of hydrogen transportation and stands on its own as an AI-based solution for leak detection management.

6.3 Assumptions and Limitations

The study operates under several assumptions and limitations. It is crucial to acknowledge these factors when interpreting the results and conclusions. These considerations are as follows:

A. Feasibility of Blended Hydrogen Transportation

The study presumed that existing pipelines operate with natural gas at their Maximum Allowable Operating Pressure (MAOP). This assumption led to the conclusion of an anticipated reduction in energy transmission when converting these pipelines to carry hydrogen blends.

However, not all pipelines may be utilized for their MAOP in the real world. In such instances, the energy transmission losses upon transition to hydrogen blends may be more negligible or potentially nonexistent.

B. Blended Gas Real-Time Transient Model (RTTM) Assumptions

Steps have been taken during the development of the RTTM to the realistic conditions of blended hydrogen transportation. However, it is important to recognize that no model perfectly represents the real world. Assumptions are necessary to solve the PDEs that represent transient flow but represent a trade-off between simulation complexity and realism.

The first decision is to assume that flow is one-dimensional, with constant fluid properties throughout the pipeline cross-sectional area. This is a common assumption for commercial models and is assumed to be reasonable. Furthermore, any blended hydrogen and natural gas mixture is assumed to be perfectly mixed. The literature review revealed that, for turbulent flow conditions, hydrogen, and natural gas do not stratify in the pipeline.

Additionally, the model assumes isothermal flow conditions for computational efficiency. While modeling temperature variations would offer a more precise representation of a real-world system, the added computational complexity and increased intricacy in the equations of state may not yield a correspondingly significant improvement in accuracy. This assumption is further

justified by considering that the pipelines are insulated and buried, minimizing the influence of thermal fluctuations on the system's behaviour. Finally, the effects of the Joule-Thomson effect were already studied in the feasibility study.

Despite these assumptions and limitations, the model includes multiple features specifically developed to increase the representation of real-world systems, such as random transient event modeling, the addition of artificial noise in the signal, and the proportional integral derivative (PID) response from the compressors at the boundary locations. However, direct data from actual gas pipeline operations would offer a much more accurate representation of gas flow behavior.

C. AI-Based Leak Detection System Assumptions and Limitations

Despite its novelty and effectiveness, the Leak Detection System (LDS) developed in this study does face some limitations, primarily rooted in the source of its training data. The system was primarily trained on simulated data. This could lead to deviations when applied to actual gas pipeline conditions.

Furthermore, the system underwent testing at a sampling rate that complies with the Courant-Friedrichs-Lewy (CFL) condition. The rate varied between 0.5 to 2 Hertz, depending on the length of the pipeline segment under consideration. Some pipeline systems may lack the field instrumentation to acquire data at these specified rates. In such cases, re-designing the sampling window interval would be necessary to ensure accurate results.

The performance of the leak location component was notably robust in scenarios devoid of transient events. However, its efficacy declined significantly in the presence of such transients, indicating a clear avenue for future improvement.

Regardless of those limitations, the architecture of the LDS is designed for adaptability and continuous improvement. The system is built to accommodate fine-tuning and learning from new data, making it capable of evolving and improving its performance as more real-world data becomes available. This inherent adaptability positions the LDS as a valuable tool that can mature and become even more reliable with further research and application.

6.4 Future Work

The findings and conclusions of this study provide a stepping stone for future research in advancing the transition to a hydrogen economy. One of the critical areas to focus on is the initiation of real-world pipeline experiments. While informative, the real-time transient model (RTTM) developed in this study may not fully capture some of the complex natures of the real-world pipeline infrastructure. Experiments involving pipelines transporting blended hydrogen will offer a more precise insight into the variables at play, including the effects of pipeline dimensions, different operational conditions, and variable flow rates. Future studies could benefit significantly from collaborations with industry partners with access to such data and operational scenarios, enhancing the translatability of the research findings to the field.

An economic feasibility study on implementing blended hydrogen in pipeline networks is also valuable. This study would assess the costs of any necessary pipeline modifications or upgrades. By modeling potential savings from reduced GHG emissions, improved energy transmission, and the potential revenue from hydrogen compared to natural gas, we can gain a clearer picture of the financial implications of transitioning to a hydrogen economy.

There is also an opportunity to explore the benefits and challenges of using ammonia as a hydrogen carrier. A comparative study between the direct use of hydrogen and ammonia as a carrier could offer valuable insights and reveal more efficient hydrogen transport and utilization

methods.

