

Optimization and Prediction in Natural Gas Networks Using Graph Neural Networks and MPCC-Based Models

Cristian Alejandro Blanco Martínez

Trabajo de grado presentado como requisito
parcial para optar al título de
Magíster en Ingeniería Eléctrica

Director
Ph.D David Augusto Cárdenas Peña

UNIVERSIDAD TECNOLÓGICA DE PEREIRA
Facultad de Ingenierías
Programa de Ingeniería Eléctrica
Pereira, August 27, 2025



Nota de Aceptación

Ph.D David Augusto Cárdenas Peña. Director

Jurado

Pereira, August 27, 2025

Optimization and Prediction in Natural Gas Networks Using Graph Neural Networks and
MPCC-Based Models
©Cristian Alejandro Blanco Martínez

Pereira, August 27, 2025
Programa de Ingeniería Eléctrica.
Universidad Tecnológica de Pereira
La Julita. Pereira(Colombia)
TEL: (+57)(6)3137122
www.utp.edu.co
Versión web disponible en: *http://recursosbiblioteca.utp.edu.co*

Acknowledgments

This work was carried out with the support of several institutions and funding sources to which I am deeply grateful.

I would like to express my sincere appreciation to the research projects 1110-852-69982, funded by the Ministry of Science, Technology and Innovation of Colombia (MINCIENCIAS), and E6-24-1, funded by the Universidad Tecnológica de Pereira (UTP), for providing the financial support necessary to conduct this research.

Additionally, I extend my gratitude to the graduate program of Maestría en Ingeniería Eléctrica at UTP for the academic environment and resources that made this work possible.

Abstract

This thesis explores novel approaches to modeling and optimizing natural gas network systems, focusing on integrating Graph Neural Networks (GNNs) and Mathematical Programs with Complementarity Constraints (MPCCs). The increasing complexity of natural gas networks and energy systems demands predictive models that capture detailed system behaviors while adhering to physical laws. Traditional models often need help to account for intricate network dynamics and interconnected pipelines under diverse operating conditions. This research addresses these limitations through a multi-chapter progression, with each chapter advancing the modeling framework based on the unique strengths of GNNs and MPCCs.

In the second chapter, a GNN-based model was developed to learn the system responses obtained from a linear optimization model of the natural gas network, which modeled flow conditions without incorporating pressures. The GNN demonstrated the ability to approximate these responses effectively and generalize to cases not seen in the training phase, highlighting its potential for rapid, approximate solutions when computational efficiency is a priority.

The third chapter introduces an MPCC-based optimization model tailored for natural gas systems. It extends the modeling framework to include the Weymouth equation, which governs pressure-flow relationships in interconnected networks. This MPCC model provides a highly accurate solution by embedding non-linear pressure constraints directly into the optimization process, improving accuracy compared to traditional approaches.

Building on these advancements, the fourth chapter integrates the strengths of both approaches by using the MPCC-based model to generate accurate training data for a new, enhanced GNN-based model that incorporates pressure considerations. This hybrid model benefits from the robust physical fidelity of the MPCC-based approach, enabling the GNN to learn pressure-related responses effectively. As a result, this GNN-based model can generate predictions for scenarios not previously encountered in

training, a feature it shares with the initial GNN model from the second chapter but now with greater accuracy due to the inclusion of pressure constraints.

The results demonstrate that, while the GNN-based model may offer slightly lower accuracy than the MPCC model, it achieves predictions with a significant reduction in computational time, making it valuable for applications requiring rapid response. The MPCC-based optimization model, in contrast, provides the lowest error response, with superior accuracy in modeling non-linear pressure dynamics.

This thesis establishes that combining MPCC and GNN-based modeling, particularly with physics-informed loss functions, offers a scalable and computationally efficient framework for optimizing natural gas networks. Future research could extend this approach to incorporate transient dynamics, implement high-complexity Weymouth loss functions, and adopt a fully physics-informed neural network (PINN) approach, advancing predictive capabilities for resilient energy system operations under variable conditions.

Contents

1	Introduction	1
1.1	Justification	1
1.2	Problem statement	3
1.2.1	Objectives	5
1.2.2	General Objective	5
1.2.3	Specific Objectives	5
2	Natural Gas System Prediction Using Graph Neural Networks	6
2.1	Mathematical framework	7
2.1.1	Graph definition	7
2.1.2	Multi-Layered Perceptrons	8
2.1.3	Graph Neural Networks	10
2.1.4	Convolution with Edge-Node Switching (Cens Block)	13
2.1.5	Gas flow estimation through CensNet	16
2.1.6	Task-Dependent Loss Functions	19
2.2	Experimental Setup and Results	19
2.3	Results and Discussion	22
2.3.1	Prediction Accuracy	23
2.3.2	Gas Balance Consistency	25
2.4	Discussion	27
3	Optimization Using Mathematical Programs with Complementarity Constraints	29
3.1	Formulation of Interconnected Power and Gas Systems	29
3.2	Mathematical Programming with Complementarity Constraints for Weymouth Approximation	33
3.3	Case studies	35

CONTENTS

3.3.1	Case Study I: 9/8 System	36
3.3.2	Case II: 118/48 System	37
3.3.3	Case Study III: 96/63 System	40
3.4	Discussion	43
4	Stochastic Modeling of Natural Gas Flows Using Physics-Guided Neural Networks	45
4.1	Methodology	45
4.2	Results	47
4.2.1	Case Study I: 8-node Network	48
4.2.2	Case Study II: 63-node Network (Colombia)	66
4.3	Discussion	71
5	Conclusions and future work	73
5.1	Conclusions	73
5.2	Future Work	75

List of Figures

2.1	Types of graphs	7
2.2	General diagram of a multilayer perceptron, showing the input layers in green, the hidden layers in blue and the outputs in red.	9
2.3	General outline of the CensNet-based model used.	21
2.4	Model results using only the loss associated with nodal flow predictions in the 8-node network.	23
2.5	Model results using only the loss associated with nodal flow predictions in the colombian 63-node network.	24
2.6	Model results using the losses associated with the flows in nodes and edges predictions in the 8-node network.	24
2.7	Model results using the losses associated with the predicted flows injected at the nodes and transported by the pipelines and compressors in the colombian 63-node network.	25
3.1	Integrated system 9/8 used in Case Study I, modified from the MATPOWER-NATURAL GAS (MPNG) software.	37
3.2	Cost function histogram for the Taylor, SOC, and MPCC Weymouth approximation approaches in the 9-bus 8-node system.	38
3.3	Boxplot of Weymouth error distribution for each pipeline in the 8/9 system attained by contrasted approximation approaches.	38
3.4	Histogram depicting the relative frequencies of cost differences obtained between MPCC and the other approaches in the 48-node 118-bus system.	39
3.5	Weymouth approximation errors for each pipeline p reached by the contrasted approaches in the 118/48 study case.	39
3.6	Daily operating cost obtained with each of the approaches in the 63-node 96-bus system.	40

LIST OF FIGURES

3.7	Weymouth error density on the Colombian case versus the gas flow and operation day.	41
3.8	Output-to-input pressure ratio at the compressor in the 96/63 system.	42
3.9	Outlying connection of well-compressor-pipeline on the system 96/63 used in Case Study III.	42
4.1	General outline of the CensNet-based model with physics guidance.	47
4.2	Model results using the losses associated with nodal and edge flow predictions in the 8-node network.	48
4.3	Model results using the losses associated with nodal and edge flow predictions along with the gas balance loss in the 8-node network.	49
4.4	Model results using the losses associated with nodal and edge flow predictions along with the gas balance and Weymouth losses in the 8-node network.	50
4.5	Model results using the loss associated with nodal flow predictions along with the gas balance and Weymouth losses in the 8-node network.	51
4.6	Bimodal joint PDFs identified in the training inputs of the 8-node gas network.	54
4.7	Bimodal joint PDFs from the KDE-sampled input data propagated through the CensNet model for the 8-node network.	55
4.8	Joint PDFs between the inputs used in the 8-node network with a high probability zone concentrated in a small region.	56
4.9	Joint PDFs between the inputs used in the 8-node network with a high probability zone concentrated in a small region in the KDE sampled variables.	57
4.10	Joint PDFs between the inputs used in the 8-node network, which have a considerably wide mode with respect to the entire distribution.	58
4.11	Joint PDFs between the outputs used in the 8-node network, which have a negative flow.	59
4.12	Joint PDFs between the outputs used in the 8-node network, which have a negative flow in the KDE sampled variables.	60
4.13	Joint PDFs between the outputs used in the 8-node network have a behavior quite similar to that of a bivariate gamma distribution.	61
4.14	Joint probability density functions (PDFs) between nodal pressure outputs in the 8-node network, highlighting a concentration of values within mid-range operating levels (30%–40% of normalized limits)	62

LIST OF FIGURES

4.15 Joint PDFs between the inputs and outputs used in the 8-node network, which present a wide dispersion along second variable.	63
4.16 Joint PDFs between the inputs and outputs used in the 8-node network, which present a wide dispersion along second variable.	64
4.17 Joint PDFs between the inputs and outputs used in the 8-node net- work, which present a linear behavior, although with different disper- sions among them.	65
4.18 Model results using the losses associated with nodal and edge flows in the Colombian case network.	67
4.19 Model results using the losses associated with nodal and edge flow pre- dictions along with the gas balance and Weymouth losses in the Colombian case network.	68
4.20 Model results using the losses associated with nodal and edge flow pre- dictions along with the gas balance loss in the Colombian case network.	69

Chapter 1

Introduction

1.1 Justification

Natural gas is an energy source that has gained great relevance worldwide, due to two fundamental causes. Firstly, it has been observed that a country's economic growth is closely related to its energy consumption [1]. Therefore, as nations develop and grow economically, it is expected that they will seek energy security to meet their own demand and continue their progress without interruptions. The second major motivation for the use of natural gas is its lower greenhouse gas emissions compared to other fuels, making it an attractive choice, especially in a context where there is a growing interest in environmental care. Natural gas emits fewer greenhouse gases compared to other fossil fuels, making it a favorable option for climate change mitigation [2]. In this context, the natural gas constitutes an energy source more efficient and less polluting than coal and oil [3], that supports heating and electricity for the intensive demand of houses and industry [4].

According to the U.S. Energy Information Administration (EIA), global demand for natural gas is projected to increase steadily through 2050, driven by population growth, rising incomes, and industrial expansion in emerging regions. In most modeled scenarios, demand for natural gas rises by between 2% and 10% by 2030 and between 11% and 57% by 2050, relative to 2022 levels. Despite significant gains in renewables and efficiency, natural gas continues to play a critical role in meeting the world's growing energy needs.[5].

In Colombia, natural gas remains a cornerstone of national energy planning, with its importance reaffirmed by the *Plan de Abastecimiento de Gas Natural (PAGN)*. This

CHAPTER 1. INTRODUCTION

plan incorporates the technical recommendations made by the Unidad de Planeación Minero-Energética (UPME) in the 2019–2028 study and continues to guide infrastructure priorities. According to the updated projections from UPME (June 2021), natural gas demand is expected to grow by 17% between April 2021 and December 2035, reflecting an average annual growth rate of approximately 1.18%. This demand projection includes aggregated consumption from sectors such as residential, industrial, tertiary, transportation, petrochemical, and compression. Notably, more than 75% of the aggregated demand is concentrated in the industrial and residential sub-sectors, reinforcing the seasonal nature of consumption patterns, with increased usage in the latter months of each year aligned with economic activity trends. While demand from the oil and thermoelectric sectors has been revised downward in medium and low scenarios—due to increased energy efficiency and favorable hydrological forecasts—potential deficits in supply between 2021 and 2030 have been identified. These gaps are expected to be addressed primarily through imported and regasified LNG from facilities like SPEC in Cartagena. Natural gas remains a key energy source for Colombia, balancing reliability, lower environmental impact compared to other fossil fuels, and its strategic role in enabling a gradual transition toward cleaner energy sources [6].

Although most of the country's electricity demand is commonly met by hydroelectric plants [7], this type of generation presents a significant source of uncertainty in the energy system since its effectiveness and generation capacity are directly linked to the country's climatic and meteorological conditions, especially in extreme cases such as the El Niño phenomenon [8]. Variations in precipitation, droughts, or floods can have a significant impact on the availability of water for hydroelectric power production, affecting the balance between supply and demand in the electrical system [9]. Additionally, the increase in energy demand and the transition to renewable energy sources pose significant challenges in the efficient and reliable transportation of gas. Optimizing the natural gas transportation system, considering the uncertainty associated with renewable energy generation and demand variability, is essential to ensure a reliable, sustainable, and environmentally friendly energy supply [10].

Therefore, it is necessary for the country not only to have a national gas transportation system, but also to ensure that it is operated as efficiently as possible, in order to make the best use of available natural resources. In the Colombian context, natural gas is a very important energy source as it is used in various sectors such as residential, commercial, industrial, and thermal [11]. It is especially in the latter sector that this fuel becomes more relevant during dry seasons, as it is when reservoir levels drop and thus hydroelectric power generation decreases. This problem is exacerbated in years

when the El Niño phenomenon occurs [12], making it of great interest to have tools that allow for the optimal injection and transportation of natural gas to fully meet demand.

1.2 Problem statement

Natural gas transportation is an integral part of the natural gas industry, relying on a pipeline network to transfer natural gas from various sources to consumers, fulfilling their demand. In general, natural gas transmission systems are composed of four fundamental elements: injection fields, responsible for injecting the hydrocarbon from extraction fields or regasification plants into the system; pipelines, which transport the gas from a sending node to a receiving node; compressors, which are responsible for raising the pressure at the outlet node relative to the inlet node; and end user, which are the main consumers of natural gas [13]. Ensuring gas flow to meet end-user demand, minimizing network operating costs, and maintaining system elements within appropriate technical operating limits are critical factors in natural gas transportation. Coordinating these factors requires efficient solutions of optimization problems with a large number of variables and different nature constraints [14].