Moreover, the leak localization aspect of the leak detection system can be further refined. The current utilization of the negative pressure wave method for leak localization proves effective during regular operations but encounters challenges distinguishing between leak-induced pressure drops and transient event pressure drops. A potential remedy is employing different filtering techniques that could help in isolating the characteristic signature of leaks from transient events, allowing the algorithm to better discern the time of arrival. Another avenue worth exploring is integrating machine learning algorithms trained on transient event data.

Lastly, while this study focused primarily on the technical aspects of a hydrogen economy, socio-economic factors also play a crucial role. Future work could involve studies on public acceptance, regulatory frameworks, economic implications, and job creation related to hydrogen-based energy solutions. Such socio-economic analyses can provide a more comprehensive view of the transition towards a hydrogen economy, enabling us to navigate this transition more effectively.

References

1. Kamarudin, S. K., Daud, W. R. W., Yaakub, Z., Misron, Z., Anuar, W., & Yusuf, N. N. A. N. (2009). Synthesis and optimization of future hydrogen energy infrastructure planning in Peninsular Malaysia. *International Journal of Hydrogen Energy*, *34*(5), 2077–2088. https://doi.org/10.1016/j.ijhydene.2008.12.086

2. Mazloomi, K., & Gomes, C. (2012). Hydrogen as an energy carrier: Prospects and challenges. Renewable and Sustainable Energy Reviews, 16(5), 3024–3033. https://doi.org/10.1016/j.rser.2012.02.028

3. Götz, M., Lefebvre, J., Mörs, F., McDaniel Koch, A., Graf, F., Bajohr, S., Reimert, R., & Kolb, T. (2016). Renewable Power-to-Gas: A technological and economic review. Renewable Energy, 85, 1371–1390. https://doi.org/10.1016/j.renene.2015.07.066

4. Oil & Gas Authority. (2018). Gas to Wire report: UK SNS and EIS. https://www.ogauthority.co.uk/news-publications/publications/2018/gas-to-wire-report-uk-sns-and-eis/

5. Bloomberg New Energy Finance (BNEF). (2021). New Energy Outlook 2021. https://about.bnef.com/new-energy-outlook/

6. Canadian Energy Research Institute (CERI). (2021). Canadian Natural Gas Market Supply and Demand Pathways of Change. CERI, Calgary, AB.

7. Canada, N. R. (2016, July 22). Pipelines Across Canada. Natural Resources Canada. https://www.nrcan.gc.ca/our-natural-resources/energy-sources-distribution/clean-fossil-fuels/pipelines/pipelines-across-canada/18856

8. Pipelines. (2022, June 15). Alberta Energy Regulator. https://www.aer.ca/providing-information/by-topic/pipelines

9. Tahir, M. Mohd., Ali, M. S., Salim, M. A., Bakar, R. A., Fudhail, A. M., Hassan, M. Z., & Muhaimin, M. S. A. (2015). Performance Analysis of A Spark Ignition Engine Using Compressed Natural Gas (CNG) as Fuel. Energy Procedia, 68, 355–362. https://doi.org/10.1016/j.egypro.2015.03.266

10. Verhelst, S., Demuynck, J., Sierens, R., Scarcelli, R., Matthias, N. S., & Wallner, T. (2013). Update on the Progress of Hydrogen-Fueled Internal Combustion Engines. In Renewable Hydrogen Technologies (pp. 381–400). Elsevier. https://doi.org/10.1016/B978-0-444-56352-1.00016-7

11. Newsom, G. (2020). Implications of Increased Renewable Natural Gas on Appliance Emissions and Stability. CEC-500-2020-070.

12. Cristello, J. B., Yang, J. M., Hugo, R., Lee, Y., & Park, S. S. (2023). Feasibility analysis of blending hydrogen into natural gas networks. International Journal of Hydrogen Energy, S0360319923002951. https://doi.org/10.1016/j.ijhydene.2023.01.156

13. Hormaza Mejia, A., Brouwer, J., & Mac Kinnon, M. (2020). Hydrogen leaks at the same rate as natural gas in typical low-pressure gas infrastructure. International Journal of Hydrogen Energy, 45(15), 8810–8826. https://doi.org/10.1016/j.ijhydene.2019.12.159

14. Miao, H., Lu, L., & Huang, Z. (2011). Flammability limits of hydrogen-enriched natural gas. International Journal of Hydrogen Energy, 36(11), 6937–6947. https://doi.org/10.1016/j.ijhydene.2011.02.126

15. Lee, S.-W., Lee, H.-S., Park, Y.-J., & Cho, Y.-S. (2011). Combustion and emission characteristics of HCNG in a constant volume chamber. Journal of Mechanical Science and Technology, 25(2), 489–494. https://doi.org/10.1007/s12206-010-1231-5

16. Dong, C., Zhou, Q., Zhang, X., Zhao, Q., Xu, T., & Hui, S. (2010). Experimental study on the laminar flame speed of hydrogen/natural gas/air mixtures. Frontiers of Chemical Engineering in China, 4(4), 417–422. https://doi.org/10.1007/s11705-010-0515-8

17. American Petroleum Institute. (2002). API RP 1130 Computational Pipeline Monitoring for Liquids, 2nd edition. Washington D.C.

18. Zhao, Y., McDonell, V., & Samuelsen, S. (2019). Influence of hydrogen addition to pipeline natural gas on the combustion performance of a cooktop burner. International Journal of Hydrogen Energy, 44(23), 12239–12253. https://doi.org/10.1016/j.ijhydene.2019.03.100

19. Kong, M., Feng, S., Xia, Q., Chen, C., Pan, Z., & Gao, Z. (2021). Investigation of Mixing Behavior of Hydrogen Blended to Natural Gas in Gas Network. Sustainability, 13(8), Article 8. https://doi.org/10.3390/su13084255

20. H21. (2019). H21 Leeds City Gate Report. Westminster, UK. https://h21.green/projects/h21-leeds-city-gate/

21. Tarkowski, R. (2019). Underground hydrogen storage: Characteristics and prospects. Renewable and Sustainable Energy Reviews, 105, 86–94. https://doi.org/10.1016/j.rser.2019.01.051

22. Krieg, D. (2012). Konzept und Kosten eines Pipelinesystems zur Versorgung des deutschen Straßenverkehrs mit Wasserstoff. Forschungszentrum Jülich.

23. Barbir, F., Basile, A., & Veziroğlu, T. N. (Eds.). (2016). Compendium of hydrogen energy. Volume 3: Hydrogen energy conversion / edited by Frano Barbir, Angelo Basile and T. Nejat Veziroglu. Elsevier ; WP Woodhead Publishing.

24. Australian Gas Infrastructure Group. (2019). Renewable Hydrogen, Hydrogen Park SA. World Plumbing Conference. https://www.worldplumbing.org/wp-content/uploads/2020/06/VikramSingh_RenewableHydrogenHydrogenParkSA.pdf

25. ATCO Gas Australia. (2019). Clean Energy Innovation Hub Lessons. Arena Insights Forum. https://arena.gov.au/assets/2019/12/atco-clean-energy-innovation-hub-lessions.pdf

26. Isaac, T. (2019). HyDeploy: The UK's First Hydrogen Blending Deployment Project. Clean Energy, 3(2), 114–125. https://doi.org/10.1093/ce/zkz006

27. GRTgaz, Immeuble Bora, 6, rue Raoul-Nordling, 92277 Bois-Colombes Cedex. (2019). Technical and economic conditions for injecting hydrogen into natural gas networks—Final report June 2019. France, INIS-FR--20-0156. https://www.afgaz.fr/wp-content/uploads/Technical-economic-conditions-for-injecting-hydrogen-into-natural-gas-ne.pdf

28. MARCOGAZ. (2021). Odorisation of Natural Gas and Hydrogen Mixtures. https://www.marcogaz.org/wp-content/uploads/2021/07/ODOR-Hydrogen-and-odorisation.pdf

29. Melaina, M. W., Antonia, O., & Penev, M. (2013). Blending Hydrogen into Natural Gas Pipeline Networks: A Review of Key Issues. Renewable Energy, 131.

30. Alqaheem, Y., Alomair, A., Mari, V., & Pérez, A. (2017). Polymeric Gas-Separation Membranes for Petroleum Refining. International Journal of Polymer Science, 2017, 1–19. https://doi.org/10.1155/2017/4250927

31. Du, Z., Liu, C., Zhai, J., Guo, X., Xiong, Y., Su, W., & He, G. (2021). A Review of Hydrogen Purification Technologies for Fuel Cell Vehicles. Catalysts, 11(3), Article 3. https://doi.org/10.3390/catal11030393

32. The American Society of Mechanical Engineers. (2019). ASME B31.12 Hydrogen Piping & Pipelines. New York, NY.

33. Najjar, Y. S. H. (2013). Hydrogen safety: The road toward green technology. International Journal of Hydrogen Energy, 38(25), 10716–10728. https://doi.org/10.1016/j.ijhydene.2013.05.126

34. Schmura, E., Klingenberg, M., & Corporation, C. T. (2005). Existing Natural Gas Pipeline Materials and Associated Operational Characteristics. DE-FC36-04GO14229.