The optimization problem consists of finding the best operational configurations to meet consumer demand while ensuring the technical and physical constraints of the natural gas transportation system. It must also be considered that these transportation systems are usually interconnected with electricity systems since the latter usually require natural gas as fuel for the thermal power plants. These power plants are significant natural gas consumers, relying on a steady gas supply to generate electricity [15]. In Colombia, despite holding 70% of hydroelectric plants, the remainder consists primarily of thermoelectric plants [16]. These thermoelectric plants are key to complementing the hydro plants to meet energy demand, especially during periods of drought, e.g., during El Niño phenomenon when reduced water availability limits hydroelectric generation [17]. As other studies have shown, variations in rainfall, droughts, or floods in countries with significant hydroelectric power plants can significantly affect water availability for hydropower production [18]. This dynamic introduces uncertainty into the demand for natural gas, as the variability of renewable energy sources such as hydropower directly influences the operational requirements of thermoelectric plants [19]. Therefore, natural gas systems must respond flexibly and rapidly to changing energy demands, which further underscores the need for computationally efficient tools capable of handling frequent re-optimizations under variable conditions [20].

The above situation necessitates solving the optimization problem multiple times

to ensure the system’s correct operation across various scenarios. Consequently, this process takes considerable time and is both resource-intensive and time-consuming. Despite the high computational cost of each model execution, the resulting solutions are not utilized in subsequent optimization processes, even in similar operational scenarios. Therefore, there is a pressing need to develop a machine learning strategy that leverages historical solutions to provide faster responses to different operational scenarios by learning from past optimization outcomes. However, the effectiveness of such a strategy also depends on how accurately the underlying physical components of the system are modeled—particularly the transmission pipelines, which pose significant challenges due to their nonlinear behavior.

Although production fields, compressors, and end users of natural gas are well-represented, modeling transmission pipelines remains complex due to the nonlinear relationship between flow and pressures at its ending nodes. This complexity arises from the Weymouth equation, which includes a nonconvex and discontinuous sign function that determines flow direction based on differential pressure. These nonconvexities introduce discontinuities lead to numerical issues and optimization instability [21, 22]. Various authors have approached the challenge posed by the Weymouth equation differently. One of the approaches used to address this problem is to approximate this equation due to its inherent complexity and non-convex nature [23]. However, since it is an approximation, this solution introduces errors that impact the accuracy of optimization outcomes. Mitigating these errors remains critical for further research and development in natural gas transportation systems [13].

This research seeks to address these challenges by developing a machine learning-based strategy that leverages solutions from past optimizations to provide rapid, reliable predictions for various operational scenarios. Specifically, this thesis explores using Graph Neural Network (GNN) models and MPCC-based optimization formulations as complementary tools. The GNN model, by learning the structure and patterns from historical optimization outcomes, enables faster scenario evaluations while maintaining acceptable error margins. Furthermore, the Mathematical Programs with Complementarity Constraints (MPCC) approach introduces an accurate modeling of the Weymouth equation, enhancing the fidelity of flow and pressure calculations without compromising computational efficiency.

1.2.1 Objectives

1.2.2 General Objective

To develop an optimization tool that integrates knowledge of the gas transportation network topology, a suitable approximation of the Weymouth equation and stochastic optimization techniques to address the gas transportation task taking into account the uncertainties related to hydroelectric generation and the growth of alternative energy sources.

1.2.3 Specific Objectives

- Design a Graph Neural Networks-based approach of regression that integrates knowledge of natural gas network topology to reduce computational time for operation estimation.
- Develop an optimization model for natural gas transportation systems that takes into account the Weymouth equation for reducing that reduces the approximation error in pipeline gas flow calculations.
- Develop a stochastic gas flow dispatch optimization strategy that quantifies the uncertainty in the objective variables and decision variables associated with the operation of the gas system taking into account the constraints of the transportation problem.

Chapter 2

Natural Gas System Prediction Using Graph Neural Networks

In recent years, the increasing complexity of natural gas transportation systems has driven the need for more efficient and scalable methods to estimate their operational behavior. Traditional optimization approaches, while accurate, can become computationally expensive, particularly when applied to large-scale or real-time scenarios. This chapter examines the application of Graph Neural Networks (GNNs) as a surrogate modeling technique for forecasting the operational state of natural gas networks. By leveraging the inherent graph structure of these systems, where nodes represent physical elements such as wells or demand points, and edges represent pipelines or compressors. The GNN model is trained to approximate nodal and edge-level variables with high accuracy. Special attention is given to incorporating physical constraints into the learning objective, particularly losses at both the node and edge levels, to ensure consistency with the underlying physics. The approach is validated on two test systems of varying complexity: a small 8-node network and a 63-node model representing the Colombian gas infrastructure. The results demonstrate that this topology-aware GNN framework can significantly reduce computation times while maintaining high prediction quality.

2.1 Mathematical framework

2.1.1 Graph definition

A graph G is a mathematical structure that represents a set of interconnected objects. These objects are known as vertices (or nodes), denoted by the set V , and the connections between them are called edges (or arcs), denoted by the set E . Formally, a graph is defined as an ordered pair $G = (V, E)$, where V is a non-empty set of vertices, and $E \subseteq \{(u, v) \mid u, v \in V(G), u \neq v\}$ is the set of edges, where each edge connects two distinct vertices [24].

Graphs can be categorized based on the properties of their edges. An undirected graph has edges that do not have a direction, so the pair $(u, v) = (v, u)$ represents an edge that simply connects vertices u and v . In contrast, in a directed graph (or digraph), each edge $(u, v) \in E$ has a direction, meaning it goes from vertex u to vertex v . This implies that $(u, v) \neq (v, u)$ unless $u = v$ [25].

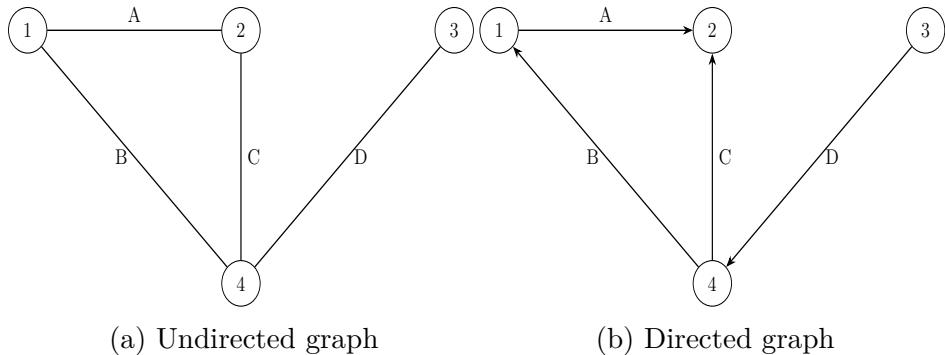


Figure 2.1: Types of graphs

In Figure 2.1, two graphs are represented, each composed of four nodes labeled 1, 2, 3, and 4, and six edges labeled A, B, C, D, E , and F . The difference between them lies in the type of graph they represent. For example, in Figure 2.1a, the edge C connects nodes 2 and 4. However, in Figure 2.1b, this connection provides additional information: a direction, which, in the context of this study, could represent the direction of a specific element, such as electric power or gas flow.

A graph can be represented in various ways using matrices, each capturing different aspects of the graph structure. The two most common matrix representations are the adjacency matrix and the incidence matrix.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

The adjacency matrix of a graph is a square matrix used to represent the connections between vertices [26]. For a graph G with n vertices, the adjacency matrix A is an $n \times n$ matrix where the entry a_{ij} is defined as follows:

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge from vertex } u \text{ to vertex } v, \\ 0 & \text{otherwise.} \end{cases} \quad (2.1)$$

For a directed graph, the adjacency matrix captures the direction of the edges. Below is the adjacency matrix for the directed graph shown earlier:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

The incidence matrix of a graph represents the relationship between vertices and edges [26]. For a graph G with n vertices and m edges, the incidence matrix I is an $n \times m$ matrix where the entry i_{ij} is defined as follows:

$$b_{ij} = \begin{cases} 1 & \text{if vertex } i \text{ is the starting point of edge } j \text{ in a directed graph,} \\ -1 & \text{if vertex } i \text{ is the endpoint of edge } j \text{ in a directed graph,} \\ 0 & \text{if vertex } i \text{ is not connected to edge } j. \end{cases} \quad (2.2)$$

For the directed graph previously described, the incidence matrix is given by:

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}$$

2.1.2 Multi-Layered Perceptrons

A Multilayer Perceptron (MLP) is a fundamental type of artificial neural network, often regarded as one of the building blocks of deep learning. At its core, an MLP consists of multiple layers of nodes, or neurons, where each layer is fully connected to the next one. The architecture typically includes an input layer, one or more hidden layers, and an output layer, as can be seen in the Figure 2.2.

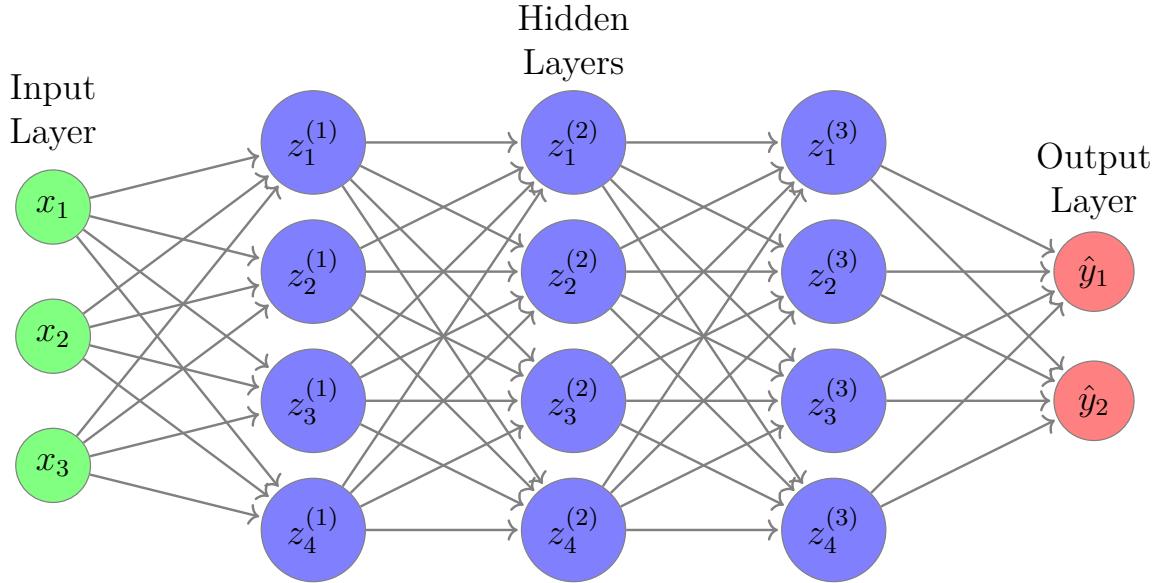


Figure 2.2: General diagram of a multilayer perceptron, showing the input layers in green, the hidden layers in blue and the outputs in red.

The neurons in each layer are connected to the neurons in the subsequent layer through weighted connections, the key parameters learned during the training process [27]. One of the most significant properties of an MLP is its ability to work as a universal approximator. Given sufficient neurons in the hidden layers, an MLP can approximate any continuous function to an arbitrary degree of accuracy, provided the network is correctly trained [28].

Mathematically, an MLP can be defined as follows. Let $\mathbf{x} \in \mathbb{R}^n$ the input vector, where n is the number of features. The output of each neuron in the first hidden layer is calculated as:

$$\mathbf{z}^{(1)} = \sigma(\mathbf{W}^{(1)}\mathbf{x} + \mathbf{b}^{(1)}) \quad (2.3)$$

where $\mathbf{W}^{(1)} \in \mathbb{R}^{m_1 \times n}$ is the weighting matrix for the first hidden layer, with m_1 being the number of neurons in this layer, $\mathbf{b}^{(1)} \in \mathbb{R}^{m_1}$ is the bias vector, and $\sigma(\cdot)$ is the activation function, typically a non-linear function such as a ReLU (Rectified Linear Unit) or sigmoid function.

This process is repeated for each subsequent hidden layer k , where the output of the k -th layer is given by:

$$\mathbf{z}^{(k)} = \sigma(\mathbf{W}^{(k)}\mathbf{z}^{(k-1)} + \mathbf{b}^{(k)}) \quad \forall i \in \{2, \dots, L-1\} \quad (2.4)$$

Here, $\mathbf{W}^{(k)} \in \mathbb{R}^{m_k \times m_{k-1}}$ represents the weight matrix connecting layer $k-1$ to layer k , $\mathbf{b}^{(k)} \in \mathbb{R}^{m_k}$ is the bias vector for layer k , and $\mathbf{z}^{(k-1)}$ is the output of the previous layer.

Finally, the output layer produces the final prediction $\hat{\mathbf{y}}$:

$$\hat{\mathbf{y}} = \sigma(\mathbf{W}^{(L)}\mathbf{z}^{(L-1)} + \mathbf{b}^{(L)}) \quad (2.5)$$

where L denotes the number of layers in the network, including the input and output layers. Depending on the nature of the problem (e.g., classification or regression), the activation function $\sigma(\cdot)$ used in the output layer can vary, with softmax being common in multi-class classification problems, and a linear activation for regression tasks. The entire MLP is trained using a process called backpropagation, combined with an optimization algorithm like gradient descent, to minimize a loss function $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}})$, which compares the true outputs \mathbf{y} and the predicted outputs $\hat{\mathbf{y}}$.

2.1.3 Graph Neural Networks

In recent years, Graph Neural Networks (GNNs) have emerged as a powerful tool in machine learning, particularly for tasks involving data that can be naturally represented as graphs. Graphs are a universal data structure that can model various systems in numerous fields, including social networks, biological networks, knowledge graphs, and physical systems. Because of their ability to represent relationships and interactions between entities, graphs are used extensively to model complex structures where the data points are not independent but interconnected [29].

GNNs are important because they can directly operate on graph-structured data, extending neural networks' success from grid-like data structures, such as images and sequences, to more general and irregular structures [30]. Traditional neural networks, like Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), are designed to work with data that has a fixed structure. However, many real-world problems involve data that can be better described by graphs, where nodes represent entities and edges represent relationships between those entities [31].

Graph Neural Networks can be broadly defined as a class of neural networks designed to perform inference on data described by graphs. Formally, let $G = (V, E)$ represent a graph, where V is the set of nodes (or vertices) and E is the set of edges. Each node $v \in V$ can be associated with a feature vector $\mathbf{x}_v \in \mathbb{R}^{D_v}$, where D_v is the number of

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

features per node, and each edge $(u, v) \in E$ may have an associated weighting or feature vector $\mathbf{e}_{uv} \in \mathbb{R}^{D_e}$, where D_e is the number of features per edge. The goal of a GNN is to learn a representation for each node (or sometimes for the entire graph) by aggregating and transforming the feature information from the node's local neighborhood in the graph.

To move from the abstract concept to a practical implementation, the specific functions used for updating and aggregating node features must be defined [32]. In the GNN framework, message passing is understood as a series of iterations in which each node updates its representation by exchanging information with its neighbors.

The basic message passing operation, which simplifies the original GNN model proposed by [33], is expressed by the following equation:

$$\mathbf{h}_u^{(k)} = \sigma \left(\mathbf{W}_{\text{self}}^{(k)} \mathbf{h}_u^{(k-1)} + \mathbf{W}_{\text{neigh}}^{(k)} \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v^{(k-1)} + \mathbf{b}^{(k)} \right) \quad (2.6)$$

In this equation:

- $\mathbf{h}_u^{(k)} \in \mathbb{R}^{D_h^{(k)}}$ represents the updated feature vector of node u at layer k , where $D_h^{(k)}$ is the dimensionality of the hidden representation at that layer.
- The term $\mathbf{W}_{\text{self}}^{(k)} \mathbf{h}_u^{(k-1)}$, where $\mathbf{W}_{\text{self}}^{(k)} \in \mathbb{R}^{D_h^{(k)} \times D_h^{(k-1)}}$ and $\mathbf{h}_u^{(k-1)} \in \mathbb{R}^{D_h^{(k-1)}}$, transforms the node's own feature vector from the previous layer, enabling the node to retain and modify its own information.
- The term $\mathbf{W}_{\text{neigh}}^{(k)} \sum_{v \in \mathcal{N}(u)} \mathbf{h}_v^{(k-1)}$, where $\mathbf{W}_{\text{neigh}}^{(k)} \in \mathbb{R}^{D_h^{(k)} \times D_h^{(k-1)}}$, aggregates the feature vectors of the neighboring nodes v in the set $\mathcal{N}(u)$, and then applies a transformation via the weighting matrix.
- $\mathbf{b}^{(k)} \in \mathbb{R}^{D_h^{(k)}}$ is a bias term that can be added to the weighted sum, though it is sometimes omitted for simplicity.
- The non-linear function $\sigma(\cdot)$ is elementwise applied to introduce non-linearity into the model, which is essential for capturing complex patterns in the data.

In the context of Graph Neural Networks (GNNs), a Graph Convolutional Network (GCN) is a specialized model that applies the concept of convolution, widely used in image processing, to graphs. First introduced by [34], GCNs offer a method to perform

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

deep learning on graph-structured data by extending traditional convolution operations to the irregular domain of graphs.

Specifically, the adjacency matrix $\tilde{\mathbf{A}}$ is defined as:

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I} \quad (2.7)$$

with \mathbf{A} is the original adjacency matrix, and \mathbf{I} is the identity matrix. The identity matrix \mathbf{I} ensures that each node considers its own features when aggregating information from its neighbors.

The degree matrix $\tilde{\mathbf{D}}$ is defined as:

$$\tilde{D}_{vv} = \sum_{v' \in V} \tilde{A}_{vv'} \quad (2.8)$$

where V represents the set of all nodes in the graph. The diagonal entries of $\tilde{\mathbf{D}}$ correspond to the degree of each node, adjusted to account for the added self-loops.

The fundamental idea behind GCNs is to create a spectral filter that operates on graph data. The filter's purpose is to combine features from a node's local neighborhood, taking into account the graph's structure. This process is mathematically formalized in the following way:

$$\mathbf{H} = \sigma \left(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{X} \Theta \right) \quad (2.9)$$

Here:

- $\mathbf{H} \in \mathbb{R}^{N_v \times D_h}$ represents the matrix of node representations after applying the GCN layer. Each row $\mathbf{h}_u \in \mathbb{R}^{D_h}$ in \mathbf{H} corresponds to the updated feature vector for node u .
- $\mathbf{X} \in \mathbb{R}^{N_v \times D_v}$ is the matrix of input node features, where each row $\mathbf{x}_u \in \mathbb{R}^{D_v}$ corresponds to the feature vector for node u before applying the GCN layer.
- $\sigma(\cdot)$ denotes a non-linear activation function elementwise applied to introduce non-linearity into the model.
- $\tilde{\mathbf{A}} \in \mathbb{R}^{N_v \times N_v}$ is the adjacency matrix of the graph, with added self-loops to account for the node itself in the aggregation.
- $\tilde{\mathbf{D}} \in \mathbb{R}^{N_v \times N_v}$ is the degree matrix of the graph, modified to include the self-loops. The degree matrix is diagonal, with each diagonal entry $\tilde{D}_{ii} \in \mathbb{R}$ representing the degree of node i in the graph.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

- $\Theta \in \mathbb{R}^{D_v \times D_h}$ is a matrix of trainable parameters, which is learned during the training process to optimize the model's performance.

The expression $\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}$ is a normalized version of the adjacency matrix, ensuring that the eigenvalues of the operation are bounded between 0 and 1. This normalization step is crucial as it prevents issues such as exploding or vanishing gradients during the training of deep networks.

2.1.4 Convolution with Edge-Node Switching (Cens Block)

Graph Convolutional Networks (GCNs) have demonstrated considerable success in various graph-based machine learning tasks, particularly in their ability to generalize convolution operations to non-Euclidean data structures like graphs [35]. GCNs operate by aggregating features from a node's neighbors, thereby capturing local neighborhood information and propagating it through the network layers. Despite their effectiveness, GCNs possess certain limitations that hinder their performance in more complex scenarios.

One notable limitation of GCNs is their reliance on node features during the convolution process, which overlooks the information encoded in edge features. This limitation prevents the model from fully capturing the complexity of the graph structure and may reduce its overall effectiveness [36]. Furthermore, GCNs typically aggregate information from immediate neighbors only, which can limit their ability to capture long-range dependencies in large or densely connected graphs. This restriction can lead to an oversimplified representation of the graph structure, particularly in cases where the graph contains intricate patterns that require deeper and more nuanced analysis [37].

To overcome the limitations of traditional GCNs, which focus primarily on node features, CensNet introduces a novel approach that integrates both node and edge features into the graph convolution process. The CensNet framework consists of two primary types of layers: the *node layer* and the *edge layer*. These layers work in concert to update node and edge embeddings alternately, leveraging the information from both nodes and edges in the graph [38].

The propagation rules in CensNet are designed to incorporate edge features into the convolution process, enabling a more comprehensive feature propagation across the graph. We define the normalized node adjacency matrix with self-loops as follows:

$$\tilde{\mathbf{A}}_v = \mathbf{D}_v^{-\frac{1}{2}} (\mathbf{A}_v + \mathbf{I}_{N_v}) \mathbf{D}_v^{-\frac{1}{2}}, \quad (2.10)$$

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

In this expression $\mathbf{A}_v \in \mathbb{R}^{N_v \times N_v}$ is the adjacency matrix for the nodes, and $\mathbf{I}_{N_v} \in \mathbb{R}^{N_v \times N_v}$ is the identity matrix that introduces self-loops. The matrix $\mathbf{D}_v \in \mathbb{R}^{N_v \times N_v}$ is the diagonal degree matrix of $\mathbf{A}_v + \mathbf{I}_{N_v}$, where each diagonal entry corresponds to the degree of a node (i.e., the number of neighbors, including itself). The term N_v denotes the number of nodes in the graph. Consequently, the sum $\mathbf{A}_v + \mathbf{I}_{N_v}$ is a matrix of size $N_v \times N_v$, and so is \mathbf{D}_v .

Node Layer Propagation: In the $(l+1)$ -th layer, the node features are updated using the following propagation rule:

$$\mathbf{H}_v^{(l+1)} = \sigma \left(\mathbf{T} \Phi \left(\mathbf{H}_e^{(l)} \mathbf{p}_e \right) \mathbf{T}^\top \odot \tilde{\mathbf{A}}_v \mathbf{H}_v^{(l)} \mathbf{W}_v \right) \quad (2.11)$$

With,

- $\mathbf{T} \in \{0, 1\}^{N_v \times N_e}$ is an incidence matrix that represents the connections between nodes and edges. Each element $T_{v,e}$ indicates whether edge e connects to node v . Specifically, if edge e is connected to node v , then $T_{v,e} = 1$; otherwise, $T_{v,e} = 0$. Given that each edge is formed by two nodes, every column of the matrix \mathbf{T} will have exactly two elements equal to 1, corresponding to the two nodes that the edge connects.
- $\mathbf{H}_e^{(l)} \in \mathbb{R}^{N_e \times D_e}$ is the edge feature matrix from the l -th layer, where D_e is the dimensionality of the edge features. $\mathbf{p}_e \in \mathbb{R}^{D_e}$ is a learnable vector of weights that projects the edge features into a scalar value per edge. The matrix-vector product $\mathbf{H}_e^{(l)} \mathbf{p}_e \in \mathbb{R}^{N_e}$ results in a vector with one scalar value per edge. The operation $\text{diag}(\mathbf{H}_e^{(l)} \mathbf{p}_e) \in \mathbb{R}^{N_e \times N_e}$ converts this vector into a diagonal matrix where each element of the vector is placed on the diagonal.
- The Hadamard product, denoted by \odot , represents element-wise multiplication between matrices. In this context, it combines the transformed edge features with the node adjacency matrix, integrating information from both the original graph and its line graph.
- $\tilde{\mathbf{A}}_v \in \mathbb{R}^{N_v \times N_v}$ is the normalized adjacency matrix for nodes, as shown in eq. (2.10), where $\mathbf{A}_v \in \mathbb{R}^{N_v \times N_v}$ is the original node adjacency matrix and $\mathbf{I}_{N_v} \in \mathbb{R}^{N_v \times N_v}$ is the identity matrix that introduces self-loops. This normalization ensures that the contributions from each node's neighbors are appropriately scaled.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

- $\mathbf{H}_v^{(l)} \in \mathbb{R}^{N_v \times D_v}$ represents the node feature matrix from the l -th layer. $\mathbf{W}_v \in \mathbb{R}^{D_v \times D'_v}$ is a learnable weight matrix that is applied to the node features during the propagation process.
- The activation function σ (typically a non-linear function such as ReLU) is applied element-wise to the resulting matrix to introduce non-linearity into the model.

This expression can be understood as a mechanism for integrating node and edge information. The matrix \mathbf{T} is responsible for transferring edge features into the node domain, allowing these edge-derived features to be merged with the normalized node adjacency matrix $\tilde{\mathbf{A}}_v$.

Edge Layer Propagation: Similarly, the normalized edge adjacency matrix is defined as:

$$\tilde{\mathbf{A}}_e = \mathbf{D}_e^{-\frac{1}{2}} (\mathbf{A}_e + \mathbf{I}_{N_e}) \mathbf{D}_e^{-\frac{1}{2}}, \quad (2.12)$$

where $\mathbf{D}_e \in \mathbb{R}^{N_e \times N_e}$ is the degree matrix corresponding to the edge adjacency matrix $\mathbf{A}_e + \mathbf{I}_{N_e}$, with $\mathbf{A}_e \in \mathbb{R}^{N_e \times N_e}$ and $\mathbf{I}_{N_e} \in \mathbb{R}^{N_e \times N_e}$. The matrix $\tilde{\mathbf{A}}_e \in \mathbb{R}^{N_e \times N_e}$ serves as the normalized version of the edge adjacency matrix, similarly to the node adjacency normalization. This normalization ensures that the influence of each edge is scaled appropriately, which is crucial for the stability of the propagation process.

The propagation rule for edge features is defined as follows:

$$\mathbf{H}_e^{(l+1)} = \sigma \left(\mathbf{T}^\top \Phi \left(\mathbf{H}_v^{(l)} \mathbf{P}_v \right) \mathbf{T} \odot \tilde{\mathbf{A}}_e \mathbf{H}_e^{(l)} \mathbf{W}_e \right), \quad (2.13)$$

In this expression, the following components are involved:

- $\mathbf{T}^\top \in \{0, 1\}^{N_e \times N_v}$ is the transpose of the binary transformation matrix $\mathbf{T} \in \{0, 1\}^{N_v \times N_e}$ used in the node layer propagation. The matrix \mathbf{T}^\top maps the node features back into the edge domain, allowing the edge features to be updated based on the node information.
- $\mathbf{H}_v^{(l)} \in \mathbb{R}^{N_v \times D_v}$ is the node feature matrix from the l -th layer, and $\mathbf{p}_v \in \mathbb{R}^{D_v}$ is a learnable vector that projects each node's features to a scalar. The operation $\mathbf{H}_v^{(l)} \mathbf{p}_v \in \mathbb{R}^{N_v}$ results in a scalar per node. The function $\Phi(\cdot)$ produces the diagonal matrix $\Phi \left(\mathbf{H}_v^{(l)} \mathbf{p}_v \right) \in \mathbb{R}^{N_v \times N_v}$, which is used to scale each node's influence in the projection.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

- The matrix product $\mathbf{T}^\top \Phi(\mathbf{H}_v^{(l)} \mathbf{p}_v) \mathbf{T} \in \mathbb{R}^{N_e \times N_e}$ maps this node-level scalar information into the edge domain. This matrix has the same dimensions as the normalized edge adjacency matrix $\tilde{\mathbf{A}}_e \in \mathbb{R}^{N_e \times N_e}$, which ensures the Hadamard product is dimensionally valid.
- $\mathbf{H}_e^{(l)} \in \mathbb{R}^{N_e \times D_e}$ represents the edge feature matrix from the l -th layer, while $\mathbf{W}_e \in \mathbb{R}^{D_e \times D'_e}$ is a learnable weight matrix. Their product $\mathbf{H}_e^{(l)} \mathbf{W}_e \in \mathbb{R}^{N_e \times D'_e}$ results in the updated edge features before aggregation, where D'_e denotes the dimensionality of the edge features in the $l+1$ -th layer.
- The Hadamard product \odot element-wise multiplies the matrix $\mathbf{T}^\top \Phi(\mathbf{H}_v^{(l)} \mathbf{p}_v) \mathbf{T} \in \mathbb{R}^{N_e \times N_e}$ with $\tilde{\mathbf{A}}_e \in \mathbb{R}^{N_e \times N_e}$, producing a filtered adjacency matrix in the edge domain.
- The resulting matrix (after Hadamard product) multiplies the transformed edge features $\mathbf{H}_e^{(l)} \mathbf{W}_e \in \mathbb{R}^{N_e \times D'_e}$, resulting in an output matrix in $\mathbb{R}^{N_e \times D'_e}$, consistent with the desired shape of $\mathbf{H}_e^{(l+1)}$.
- As in the node layer propagation, the activation function σ is applied element-wise to introduce non-linearity.