35. Laurent Bedel & Michel Junker. (2006). Natural gas pipelines for hydrogen transportation. WHEC16: 16 World Hydrogen Energy Conference, France.

36. Lee, J.A. (2016). NASA 2016-218602: Hydrogen Embrittlement. National Aeronautics and Space Administration, Hunstville, AL.

37. Adolf, J., Balzer, C. H., Balzer, C., Louis, J., Schabla, U., Fischedick, M., Arnold, K., & Pastowski, A. (2017). Shell Hydrogen Study. Energy of the future? Sustainable Mobility through Fuel Cells and H2. Shell Deutschland Oil GmbH.

38. Gillette, J. L., & Kolpa, R. L. (2008). Overview of interstate hydrogen pipeline systems. (ANL/EVS/TM/08-2). Argonne National Lab. (ANL), Argonne, IL (United States). https://doi.org/10.2172/924391

39. Rigas, F., & Sklavounos, S. (2005). Evaluation of hazards associated with hydrogen storage facilities. International Journal of Hydrogen Energy, 30, 1501–1510. https://doi.org/10.1016/j.ijhydene.2005.06.004

40. Huising, O. J. C., & Krom, A. H. M. (2021, January 15). H2 in an Existing Natural Gas Pipeline. 2020 13th International Pipeline Conference. https://doi.org/10.1115/IPC2020-9205

41. HySafe. (2009). D113: Initial Guidance for Using Hydrogen in Confined Spaces—Results from InsHyde. Sixth Framework Programme, Contract No SES6-CT-2004-502630. http://www.hysafe.net/documents?deliverable=113

42. Wahl, J., & Kallo, J. (2020). Quantitative valuation of hydrogen blending in European gas grids and its impact on the combustion process of large-bore gas engines. International Journal of Hydrogen Energy, 45(56), 32534–32546. https://doi.org/10.1016/j.ijhydene.2020.08.184

43. Leader, W., Partners, W., & Huld, T. (2001). EIHP2: Compilation of Existing Safety Data on Hydrogen and Comparative Fuels. ENK6-CT2000-00442.

44. Henrie, M., Carpenter, P., & Nicholas, R. E. (2016). Pipeline Leak Detection Handbook. In Pipeline Leak Detection Handbook. Elsevier. https://doi.org/10.1016/B978-0-12-802240-5.00001-7

45. Hübert, T., Boon-Brett, L., Black, G., & Banach, U. (2011). Hydrogen sensors – A review. Sensors and Actuators B: Chemical, 157(2), 329–352. https://doi.org/10.1016/j.snb.2011.04.070

46. Soundarrajan, P., & Schweighardt, F. (2008). Hydrogen Sensing and Detection. In R. Gupta (Ed.), Hydrogen Fuel (pp. 495–534). CRC Press. https://doi.org/10.1201/9781420045772.ch15

47. Geiger, G., Vogt, D., & Tetzner, R. (2006). State-of-the-Art in Leak Detection and Localisation. Oil Gas European Magazine.

48. Kida, T., Kuroiwa, T., Yuasa, M., Shimanoe, K., & Yamazoe, N. (2008). Study on the response and recovery properties of semiconductor gas sensors using a high-speed gas-switching system. Sensors and Actuators B: Chemical, 134(2), 928–933. https://doi.org/10.1016/j.snb.2008.06.044

49. Tiratsoo, J. N. H. (Ed.). (1992). Pipeline pigging technology (2nd ed). Gulf Pub. Co.

50. Tylman, W., Kolczyński, J., & Anders, G. J. (2010). Fully automatic AI-based leak detection system. Energy, 35(9), 3838–3848. https://doi.org/10.1016/j.energy.2010.05.038

51. Kopbayev, A., Khan, F., Yang, M., & Halim, S. Z. (2022). Gas leakage detection using spatial and temporal neural network model. Process Safety and Environmental Protection, 160, 968–975. https://doi.org/10.1016/j.psep.2022.03.002

52. Wang, T.-K., Lin, Y.-H., & Shen, J.-Y. (2022). Developing and Implementing an AI-Based Leak Detection System in a Long-Distance Gas Pipeline. Advances in Technology Innovation, 7(3), Article 3. https://doi.org/10.46604/aiti.2022.8904

53. Menon, E. S. (2005). Gas Pipeline Hydraulics (1st ed.). CRC Press. https://doi.org/10.1201/9781420038224

54. Mohitpour, M., Golshan, H., & Murray, A. (2000). Pipeline Design & Construction: A Practical Approach. American Society of Mechanical Engineers. https://doi.org/10.1115/1.802574

55. Demirel, Y. (2014). Nonequilibrium thermodynamics: Transport and rate processes in physical, chemical and biological systems (Third edition). Boston ; Elsevier.

56. Haeseldonckx, D., & Dhaeseleer, W. (2007). The use of the natural-gas pipeline infrastructure for hydrogen transport in a changing market structure. International Journal of Hydrogen Energy, 32(10–11), 1381–1386. https://doi.org/10.1016/j.ijhydene.2006.10.018

57. The American Society of Mechanical Engineers. (2020). ASME B31.8 Gas Transmission & Distribution Piping Systems—ASME. New York, NY.

58. Menon, E. S. (2015). Transmission Pipeline Calculations and Simulations Manual (1st edition). Gulf Professional Publishing. https://doi.org/10.1016/C2009-0-60912-0

59. Chaudhry, M. H. (2014). Applied Hydraulic Transients. Springer New York. https://doi.org/10.1007/978-1-4614-8538-4

60. Braga, A. S. (2018). Leakage Modeling Through Empirical Equations: An Experimental Approach: (021). WDSA / CCWI Joint Conference Proceedings, 1. https://ojs.library.queensu.ca/index.php/wdsa-ccw/article/view/12012

61. Cassa, A., Van Zyl, J., & Laubscher, R. (2010). A numerical investigation into the effect of pressure on holes and cracks in water supply pipes. Urban Water Journal - URBAN WATER J, 7, 109–120. https://doi.org/10.1080/15730620903447613

62. Metwally, M., & Modisette, D. J. (2022). The Atmos book of pipeline simulation (A. Hoffman & D. J. Zhang, Eds.). Atmos International Limited.

63. Stephenson, G., & Stephenson, G. (1985). Partial differential equations for scientists and engineers (3rd ed). Longman.

64. Coughanowr, D. R. (1991). Process systems analysis and control (2. ed). McGraw-Hill.

65. Kunz, O., & Wagner, W. (2012). The GERG-2008 Wide-Range Equation of State for Natural Gases and Other Mixtures: An Expansion of GERG-2004. Journal of Chemical & Engineering Data, 57(11), 3032–3091. https://doi.org/10.1021/je300655b

66. Bell, I. H., Wronski, J., Quoilin, S., & Lemort, V. (2014). Pure and Pseudo-pure Fluid Thermophysical Property Evaluation and the Open-Source Thermophysical Property Library CoolProp. Industrial & Engineering Chemistry Research, 53(6), 2498–2508. https://doi.org/10.1021/ie4033999

67. Wilke, C. R. (2004). A Viscosity Equation for Gas Mixtures. The Journal of Chemical Physics, 18(4), 517–519. https://doi.org/10.1063/1.1747673

68. Hayes, M. H. (1996). Statistical digital signal processing and modeling. John Wiley & Sons.

69. Pipe Flow Software ® Official—Pipe Flow & Pressure Drop Calculator Software. (n.d.). Retrieved August 25, 2023, from https://www.pipeflow.com/

70. Our Pipelines | Trans-Northern Pipelines Inc. (n.d.). Retrieved August 24, 2023, from https://tnpi.ca/our-pipelines/

71. Google Earth. (n.d.). Retrieved June 11, 2022, from https://earth.google.com/web/@0,0,0a,22251752.77375655d,35y,0h,0t,0r

72. Google Maps Platform Documentation. (n.d.). Google Developers. Retrieved June 11, 2022, from https://developers.google.com/maps/documentation

73. Hou, Q., Ren, L., Jiao, W., Zou, P., & Song, G. (2013). An Improved Negative Pressure Wave Method for Natural Gas Pipeline Leak Location Using FBG Based Strain Sensor and Wavelet Transform. Mathematical Problems in Engineering, 2013, e278794. https://doi.org/10.1155/2013/278794

74. Mallat, S. G. (1999). A wavelet tour of signal processing (2nd ed). Academic Press.

75. Gwinn, R., Matties, M., & Rubin, A. D. (2021). Wavelet Selection and Employment for Side-Channel Disassembly. 2021 IEEE Physical Assurance and Inspection of Electronics (PAINE), 1–8. https://doi.org/10.1109/PAINE54418.2021.9707707

76. Calin, O. (2020). Deep Learning Architectures: A Mathematical Approach. Springer International Publishing. https://doi.org/10.1007/978-3-030-36721-3