This propagation rule updates the edge embeddings by integrating information from the node features and the edge structure, thereby enhancing the expressiveness of the edge representations. The alternating updates between node and edge embeddings allow the model to effectively bridge signals across nodes and edges, resulting in more informative graph embeddings compared to other neural network approaches.

2.1.5 Gas flow estimation through CensNet

The natural gas system consists of a network of production centers, pipelines, compressor stations, storage facilities, and distribution points that ensure reliable gas delivery from producers to consumers. Mathematically, this system can be represented as a directed graph defined as $G = (V, E)$ where V is the set of units within the gas system, and E is the set of different elements linking them. This set of units includes gas supply nodes or wells $\mathcal{W} \subset V$, gas demand nodes or users $\mathcal{U} \subset V$, and gas storage facilities $\mathcal{S} \subset V$. Similarly, the set of directed gas adjacency edges $\mathcal{A} = \{(n, m) \mid n, m \in V\} \subset E$ delineates the network structure through two kinds of

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

transmission elements: transport pipelines $\mathcal{P} = \{p = (n, m) \mid n, m \in V\}$ and compressing stations $\mathcal{C} = \{c = (n, m) \mid n, m \in V\}$, so that $\mathcal{P} \cup \mathcal{C} = \mathcal{A}$ and $\mathcal{P} \cap \mathcal{C} = \emptyset$.

Natural gas transportation requires coordination to manage the flow through the different elements to maintain safe operating ranges. In optimizing this network, mathematical models minimize overall operating costs associated with the various stages of natural gas transportation, compression, storage, and handling unsupplied demand, ensuring compliance with technical and physical constraints. The cost function to minimize is expressed as:

$$\min_{\mathcal{P}, \mathcal{F}} \quad \sum_{t \in \mathcal{T}} \sum_{w \in \mathcal{W}} C_w^t f_w^t + \sum_{p \in \mathcal{P}} C_p^t f_p^t + \sum_{c \in \mathcal{C}} C_c^t f_c^t + \sum_{u \in \mathcal{U}} C_u^t f_u^t \quad (2.14)$$

Here, \mathcal{T} denotes the set of discrete time periods considered in the optimization model. The term $\sum_{w \in \mathcal{W}} C_w^t f_w^t$ represents the total cost of gas production at the wells, where C_w^t denotes the cost per unit flow of gas at a specific well w during time period t , and f_w^t corresponds to the flow of gas from well w . Similarly, the transportation of gas through pipelines is captured by the term $\sum_{p \in \mathcal{P}} C_p^t f_p^t$, where C_p^t is the cost per unit flow through pipeline p during time period t , and f_p^t represents the flow of gas through pipeline p . In addition, the total cost associated with gas compression at compressor stations is accounted for by $\sum_{c \in \mathcal{C}} C_c^t f_c^t$, where C_c^t is the cost per unit flow at compressor station c during time period t , and f_c^t is the flow of gas through compressor station c . Beyond production, transportation, and compression, the model also considers the costs related to unmet gas demand. The term $\sum_{u \in \mathcal{U}} C_u^t f_u^t$ reflects the penalty cost associated with unsupplied gas demand, where C_u^t is the penalty cost per unit of unsupplied gas at location u during time period t , and f_u^t represents the volume of unmet demand.

The constraint set models the gas architecture of the transportation system and its technical limits: Equation (2.15) forces each production well to inject the flow $f_w^t \in \mathbb{R}^{N_v}$ over the technical minimum $\underline{f}_w^t \in \mathbb{R}^{N_v}$ and under the maximum capacity $\overline{f}_w^t \in \mathbb{R}^{N_v}$. Equation (2.16) upper-bounds the gas flow through pipelines $f_p^t \in \mathbb{R}^{N_e}$ to the structural capacity $\overline{f}_p^t \in \mathbb{R}^{N_e}$. Equation (2.17) ensures that the unsupplied demand $f_u^t \in \mathbb{R}^{N_v}$ is lower than the corresponding user demand $\overline{f}_u^t \in \mathbb{R}^{N_v}$. The nodal gas balance in Equation (2.18) guarantees that the gas entering the node n equals the gas leaving it.

$$\underline{f}_w^t \leq f_w^t \leq \overline{f}_w^t \quad \forall w \in \mathcal{W} \quad (2.15)$$

$$-\overline{f}_p^t \leq f_p^t \leq \overline{f}_p^t \quad \forall p \in \mathcal{P} \quad (2.16)$$

$$0 \leq f_u^t \leq \overline{f}_u^t \quad \forall u \in \mathcal{U} \quad (2.17)$$

$$\sum_{m:(m,n) \in \mathcal{A}} f_m^t = \sum_{m':(n,m') \in \mathcal{A}} f_{m'}^t \quad \forall n \in \mathcal{N}_f \quad (2.18)$$

The optimization model described above is used to generate flow scenarios across the network. These scenarios serve as supervised training data for a CensNet-based model that approximates the mapping between the structural and technical parameters of the network and the resulting gas flows. The CensNet architecture is designed to learn these mappings by taking as input a set of graph-based features and predicting the flow behavior directly through a regression task. Specifically, the network receives the following input data:

- **Node Features $\mathbf{X} \in \mathbb{R}^{N_v \times 3}$:** A matrix where each row corresponds to a node in the graph, and each column represents a node-level attribute. Each node feature vector includes the lower and upper limits for injected flow, as well as demanded flow.
- **Node Laplacian $\mathbf{L}_v \in \mathbb{R}^{N_v \times N_v}$:** An adjacency-based Laplacian matrix encoding the graph structure of the nodes. This matrix captures the relational topology among nodes and is used in spectral formulations of GNNs.
- **Edge Laplacian $\mathbf{L}_e \in \mathbb{R}^{N_e \times N_e}$:** A Laplacian matrix defined over the edges, capturing connectivity patterns between edges, relevant in edge-based message passing formulations.
- **Incidence Matrix $\mathbf{T} \in \{0, 1\}^{N_v \times N_e}$:** A binary matrix representing the node-edge incidence relationship. Each column corresponds to an edge and contains ones in the rows of its incident nodes, enabling the mapping of flows through the network.
- **Edge Features $\mathbf{E} \in \mathbb{R}^{N_e \times 5}$:** A matrix where each row corresponds to an edge in the graph, and each column represents an edge-level attribute. Each edge feature includes the K constant, the maximum compression ratio β , and the upper and lower flow limits.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

The model outputs include node-level predictions corresponding to the injected flow at each node and edge-level predictions representing the transported flow across the network edges.

2.1.6 Task-Dependent Loss Functions

The output layer and corresponding loss functions in CensNet are designed to be task-dependent. For regression tasks, the loss function can be formalized as a regularized mean square error (MSE) loss. The MSE loss measures the difference between the predicted outcomes and the actual values, providing a natural fit for regression problems.

We define the loss function for graph regression as follows:

$$\mathcal{L}(\Theta) = \sum_{r=1}^R \|Y_r - \hat{Y}_r\|_2^2 + \lambda \|\Theta\|_p, \quad (2.19)$$

where:

- $Y_r \in \mathbb{R}$ is the ground-truth value for the r -th regression target.
- $\hat{Y}_r \in \mathbb{R}$ is the predicted value for the same feature, computed from the final node representation of the graph.
- $\|Y_f - \hat{Y}_r\|_2^2$ is the squared error for feature r , and the sum is taken over all R output features.
- $\Theta \in \mathbb{R}^P$ is the set of learnable parameters in the model, and $\|\Theta\|_p$ denotes the L_p norm regularization term.
- $\lambda \in \mathbb{R}_+$ is a regularization coefficient that penalizes overly complex models to reduce overfitting.

2.2 Experimental Setup and Results

The generated samples served as training data for a CensNet-based model to solve the natural gas transportation problem, which was designed as a faster alternative to the optimization-based model. This model is built to focus on predicting node and edge-level characteristics, incorporating the structure of the network and its connectivity into the learning process.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

In this experimental setup, different network operation scenarios were generated by perturbing the users' consumption values with added noise and subsequently solving each scenario using the linear constrained optimization model presented in the previous section. The optimization was performed using APOPT (Advanced Process Optimizer) through the Python package GEKKO [39], a solver for large-scale linear and nonlinear optimization problems. APOPT was used to obtain the output variables for each scenario, serving as the ground truth data with which the CensNet model was trained. The noise levels ranged from 5% to 25%, applied to the parameters of the networks to simulate diverse operating conditions.

The first network is a small-scale test network consisting of 8 nodes, while the second represents the Colombian natural gas transportation system, a more extensive and complex network consisting of 63 nodes. These networks were used to evaluate the performance of the proposed model under varying scenarios. The model penalizes deviations in node and edge losses, which directly impact gas flow efficiency through the system. To achieve this, the architecture is structured as a CensNet-based neural networks, with customizable depth (number of layers), channels, and dense layers, ensuring flexibility in adapting to small-scale and large-scale networks, such as the Colombian system.

A general outline of the model can be seen in fig. 2.3 and the basic components of the model are explained below:

- **Normalization and Predense Layers:** The node and edge inputs receive feature-wise normalization to standardize the data. Following this, the inputs are passed through two dense layers, each with N channels. The purpose of these pre-dense layers is to transform the feature space before applying the convolutional layers, enabling the network to project the raw input data into a higher-dimensional latent space that facilitates more expressive and effective learning in the subsequent convolutional stages.
- **CensNet Blocks Layers:** The main body of the network consists of N convolutional blocks. Each block applies a CensNet convolution, which updates both node and edge features by considering the structural relationships encoded in the node and edge Laplacians, as well as the incidence matrix. Batch normalization follows each convolution to stabilize learning. This structure allows the model to capture complex interactions between nodes and edges and propagate information across the graph, learning how local features influence the broader system.
- **Post-dense Layer:** After passing through the convolutional blocks, the node

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS



Figure 2.3: General outline of the CensNet-based model used.

and edge features are further processed by a series of dense layers. The number of dense layers N_{dense} is adjustable. These layers further refine the learned features, enabling the model to output node and edge-level predictions.

- **Losses and Outputs:** The final outputs of the network are the node-level and edge-level predictions, denoted by $\hat{\mathbf{X}}_v$ and $\hat{\mathbf{X}}_e$, respectively. The node predictions $\hat{\mathbf{X}}_v$ correspond to the estimated flow at each node, while the edge predictions $\hat{\mathbf{X}}_e$ represent the flow along the edges. Both outputs are penalized based on their respective losses, which are calculated by comparing the predicted values to ground truth values and evaluating how well the physical constraints are respected.

The loss functions ensure that the model accurately predicts the node and edge flows while satisfying the physical constraints of the system. These constraints are essential for ensuring that the predicted flows are feasible within the operational limitations of the network.

- **Model Optimization:** The model is trained using backpropagation with the Adam optimizer. The training process involves minimizing the node and edge loss functions, which penalize incorrect flow predictions and deviations from the expected behavior of the network.

The network training was done under the following conditions: data was partitioned into training, validation, and test sets, with percentages of 60%, 20%, and 20% using

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

2000 samples for the 8-node network and 2400 for the 63-node network. The learning rate schedule followed an Exponential Decay approach, with an initial learning rate of 1×10^{-2} decay steps of 1000, and a decay rate of 0.9. The model was trained over 1500 epochs, ensuring the parameters had sufficient time to converge for both node- and edge-level predictions using a Leaky Relu activation function with an alpha parameter of 0.2.

The model utilized three hyperparameters: N *channels*, which corresponds to the units in the pre-dense and convolutional layers (CensNet), N *dense*, representing the number of post-dense layers, and N *layers*, denoting the number of convolutional layers. Additionally, the model penalizes deviations in the network's response, specifically targeting MSE between the actual flow and the predicted flow at both the nodes and pipelines. Two tests were performed to determine the optimal set of hyperparameters. The objective was to compare the effect of using different combinations of loss functions during training. In the first test, only the weight associated with the nodal flow loss was considered, while in the second test, both nodal and edge flow losses were included. A test using only the edge loss was not conducted, as it was considered inappropriate to disregard the nodal flow loss. This is because the nodal flows represent the injection and demand values, which are essential for the physical consistency of the system and are directly associated with the behavior of the network at each node. In each test, the hyperparameters were optimized using the open-source Optuna framework [40], with the following search space: N *channels* ranged from 16 to 64, N *layers* from 1 to 5, and N *dense* from 2 to 32.

2.3 Results and Discussion

This section presents the evaluation of the proposed CensNet-based model and its comparison with a traditional optimization model (APOPT) for two gas transportation networks: a synthetic 8-node network and a real-world 63-node network representing the Colombian gas system. Two experimental settings are analyzed in each case: one that considers only nodal flow prediction loss (CensNet (N)), and another that incorporates both nodal and edge flow losses (CensNet (N+E)). For each configuration, hyperparameter tuning was performed to optimize the performance of the CensNet models. In the 8-node network, the best hyperparameters for the CensNet (N) configuration were N *channels* = 21, N *layers* = 5, and N *dense* = 4, while for CensNet (N+E), the optimal values were N *channels* = 44, N *layers* = 4, and N *dense* = 4. In the case of the 63-node Colombian network, the chosen hyperparameters for the

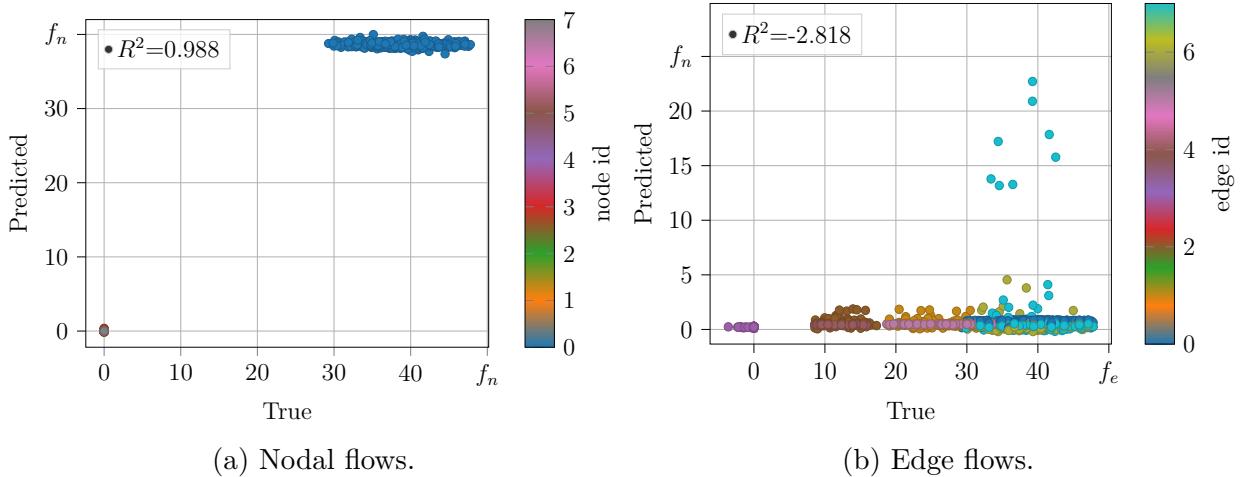


Figure 2.4: Model results using only the loss associated with nodal flow predictions in the 8-node network.

CensNet (N) configuration were $N \text{ channels} = 43$, $N \text{ layers} = 2$, and $N \text{ dense} = 2$, and for CensNet (N+E), the hyperparameters were $N \text{ channels} = 25$, $N \text{ layers} = 5$, and $N \text{ dense} = 4$. These configurations were selected based on the minimal validation losses obtained during training.

2.3.1 Prediction Accuracy

Across both networks, the CensNet model effectively identifies which nodes inject gas into the system. Even when trained using only nodal losses, the model achieves high R^2 scores for nodal flow predictions (0.988), demonstrating its ability to generalize the mapping between inputs and nodal injections. This trend holds for both the 8-node and 63-node networks, where scatter plots (Figures 2.5 and 4.5) confirm the alignment between actual and predicted values.

For the 8-node network, where gas injection occurs only at a single node, the CensNet (N) model effectively captures this behavior, with minor deviations from the actual values. Similarly, for the 63-node Colombian network, the model successfully identifies the gas injection points, demonstrating a high degree of accuracy with an R^2 value of 0.996. These results validate the model's ability to generalize well even when dealing with larger-scale networks.

However, edge flow predictions show a clear distinction between the two training

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

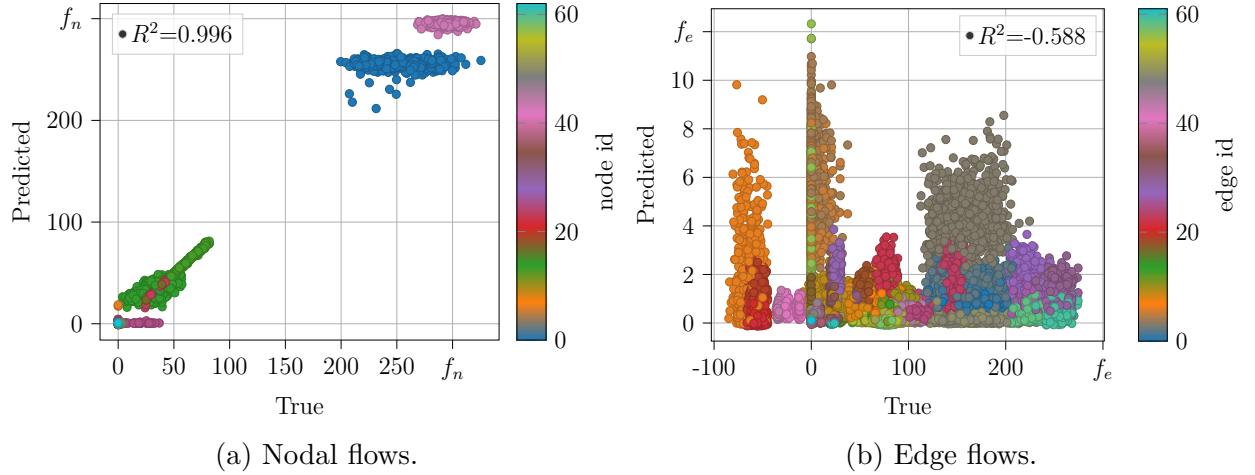


Figure 2.5: Model results using only the loss associated with nodal flow predictions in the colombian 63-node network.

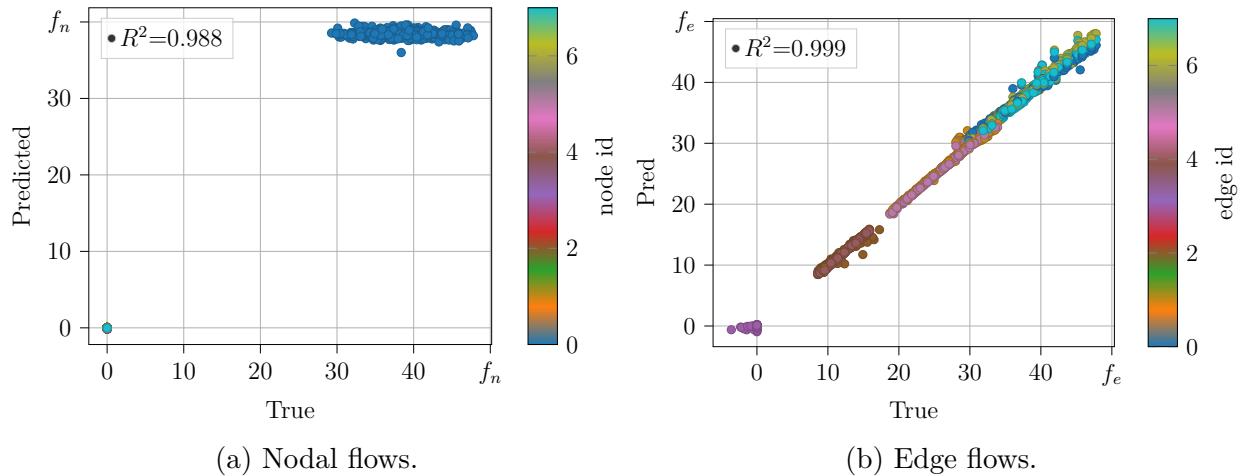


Figure 2.6: Model results using the losses associated with the flows in nodes and edges predictions in the 8-node network.

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

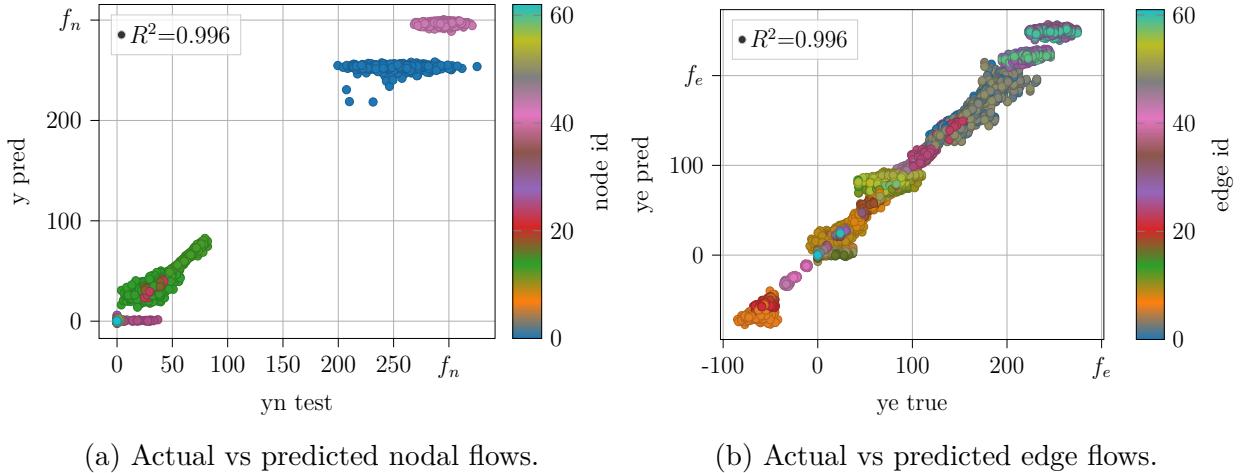


Figure 2.7: Model results using the losses associated with the predicted flows injected at the nodes and transported by the pipelines and compressors in the colombian 63-node network.

regimes. In both networks, the CensNet model trained only on nodal loss performs poorly in capturing edge-level flow dynamics, as seen in Figures 2.4b and 2.5b. This is expected since edge flow behavior was not directly optimized in these cases. Once edge loss is incorporated (Censnet (N+E)), prediction accuracy improves dramatically, as shown by the near-perfect alignment in the edge flow scatter plots (Figures 2.6 and 2.7). In the 8-node network, the R^2 score for edge flow predictions reaches 0.999, while the 63-node case achieves an R^2 of 0.996, indicating consistent model generalization across both small and large-scale networks.

The improvements in edge flow predictions are particularly notable in the Colombian network, where the inclusion of edge flow loss helps the model capture the complexities of the pipeline and compressor dynamics, which were previously overlooked in the nodal-only model. This improved prediction for edge flows also demonstrates the model's capability to scale effectively to larger and more complex systems.

2.3.2 Gas Balance Consistency

When compared to the APOPT optimization model, the CensNet achieves comparable performance in terms of flow predictions, particularly when both nodal and edge losses are used. In the 8-node case, for instance, the CensNet (N+E) model achieves

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

Method	Node Error	Edge Error	Balance Error	Time
CensNet (N)	0.000 ± 18.144	23.760 ± 14.042	5.710 ± 16.850	0.650 ± 0.442
CensNet (N+E)	0.070 ± 18.081	0.080 ± 19.884	0.080 ± 1.171	0.640 ± 0.444

Table 2.1: Differences in mean and standard deviation between APOPT and CensNet-based models for node error, edge error, balance error, and prediction time. Each value represents the difference APOPT - CensNet, computed over 100 samples.

Method	Node Error	Edge Error	Balance Error	Time
CensNet (N)	0.030 ± 69.551	62.610 ± 81.630	0.040 ± 60.837	4.880 ± 5.590
CensNet (N+E)	0.010 ± 69.438	0.070 ± 115.259	0.010 ± 23.398	4.870 ± 5.591

Table 2.2: Differences in mean and standard deviation between APOPT and CensNet-based models for node error, edge error, balance error, and prediction time, across 100 samples. Each value represents the difference APOPT - CensNet.

a significant reduction in edge error compared to the nodal-only model, with a mean edge difference of 0.08 units versus 23.76, as shown in Table 2.1. Additionally, the gas balance error improves substantially, from 5.71 to 0.08, indicating improved mass conservation when edge information is incorporated into training.

For the 63-node case, the advantage of using both nodal and edge losses remains consistent. While the nodal-only model (CensNet (N)) yields a large edge error difference of 62.61, this is reduced to just 0.07 by the CensNet (N+E) model, as shown in Table 2.2. Moreover, the balance error also decreases from 0.04 to 0.01, reinforcing the benefit of training with comprehensive physical constraints.

In terms of nodal predictions, the differences between APOPT and CensNet remain small in both networks. The maximum difference is under 0.1 units, which confirms the CensNet's capability to match optimization-based predictions for node flows when trained appropriately.

Regarding computational performance, the CensNet-based model significantly outperforms the optimization model. T-tests were conducted to compare the prediction times for the APOPT and CensNet models across 100 unseen cases. These tests revealed statistically significant differences in computational time between the two models, with the CensNet-based model taking considerably less time to generate predictions. For the 8-node network, the t-test returned a T-statistic of 14.94, with 198 degrees of freedom and a p-value of 1.32×10^{-34} . Similarly, for the 63-node network, the t-test showed a T-statistic of 47.29 and a p-value of 4.92×10^{-110} . These results confirm the significant

CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH NEURAL NETWORKS

computational advantage of the CensNet-based model over the optimization approach.

2.4 Discussion

Aligned with the first objective, this section discusses the effectiveness of a Graph Neural Networks-based regression approach that leverages the topology of natural gas networks to reduce computation time for operation estimation. In this section, we analyze and summarize the performance of the proposed CensNet-based model in predicting the operational parameters of natural gas transportation systems. The evaluation was conducted on two test cases: a simplified 8-node network and a more complex real-world system with 63 nodes representing the Colombian natural gas system. The results demonstrate the CensNet model’s ability to deliver accurate predictions with significant improvements in computational efficiency compared to traditional optimization methods.