77. Kingma, D. P., & Lei, J. (2015). Adam: A Method for Stochastic Optimization. ICLR 2015.

78. Brownlee, J. (2018). Train Faster, Reduce Overfitting, and Make Better Predictions. Machine Learning Mastery.

79. Jang, J.-S. R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics, 23(3), 665–685. https://doi.org/10.1109/21.256541

80. Gaussian combination membership function—MATLAB gauss2mf. (n.d.). Retrieved July 24, 2023, from https://www.mathworks.com/help/fuzzy/gauss2mf.html

81. Boehmke, B., & Greenwell, B. M. (2019). Hands-on machine learning with R. CRC Press.

82. Rossum, G. van. (1995). Python tutorial (CS-R9526). Centrum voor Wiskunde en Informatica (CWI).

83. Inc, T. M. (2022). MATLAB version: 9.13.0 (R2022b). The MathWorks Inc. https://www.mathworks.com

84. Chollet, F. & others. (2015). Keras. https://keras.io

85. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., & Lerer, A. (2017). Automatic differentiation in PyTorch.

86. Japkowicz, N., & Shah, M. (2014). Evaluating learning algorithms: A classification perspective (First paperback ed). Cambridge University Press.

87. Jang, J.-S. R., Sun, C.-T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing: A computational approach to learning and machine intelligence. Prentice Hall.

88. Géron, A. (2022). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, Inc.

Appendix

Appendix A: Validation Data Comparison – RTTM and Pipe Flow

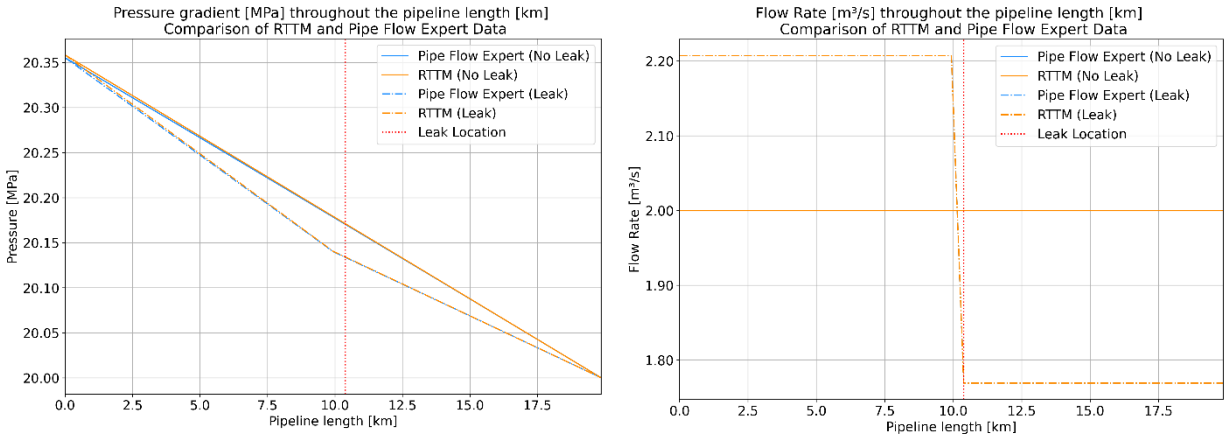


Figure A.1: Results for 100% Hydrogen 0% Methane for RTTM and Pipe Flow.