A key insight from the experiments is that incorporating both nodal and edge flow losses in the training objective is important to achieving physically consistent and accurate predictions. While the model trained using only nodal losses was able to identify gas injection nodes with high R^2 scores—up to 0.996 in both systems—its inability to enforce edge-level flow consistency led to significant errors in gas balance and edge flow predictions. This limitation is expected, as edge flows are governed by physical constraints not captured by nodal outputs alone. Once edge-related losses were included (CensNet (N+E)), the model achieved near-perfect alignment with the actual data for both nodal and edge flows, as confirmed by scatter plots and R^2 values reaching 0.999 for edge predictions (see Figures 2.6 and 2.7). Another contribution is the ability of the CensNet-based model to generalize across different network scales and topologies using topology-informed learning. Despite the increased complexity of the 63-node system, the model retained its predictive accuracy and computational advantages, highlighting the scalability of the approach.

In terms of computational performance, the CensNet model consistently outperformed the optimization-based approach. T-tests applied to prediction times in both experiments confirmed statistically significant differences, with p-values smaller than 10^{-30} in all cases. For instance, in the 8-node experiments, the CensNet achieved prediction times under 0.15 seconds on average, compared to over 0.79 seconds for the optimizer, with a T-statistic of 14.94 and p-value of 1.32×10^{-34} in the base case (see ??). Similar results were observed in the 63-node system, where the CensNet achieved an average prediction time of 0.14 seconds versus 5.01 seconds for the optimizer. The

*CHAPTER 2. NATURAL GAS SYSTEM PREDICTION USING GRAPH
NEURAL NETWORKS*

T-statistic of 47.29 and p-value of 4.92×10^{-110} (see ??) confirm the robustness of the computational advantage.

Overall, the experiments provide validation that the inclusion of physical losses, combined with GNN architectures that respect the topological structure of gas networks, leads to models that are both accurate and efficient. This supports the premise that GNNs, when appropriately guided by physical insights, can serve as powerful surrogates for traditional optimization tools in the context of operational estimation for natural gas systems.

Chapter 3

Optimization Using Mathematical Programs with Complementarity Constraints

This chapter presents an optimization approach based on Mathematical Programs with Complementarity Constraints (MPCC) to address the natural gas transportation problem, with a particular focus on modeling the nonlinear behavior of gas flow in pipelines governed by the Weymouth equation. The MPCC framework offers a mathematical structure for incorporating the nonconvexities and discontinuities inherent in gas network models, thereby enabling a more precise representation of the underlying physical constraints. This methodology is relevant in scenarios involving the interconnection between natural gas and electrical systems, where operational decisions in one network directly influence the other. The results presented in this chapter have been peer-reviewed and published in an article featured in an A1-ranked journal according to the PUBLINDEX classification.

3.1 Formulation of Interconnected Power and Gas Systems

An interconnected system can be effectively represented by a directed graph denoted as $G = \{\mathcal{N}, \mathcal{E}\}$, where the sets of units \mathcal{N} and edges \mathcal{E} consider all power and gas components along with their interconnections. On the electrical power side, the system holds power units $\mathcal{N}_P \subset \mathcal{N}$, termed buses, and power edges $\mathcal{B} \subset \mathcal{E}$ or branches.

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

The power buses comprise generators $\mathcal{G} \subset \mathcal{N}_P$ injecting power and users $\mathcal{D} \subset \mathcal{N}_P$ demanding power [41]. The branches $\mathcal{B} = \{b = (n, m) \mid n, m \in \mathcal{N}_P\}$ connect the buses to make the electrical power flow from the generators to the users. Although the physical power flow is alternating current, the system is accurately modeled using a linear direct current (DC) approximation. The DC model ignores reactive power flows and voltage magnitude fluctuations and approximates active power flows using linear transfer distribution factors [42]. Further, the linear characteristics allow stating linear programming problems. Thus, the DC model serves as an appropriate approximation for many power system operations and planning studies, providing a balance of accuracy and computational tractability [43].

Then, the optimization problem of the interconnected system seeks to minimize the operation costs for satisfying the demands of the interconnected system while encompassing the power and gas constraints. Specifically, the following cost function linearly combines the flows of power and gas through the operation costs of the interconnected system elements:

$$\begin{aligned} \min_{\mathcal{P}, \mathcal{F}} \quad & \sum_{g \in \mathcal{G}} C_g^t P_g^t + \sum_{d \in \mathcal{D}} C_d^t P_d^t + \sum_{w \in \mathcal{W}} C_w^t f_w^t + \\ & \sum_{p \in \mathcal{P}} C_p^t f_p^t + \sum_{c \in \mathcal{C}} C_c^t f_c^t + \sum_{u \in \mathcal{U}} C_u^t f_u^t + \\ & \sum_{s \in \mathcal{S}} C_{s+}^t f_{s+}^t + \sum_{s \in \mathcal{S}} C_{s-}^t f_{s-}^t + \sum_{s \in \mathcal{S}} C_s^t V_s^t \end{aligned} \quad (3.1)$$

where C_g^t denotes the generation cost by the g -th bus and C_d^t the unsupplied power demand for the d -th user. Regarding the natural gas system, terms C_w^t , C_p^t , C_c^t and C_u^t are the same as those used in Equation (2.14). However, some additional terms are considered in this case: C_{s+}^t , C_{s-}^t , and C_s^t represent the costs of injecting, extracting, and storing gas at the s -th storage station.

Therefore, the decision variables for the optimization problem are P_g^t for the generated power, P_d^t for the unsupplied power, f_w^t for the inject gas flow, f_p^t and f_c^t for the transported gas through pipeline p and compressor c , f_u^t for the unsupplied gas demand, f_{s+}^t , f_{s-}^t , and f_s^t for injecting, extracting, and storing gas. Traditionally, a transported gas with a positive value of $f_p^t > 0$ moves in the predefined direction, while a negative value flows in the opposite one, with no impact on the optimization process. On the other hand, compressor stations solely allow unidirectional gas flow, expressed as $f_c^t \geq 0$. By optimizing this integrated cost function while adhering to the system's

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

operational constraints, the proposed methodology effectively balances the demands of both energy systems, leading to a comprehensive solution that minimizes costs while ensuring reliable and efficient operation.

Optimization of the integrated cost function in Equation (3.1) while adhering to the system's operational constraints must lead to a comprehensive solution balancing the demands of both energy systems while ensuring reliable and efficient operation. Three sets of operational constraints describe the within and between power and gas interplay.

The first constraint set guarantees a stable power system operation: Equation (3.2) ensures that the generated power P_g^t lies between the technical minimum \underline{P}_g^t and maximum \overline{P}_g^t . Equation (3.3) bounds the power flow through the transmission line P_l^t , preventing damages, such as overheating. Equation (3.4) models the power flow over the electrical network through the reactance-based relationship of the power flow P_l^t , the line susceptance B_{nm} , and the voltage angles θ_n, θ_m at buses n, m . Equation (3.5) limits the unsupplied power P_d^t to the user demand \overline{P}_d^t . Equation (3.6) ensures stable operating conditions within the interconnected power grid by restricting the bus voltage angles. Equation (3.7) defines the power balance at each bus, i.e., the total input and generated power must equal the total output and unsupplied power, being $\mathcal{B}_{n+} = \{(m, n') \in \mathcal{B} : n' = n\}$ and $\mathcal{B}_{n-} = \{(n', m) \in \mathcal{B} : n' = n\}$ the set of inflow and outflow transmission lines at the n -th bus, respectively.

$$\underline{P}_g^t \leq P_g^t \leq \overline{P}_g^t \quad \forall g \in \mathcal{G}, \quad (3.2)$$

$$-\overline{P}_l^t \leq P_l^t \leq \overline{P}_l^t \quad \forall l \in \mathcal{L}, \quad (3.3)$$

$$P_l^t = B_{nm}(\theta_n - \theta_m) \quad \forall l = (n, m) \in \mathcal{L}, \quad (3.4)$$

$$0 \leq P_d^t \leq \overline{P}_d^t \quad \forall d \in \mathcal{D}, \quad (3.5)$$

$$-\overline{\theta}_n^t \leq \theta_n^t \leq \overline{\theta}_n^t \quad \forall n \in \mathcal{N}_P, \quad (3.6)$$

$$\sum_{\substack{l \in \mathcal{B}_{n+} \\ g=n}} P_l^t + P_g^t = \sum_{\substack{l \in \mathcal{B}_{n-} \\ d=n}} P_l^t + P_d^t \quad \forall n \in \mathcal{N}_P \quad (3.7)$$

The second constraint set interconnects natural gas and electrical power systems through gas-fired power plants generating electricity, as expressed by Equation (3.8), where f_n^t stands for the natural gas fuel consumption to generate a power P_n^t at generator bus $n \in \mathcal{N}_I$, the heat-rate HR_n defines the generator efficiency, and the set $\mathcal{N}_I = \mathcal{G} \cap \mathcal{U}$ holds all thermoelectrical plants as the gas users that generate power.

$$f_n^t = P_n^t \cdot \text{HR}_n, \quad \forall n \in \mathcal{N}_I, \quad (3.8)$$

The third constraint set models the gas transportation system: Besides Equations (2.15) to (2.18), Equation (3.9) fixes safe operating limits for the pressure on the n -th node π_n^t as $[\underline{\pi}_n^t, \overline{\pi}_n^t]$. The constraint in Equation (3.10) asserts that the compression ratio π_m^t/π_n^t cannot physically exceed the compressor's design limitation $\beta_c \geq 1 \forall c = (n, m) \in \mathcal{C}$, enabling the representation of different compressors by adjusting the values of $\beta_c \geq 1$. Equations (3.11) and (3.12) limit the gas injection f_{s+}^t and extraction f_{s-}^t rates at storage facilities according to the feasible operating range determined by the currently stored volume V_s^t , the maximum storage limit \overline{V}_s , and the minimum storage limit \underline{V}_s . In turn, Equation (3.13) balances the gas storage unit such that gas volume at operation period t , V_s^t , equals the volume from period V_s^{t-1} plus the difference between injected f_{s+}^{t-1} and extracted f_{s-}^{t-1} gas flow; a fundamental constraint for modeling the dynamics of gas storage over time. Lastly, Equation (3.14), known as the Weymouth equation, summarizes the physical behavior of gas flow through pipelines by relating the gas flow through the pipeline f_p^t to the pressures at the ends of the pipeline $\pi_n^t, \pi_m^t \forall p = (n, m) \in \mathcal{P}$. The constant K_{nm} in this equation is a pipeline-specific coefficient that encapsulates various physical characteristics, including its length, diameter, and friction factor, as well as the properties of the gas being transported. The Weymouth equation defines a nonlinear, nonconvex, disjunctive flow-pressure relationship that hampers the optimization of the gas transport system.

$$\underline{\pi}_n^t \leq \pi_n^t \leq \overline{\pi}_n^t \quad \forall n \in \mathcal{N}_f, \quad \forall t \in T \quad (3.9)$$

$$\pi_m^t \leq \beta_c^t \pi_n^t \quad \forall c = (n, m) \in \mathcal{C}, \quad \forall t \in T \quad (3.10)$$

$$0 \leq f_{s+}^t \leq V_{0s} - \underline{V}_s \quad \forall s \in \mathcal{S}, \quad \forall t \in T \quad (3.11)$$

$$0 \leq f_{s-}^t \leq \overline{V}_s - V_s^t \quad \forall s \in \mathcal{S}, \quad \forall t \in T \quad (3.12)$$

$$V_s^t = V_s^{t-1} + f_{s-}^{t-1} - f_{s+}^{t-1} \quad \forall s \in \mathcal{S}, \quad \forall t \in T \quad (3.13)$$

$$sgn(f_p^t)(f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2) \quad \forall p = (n, m) \in \mathcal{P}, \quad \forall t \in T \quad (3.14)$$

3.2 Mathematical Programming with Complementarity Constraints for Weymouth Approximation

The Weymouth equation is the fundamental model for gas flow through pipelines. Nevertheless, it presents a challenge for optimal interconnected operation due to its nonlinearity, which arises from the signum function determining the gas flow direction. This nonlinearity results from the complex physics of gas flow, making it challenging to find optimal solutions for gas transportation systems [44]. Traditional optimization methods have difficulties in solving the nonconvexity of the Weymouth equation. However, recent advances in optimization techniques, particularly mathematical programs with complementarity constraints (MPCC), offer a solution to address this issue. MPCC specializes in handling complementarity constraints and non-convexities, making it well-suited to tackle the intricacies of the Weymouth equation [45]. This type of formulation involves optimization problems of the general form:

$$\mathcal{O} : \min q(x, y) \tag{3.15a}$$

$$\text{s.t. } h_i(x, y) = 0 \tag{3.15b}$$

$$g_j(x, y) \geq 0 \tag{3.15c}$$

$$0 \leq G_k(x) \perp H_k(y) \geq 0 \tag{3.15d}$$

Here, $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ are the decision variables, typically representing primal and dual components, or system and auxiliary variables, respectively. The objective function $q(x, y) \in \mathbb{R}$ is to be minimized, subject to a set of equality constraints $h_i(x, y) = 0$ for all $i \in \mathbb{I}$, where $|\mathbb{I}|$ denotes the number of such constraints, and a set of inequality constraints $g_j(x, y) \geq 0$ for all $j \in \mathcal{R}$, with $|\mathcal{R}|$ indicating the number of inequality constraints. The complementarity conditions in Equation (3.15d) involve mappings $G_k : \mathbb{R}^n \rightarrow \mathbb{R}$ and $H_k : \mathbb{R}^m \rightarrow \mathbb{R}$, defined for all $k \in \mathcal{K}$, where $|\mathcal{K}|$ is the total number of complementarity pairs. The notation $0 \leq G_k(x) \perp H_k(y) \geq 0$ expresses the condition that for each $k \in \mathcal{K}$, the product $G_k(x) \cdot H_k(y) = 0$, with both terms being non-negative. This ensures that at least one element in each complementarity pair is zero at the optimum, thereby modeling the disjunctive or switching behavior characteristic of gas flow through pipelines.

To deal with the non-convexity, this work rewrites the Weymouth equation as the

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

following mathematical program with two complementarity constraints:

$$\begin{aligned}
 \mathcal{O}_W : \min_{y_p^t} & -y_p^t f_p^t \\
 \text{s.t. } & y_p^t (f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2) \\
 & -1 \leq y_p^t \leq 1 \\
 & f_p^t = f_{p+}^t - f_{p-}^t \\
 & 0 \leq f_{p+}^t \perp (y_p^t + 1) \geq 0 \\
 & 0 \leq f_{p-}^t \perp (1 - y_p^t) \geq 0
 \end{aligned} \tag{3.16}$$

where $f_{p+}^t \geq 0$ and $f_{p-}^t \geq 0$ hold the positive and negative components of the gas flow in the p -th pipeline at operation period t , for assessing directional flow.

In problem \mathcal{O}_W , the objective function minimizes the product $-y_p^t f_p^t$, which promotes alignment between the direction of gas flow and its magnitude, as defined by the auxiliary variable $y_p^t \in [-1, 1]$. The first constraint in \mathcal{O}_W enforces the Weymouth relationship in its relaxed form, where the product $y_p^t (f_p^t)^2$ equals the pressure drop across the pipeline scaled by the constant K_{nm} , such that $K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2)$ captures the non-linear physics of gas flow between nodes n and m at time t . The second constraint bounds the direction variable y_p^t between -1 and 1 , allowing it to encode bidirectional flow while keeping the model numerically stable. The third constraint defines the net gas flow f_p^t as the difference between its non-negative forward and reverse components, f_{p+}^t and f_{p-}^t , respectively.

The final two constraints are complementarity conditions: $0 \leq f_{p+}^t \perp (y_p^t + 1) \geq 0$ and $0 \leq f_{p-}^t \perp (1 - y_p^t) \geq 0$. These conditions enforce that the gas can only flow in one direction at a time. Specifically, if $f_{p+}^t > 0$, then $y_p^t = -1$, indicating flow from node m to node n , while if $f_{p-}^t > 0$, then $y_p^t = 1$, indicating flow in the opposite direction. This formulation allows the MPCC to capture the switching behavior of gas flows due to changes in pressure direction, without resorting to discrete variables or binary decision-making.

Solving MPCC presents a unique set of challenges distinguishing it from traditional optimization problems. One notable challenge is the need for regularity properties, making MPCC more complex [46]. Compared to smooth optimization problems, where gradients and Hessians provide valuable information for optimization algorithms, MPCC often lacks these properties, leading to difficulties in devising efficient numerical methods.

It is widely recognized that conventional constraint qualifications in nonlinear programming, such as LICQ and MFCQ, are typically not satisfied in the case of MPCC.

*CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH
COMPLEMENTARITY CONSTRAINTS*

As a result, KKT conditions commonly associated with MPCC may not be applicable or valid at a local minimization point [47]. Therefore, posing relaxed nonlinear programs (RNLP) deals with the numerical resolution of MPCC by introducing a positive regularization parameter $\epsilon \in \mathbb{R}^+$ that simplifies the solution and properly handles the inequalities [48]. These programs typically satisfy constraint qualifications, making them more amenable to efficient optimization techniques. Relaxing MPCC ensures that inequalities are appropriately treated as inactive, particularly when $G_k(x)H_k(y) \leq \epsilon$, enhancing their structural integrity. Besides, relaxed programs reliably approximate the original problem as $\epsilon \rightarrow 0$ [49]. Hence, instead of working with the original problem \mathcal{O}_W as can be seen in Equation (3.16), the relaxed problem \mathcal{O}_ϵ is formulated:

$$\mathcal{O}_\epsilon : \min_{y_p^t} -y_p^t f_p^t \quad (3.17a)$$

$$\text{s.t. } y_p^t (f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2) \quad (3.17b)$$

$$f_p^t = f_{p+}^t - f_{p-}^t \quad (3.17c)$$

$$-1 \leq y_p^t \leq 1 \quad (3.17d)$$

$$f_{p+}^t (y_p^t + 1) \leq \epsilon \quad (3.17e)$$

$$f_{p-}^t (1 - y_p^t) \leq \epsilon \quad (3.17f)$$

Theoretically, the relaxed problem offers fundamental properties that tackle challenging MPCC problems [50]. Firstly, the relaxed approach guarantees the convergence to the true MPCC solution as $\epsilon \rightarrow 0$. Additionally, the boundedness of Lagrange multipliers ensures numerical stability and avoids issues with infinitely large values during optimization. Lastly, the local uniqueness of the \mathcal{O}_ϵ solution under specific conditions guarantees a single and well-defined solution. Therefore, the proposed relaxed optimization problem deals with the non-convexity in the Weymouth equation while guaranteeing the KKT conditions around ϵ , posing a standard optimization problem, and avoiding ambiguity in interpreting results.

3.3 Case studies

The current section validates the proposed MPCC approach by comparing its performance against two well-established methods for approximating the Weymouth equation: i) The Taylor series approach that piecewise approximates Weymouth with line

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

segments [51] and ii) The Second-order cone (SOC) programming that introduces a two-stage optimization, namely, flow direction estimation and cost minimization [52].

The validation aims to quantify the inherent errors and the relationship between cost and error in the contrasted approaches to support their real-world pertinence. Therefore, this work reports two performance metrics: the cost function in Equation (3.1) that assesses the capacity for optimally operating an integrated system and the Weymouth error metric ($WE_p^t \in \mathbb{R}^+$) for quantifying the required flow to guarantee equality for pipeline p at time instant t in Equation (3.18):

$$WE_p^t = \left| f_p^t - (K_{nm}|(\pi_n^t)^2 - (\pi_m^t)^2|)^{1/2} \right|, \quad \forall p = (n, m) \in \mathcal{P}. \quad (3.18)$$

Hence, the WE_p^t metric, measured in million standard cubic feet per day (MMSCFD), explains the approximations' inherent sensitivity and validates the significance of their differences. The validation contrasts Taylor, SOC, and MPCC approaches in three case studies of interconnected systems with different complexities.

3.3.1 Case Study I: 9/8 System

The first case study analyzes a small-scale integrated energy system with a closed-loop topology that necessitates modeling bidirectional flows. The electrical component of the system has $|\mathcal{G}| = 3$ generators, $|\mathcal{N}_p| = 9$ buses, $|\mathcal{B}| = 9$ branches, and $|\mathcal{D}| = 3$ demand nodes. This electricity network is coupled with a natural gas network comprising 8 nodes, including $|\mathcal{W}| = 1$ gas supply well, $|\mathcal{P}| = 6$ pipelines, $|\mathcal{C}| = 2$ compressor stations, and $|\mathcal{U}| = 2$ gas demand nodes. The integration is achieved through $|\mathcal{I}| = 1$ interconnection point between the two networks [53].

To assess the performance of Weymouth approximation approaches on the 9/8 system, a Monte Carlo experiment estimates the cost function and Weymouth error distributions by solving the optimization problem for one day ($\mathcal{T} = \{1\}$) one hundred times with uniformly sampled natural gas demands. Further network parameter details can be found in the publicly available repository OptiGasFlow (<https://github.com/cblancom/optigasflow>, accessed on 05 April 2024). Figure 3.2 depicts the cost function histogram for Taylor, SOC, and MPCC approaches. Remarkably, the three histograms evidence identical distribution patterns, leading to regular solutions across approaches.

The boxplots in Figure 3.3 show the Weymouth approximation error distribution for each pipeline using the three approaches. The error distributions, including median and interquartile range, indicate that MPCC consistently maintains accuracy throughout

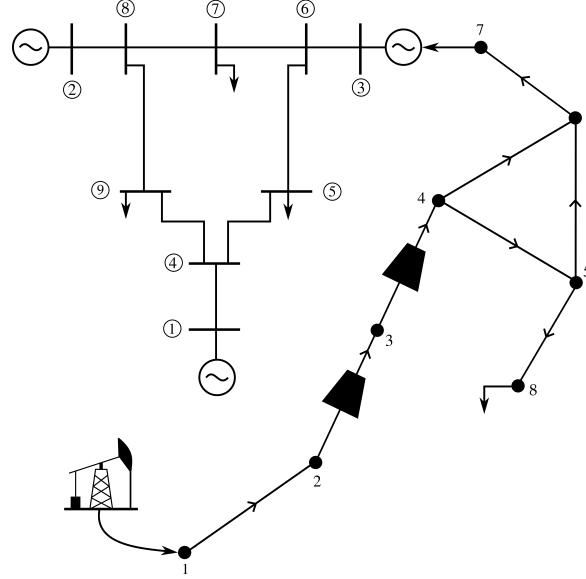


Figure 3.1: Integrated system 9/8 used in Case Study I, modified from the MATPOWER-NATURAL GAS (MPNG) software.

the network. In contrast, the widely varying errors of the Taylor and SOC approaches suggest a lack of consistency in the achieved solution. Therefore, in a small network, the proposed MPCC approach converges to identical operational costs of Taylor and SOC, even in rationing, while meeting all linear constraints and improving the Weymouth approximation, achieving error reductions by several orders of magnitude compared to the other methods.

3.3.2 Case II: 118/48 System

The following case simulates a complex, large-scale electric grid system, the widely studied IEEE 118 bus system [54], consisting of 54 generator buses, 9 fed by the gas system, 186 transmission lines, and 99 users, that is, $|\mathcal{G}| = 54$, $|\mathcal{B}| = 186$, $|\mathcal{D}| = 99$. This electric grid interconnects with a 48-node natural gas system featuring 9 supply wells, 46 pipelines, eight compressor stations, and 22 user nodes through 9 connection points, i.e., $|\mathcal{W}| = 9$, $|\mathcal{P}| = 46$, $|\mathcal{C}| = 8$, $|\mathcal{U}| = 22$, $|\mathcal{I}| = 9$ [14]. The network topology deliberately introduces closed flow loops to stress the solver and the constraint approximations, as do real-world systems.

Figure 3.4 depicts the histogram of relative cost differences for the MPCC proposal

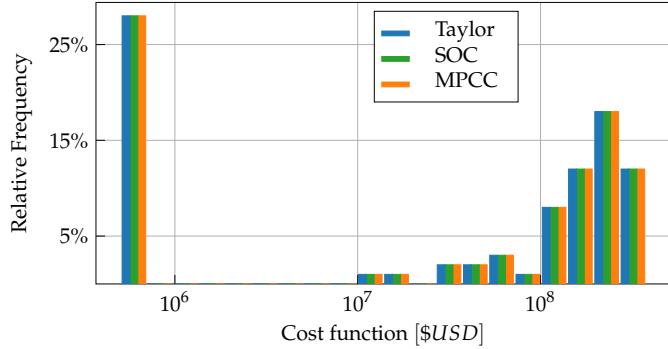


Figure 3.2: Cost function histogram for the Taylor, SOC, and MPCC Weymouth approximation approaches in the 9-bus 8-node system.

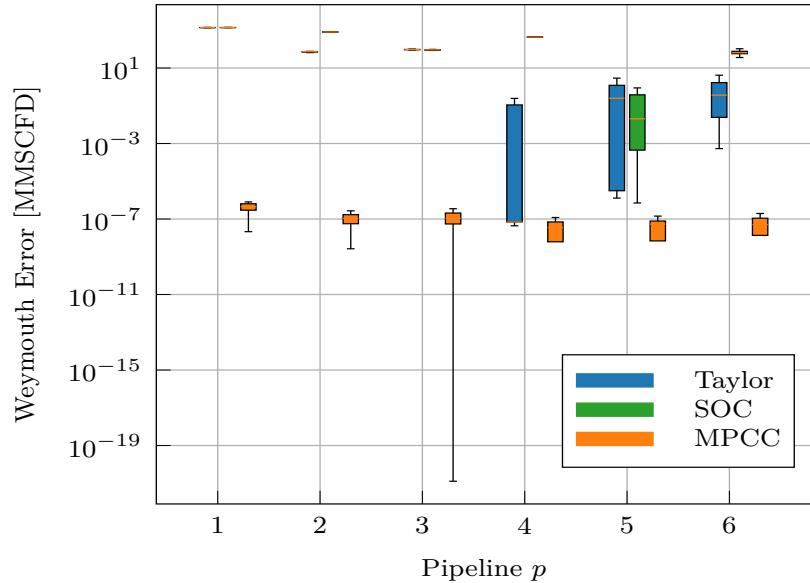


Figure 3.3: Boxplot of Weymouth error distribution for each pipeline in the 8/9 system attained by contrasted approximation approaches.

to Taylor and SOC baselines from a hundred trials of the Monte Carlo experiment and a considered operation of one day ($T = 1$). It is worth noting that both baselines yielded the same cost function values. The relative difference between MPCC and the baselines is always positive, indicating that the complementarity constraint formulation consis-

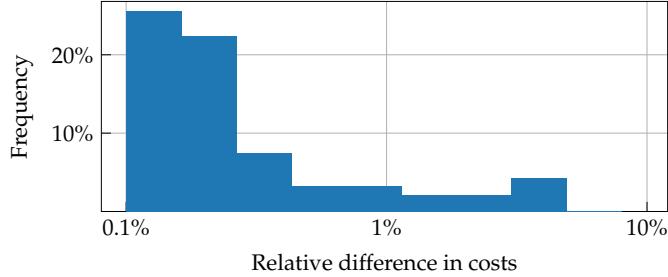


Figure 3.4: Histogram depicting the relative frequencies of cost differences obtained between MPCC and the other approaches in the 48-node 118-bus system.

tently produces larger cost values in this system. However, the maximum difference of 6% falls within the range of real-world variations due to the dispatcher's practical decisions in line with the actual pressure-flow relationship [55].