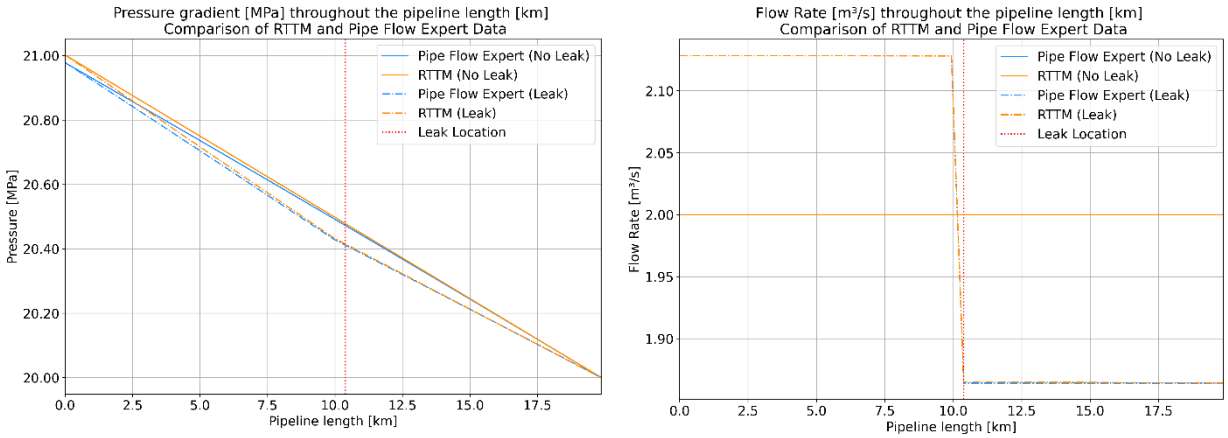


Figure A.2: Results for 75% Hydrogen 25% Methane for RTTM and Pipe Flow.

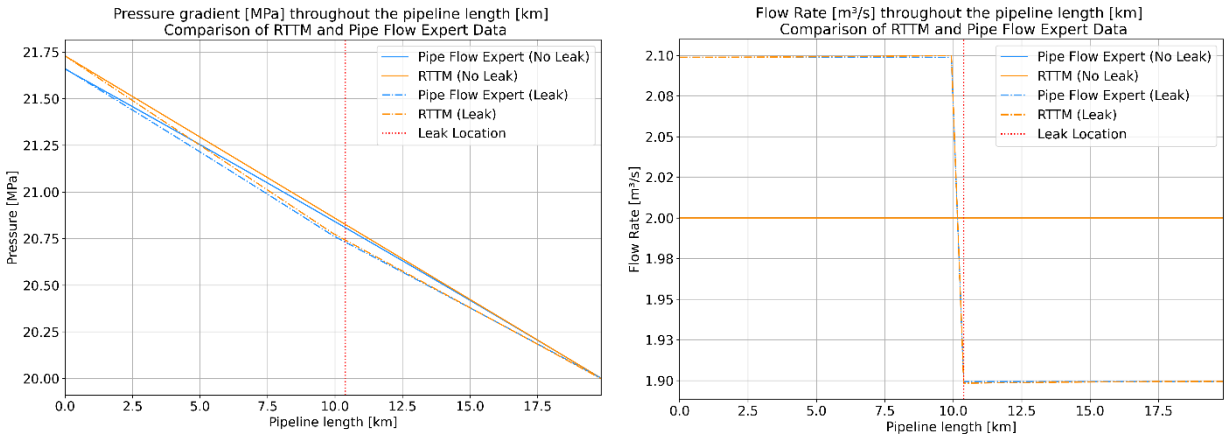


Figure A.3: Results for 50% Hydrogen 50% Methane for RTTM and Pipe Flow.

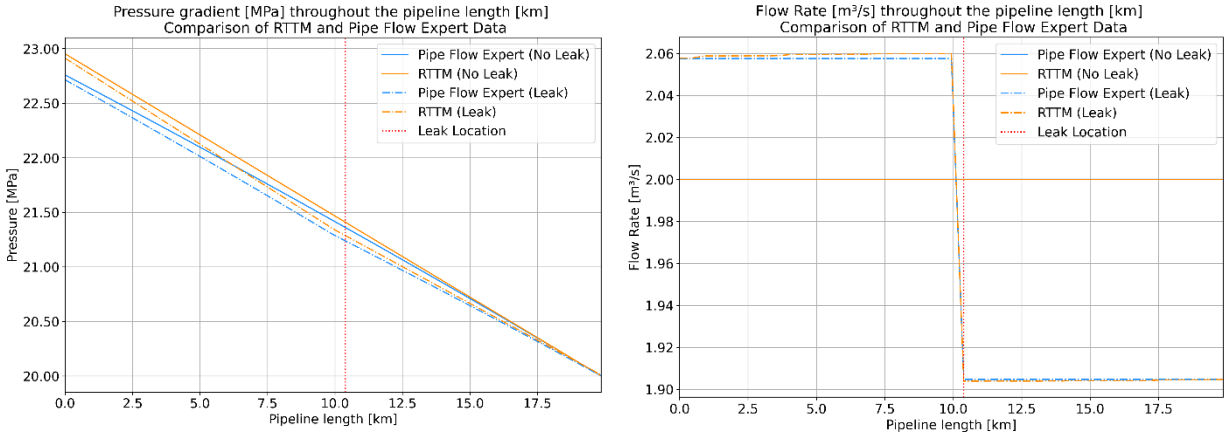


Figure A.4: Results for 25% Hydrogen 75% Methane for RTTM and Pipe Flow.

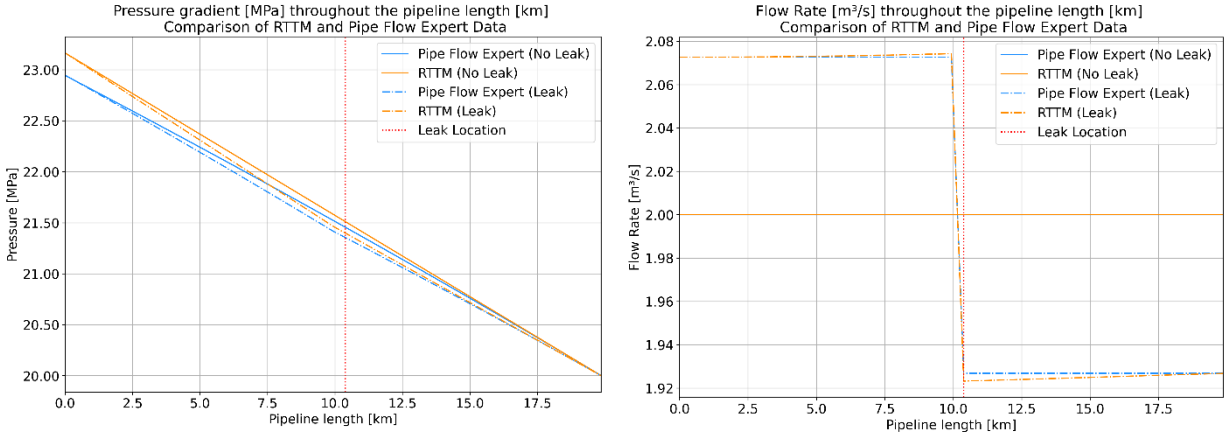


Figure A.5: Results for 0% Hydrogen 100% Methane for RTTM and Pipe Flow.

Appendix B: Optimal Blends (Hydrogen – Methane) Concentration and energy content versus System Modification.

Energy Content	Gas Contents (H_2 – CH_4)	Hydrogen Concentration without Modification (%)							
		Pressure (MPa)		Compression Ratio		Compressor Head (m)		No. of Compressors required	Modification
		Limit	Simulation Results	Limit	Simulation Results	Limit	Simulation Results		
100%	0 – 100%	11.23	12.27	1.5	2.20	2438	1915.95	2	Pipe, Compressor
	20 – 80%		13.08		2.35		2807.14	2	Pipe, Compressor
	40 – 60%		13.93		2.50		4171.38	2	Pipe, Compressor
	60 – 40%		14.80		2.65		6511.26	3	Pipe, Compressor
	80 – 20%		15.51		2.78		11372.60	5	Pipe, Compressor
	100 – 0%		14.83		2.66		26730.60	11	Pipe, Compressor
86%	0 – 100%	11.23	10.92	1.5	1.96	2438	1617.67	1	Compressor
	20 – 80%		11.59		2.08		2384.00	1	Pipe, Compressor
	40 – 60%		12.30		2.21		3560.97	2	Pipe, Compressor
	60 – 40%		13.04		2.34		5582.66	3	Pipe, Compressor
	80 – 20%		13.63		2.44		9775.49	5	Pipe, Compressor
	100 – 0%		13.06		2.34		22846.72	10	Pipe, Compressor
65%	0 – 100%	11.23	9.01	1.5	1.62	2438	1140.21	2	Compressor
	20 – 80%		9.48		1.71		1700.39	2	Compressor
	40 – 60%		9.98		1.79		2567.37	2	Compressor
	60 – 40%		10.50		1.89		4062.70	2	Compressor
	80 – 20%		10.92		1.96		7155.38	3	Compressor
	100 – 0%		10.51		1.89		16540.60	7	Compressor
51%	0 – 100%	11.23	8.42	1.5	1.42	2438	810.52	1	-
	20 – 80%		8.82		1.48		1221.56	1	-
	40 – 60%		9.25		1.54		1862.81	2	Compressor
	60 – 40%		9.69		1.61		2974.17	2	Compressor
	80 – 20%		10.06		1.66		5269.17	3	Compressor
	100 – 0%		9.709		1.61		12064.05	5	Compressor
30%	0 – 100%	11.23	6.43	1.5	1.16	2438	344.12	1	-
	20 – 80%		6.57		1.19		528.94	1	-
	40 – 60%		6.73		1.22		822.41	1	-
	60 – 40%		6.90		1.25		1337.38	1	-
	80 – 20%		7.04		1.27		2400.37	1	-
	100 – 0%		6.90		1.25		5395.90	3	Compressor