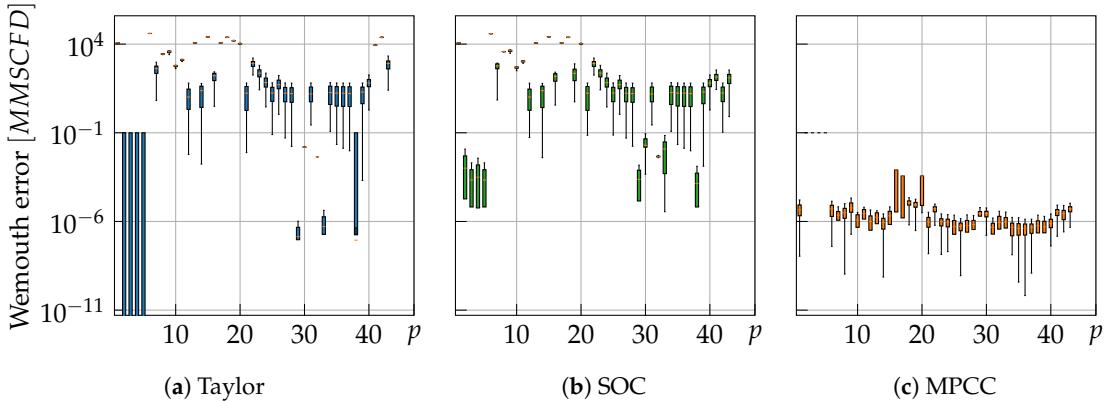


Figure 3.5: Weymouth approximation errors for each pipeline p reached by the contrasted approaches in the 118/48 study case.

Contrarily to cost function analysis, results in Figure 3.5 reveal a significant error reduction of about seven orders of magnitude (from 10^1 to 10^{-6}) under the proposed complementarity constraints. As an additional benefit, MPCC exhibits a shorter error dispersion than Taylor and SOC at most of the 46 pipelines in the network. Such behavior in the 118/48 system, also evidenced in the small 9/8 case study, proves the reliability of MPCC in effectively addressing more complex network configurations and interconnected dynamics.

3.3.3 Case Study III: 96/63 System

The last case study focuses on the Colombian power system, a complex network comprising 96 nodes ($|\mathcal{N}_P| = 96$), 49 generators ($|\mathcal{G}| = 49$), 207 transmission lines ($|\mathcal{F}| = 207$), and 80 power users ($|\mathcal{D}| = 80$). From the 49 generators, 10 are thermal power plants ($|\mathcal{I}| = 10$) fed by the natural gas transportation system, including 13 wells ($|\mathcal{W}| = 13$), 48 pipelines ($|\mathcal{P}| = 48$), 14 compressor stations ($|\mathcal{C}| = 14$), and 26 consuming users ($|\mathcal{U}| = 26$), yielding 63 nodes ($|\mathcal{N}_F| = 63$). Despite its radial structure, the gas system supports bidirectional flows in its pipelines due to the highly varying demand by thermal power plants influenced by meteorological conditions: On rainy seasons, thermal power plants dramatically reduce their demand; while on dry seasons, a large amount of gas must flow to them.

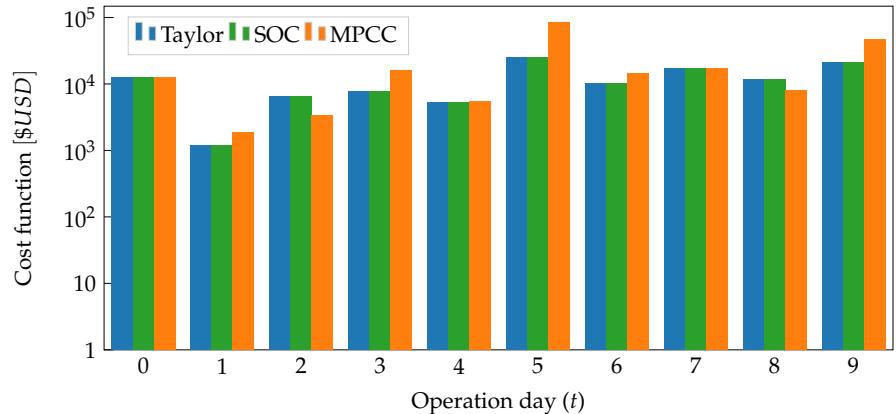


Figure 3.6: Daily operating cost obtained with each of the approaches in the 63-node 96-bus system.

Instead of estimating the distributions of the cost function and Weymouth error as in cases 9/8 and 118/48, the 96/63 case validates the Weymouth approximations in an operation case of ten consecutive days ($|\mathcal{T}| = 10$) with randomly changing gas extraction costs. Such a complementary validation strategy allows the interconnected system to reduce gas transportation costs by exploiting its single storage station, extending the performance analysis to scheduling scenarios. Figure 3.6 illustrates the daily optimized operating cost of the integrated system over the ten-day scheduling horizon for each tested approach. The daily cost values reveal notable similarities between the Taylor series and SOC relaxations. In average, the MPCC approach yields a 2.7% more expensive solution, from 8% cheaper to 12% more expensive, with a difference standard

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

deviation of 6%. The above results indicate that the difference between the proposed MPCC and baseline approximations is statistically negligible and will disappear after the empirical corrections.

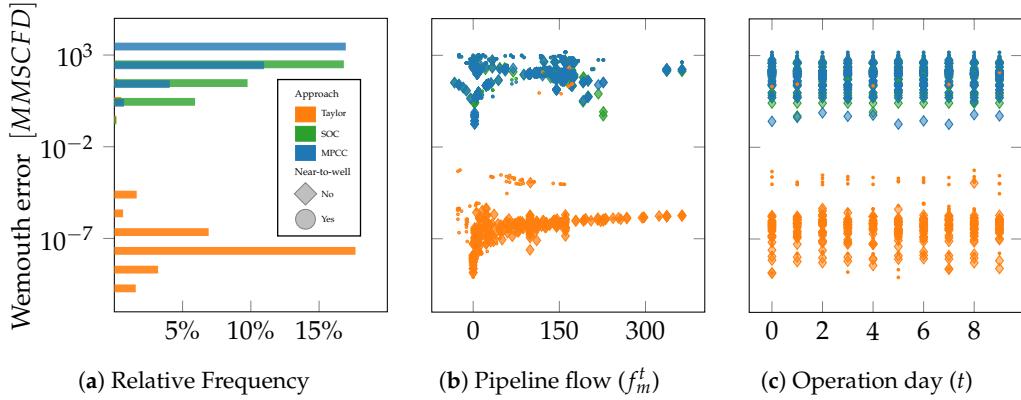


Figure 3.7: Weymouth error density on the Colombian case versus the gas flow and operation day.

Regarding the Weymouth approximation analysis, Figure 3.7 presents the error distribution and its relationship with the gas flow and the scheduled day for Taylor, SOC, and MPCC. Firstly, the error histogram in Figure 3.7a proves that the proposed MPCC formulation (in orange) exhibits superior approximation accuracy to Taylor and SOC for most pipelines and days. Secondly, the scatter plot in Figure 3.7b illustrates the relationship between Weymouth error and gas pipeline flow for each approach. Note that the benchmark techniques of Taylor (blue) and SOC (green) hold a steady error regardless of the flow rate. In the case of MPCC (orange), the larger the flow rate, the shorter the error dispersion. In addition, despite a larger error dispersion at low flow rates, MPCC still delivers much lower errors than benchmark methodologies. Hence, the complementarity constraints improve the error rates of Taylor and SOC and become more reliable for higher gas flow rates.

Lastly, Figure 3.7c suggests independence between the Weymouth error and each scheduled day, with a stationary error distribution for all approximations. Nonetheless, MPCC holds two groups of outlying errors. The higher ones align with typical magnitudes of the benchmark techniques. The second group of errors, lying around 10^{-2} , corresponds to pipelines connected to injection wells (denoted as dots in Figure 3.8c). Since the wells are technically regulated, their fixed injection pressure hampers the flexibility of MPCC for approximating the Weymouth equation.

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

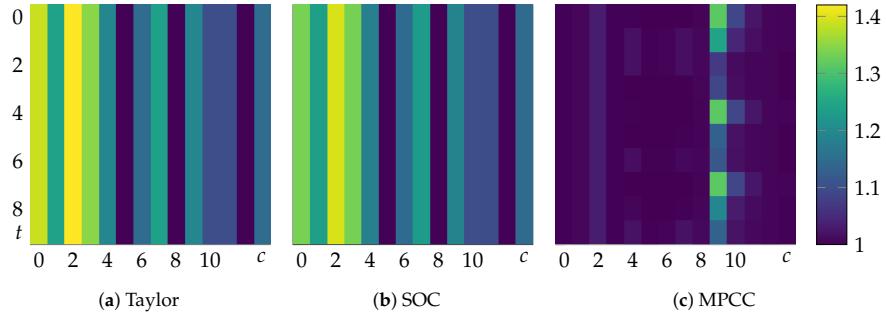


Figure 3.8: Output-to-input pressure ratio at the compressor in the 96/63 system.

The heatmaps in Figure 3.8 illustrate the output-to-input pressure ratio for each of the 14 compressors over the ten days of the scheduled operation. The baseline approaches of Taylor and SOC (Figure 3.8a,b) yield constant pressure ratios stemming from an over-relaxation of the Weymouth equation that extends the feasible region to unpractical solutions. In contrast, the MPCC approach in Figure 3.8c exhibits day-to-day pressure ratio changes within each compressor. The above is because the complementarity constraints closely align with the gas transport system's real physics, restraining the range of the feasible pressure values to trade off the daily varying injection cost.

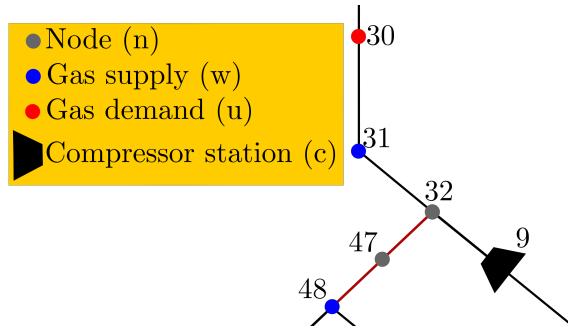


Figure 3.9: Outlying connection of well-compressor-pipeline on the system 96/63 used in Case Study III.

As a remark, compressor nine in Figure 3.8c reaches large pressure ratios on Days 0, 1, 2, 4, and 7, overlapping with the time instants with the highest approximation errors for MPCC in Figure 3.7c. A detailed examination of these outcomes detects that compressor nine and the outlying pipeline are the two outputs of a bifurcation, the latter

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

being followed by an injection well. Figure 3.9 exemplifies that such an interconnection is the sole over the gas network. As a hypothesis, fixing the pressure at the injection well and the flow direction at compressor nine pushes the complementarity constraints to the limits and forces the compressor to augment the pressure ratio to satisfy the forthcoming branch demand.

3.4 Discussion

This chapter presented a novel approximation for the Weymouth constraint by representing the nonconvex pressure–flow relationship as an MPCC. The MPCC-based formulation significantly benefits the optimization problems in interconnected power and gas systems by using continuous variables that emulate binary behavior to model flow direction, thus avoiding the need for computationally expensive discrete-variable formulations. Additionally, the MPCC inherently captures the complexity of the signum function, resulting in a rigorous approximation of the Weymouth equation.

The validation compared the proposed MPCC approach against the Taylor series and SOC programming approximations on optimizing the operation of interconnected power and gas transport systems. Monte Carlo experiments validated the solution reliability in two well-known case studies, while a ten-day operation planning assessed the scheduling task in the real-world case study of Colombian interconnected systems.

Regarding cost function, the MPCC approach demonstrated a remarkable ability to balance operational costs effectively. Results on the 9/8 system proved that MPCC converges to the exact cost of Taylor and SOC in small-scale cases. For more complex networks (cases 118/48 and 96/63), MPCC yields higher operational costs than baselines due to the more rigorous Weymouth equation modeling. Nonetheless, the cost differences among approaches lie within reasonable limits and align with the dispatcher’s empirical decisions.

In the case of Weymouth approximation, MPCC significantly outperforms Taylor and SOC in the tested cases. In the 118/48 and 96/63 systems, MPCC substantially reduces Weymouth approximation errors, often by several orders of magnitude, compared to traditional linearization and convex relaxation strategies. Such an accuracy improvement becomes crucial in large-scale, complex systems where precise approximation directly influences operational efficiency and system reliability. Hence, the introduced pressure–flow model mathematically benefits the optimization task, asserting its cost-effectiveness at various system scales.

The analysis of the scheduling task in the 96/63 Colombian interconnected system

CHAPTER 3. OPTIMIZATION USING MATHEMATICAL PROGRAMS WITH COMPLEMENTARITY CONSTRAINTS

highlights the robustness and reliability of the MPCC approach. Despite the complexities of bidirectional flows and time-varying demand scenarios, MPCC maintains high accuracy levels in Weymouth approximation. Furthermore, the nearly negligible cost differences among approximation approaches establish MPCC as the most robust and reliable approach for short-term operational scheduling.

Chapter 4

Stochastic Modeling of Natural Gas Flows Using Physics-Guided Neural Networks

This chapter aims to explore the integration of physical knowledge and stochastic modeling into graph-based neural network architectures to improve natural gas flow predictions under uncertainty. Building upon previous deterministic configurations, the proposed approach introduces physical constraints, such as nodal balances and the Weymouth equation, directly into the model's loss function to enhance prediction accuracy. Additionally, a stochastic framework is developed to quantify the uncertainty in both objective and decision variables related to gas system operations. Through a series of experiments, including applications to a simplified 8-node network and the Colombian national gas network, the chapter evaluates the trade-offs introduced by different loss function combinations. It demonstrates the model's ability to generalize under stochastic sampling, thereby advancing the development of a physics-guided and uncertainty-aware gas dispatch optimization strategy.

4.1 Methodology

Physics-Informed Neural Networks (PINNs) represent a class of neural networks where physical laws are incorporated into the learning process, guiding the model to respect these constraints. Unlike traditional neural networks, where the loss function is typically based on the discrepancy between predicted and actual data, PINNs introduce

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

additional terms in the loss function that penalize the model for deviating from known physical principles.

In this case, the physical constraints are derived from the gas balance and the Weymouth equations, which describe the flow and pressure behavior within the gas transportation network. These constraints are integrated into the neural network model introduced in Chapter 2 as additional loss terms. Specifically, we define two layers within the network: one that calculates the error in gas balance and another that calculates the error in the Weymouth equation. The outputs of these layers are then used to adjust the network's predictions, ensuring that they adhere to the physical laws governing the system.

The inclusion of these physics-informed layers allows the network to achieve better generalization, as it is not only trained on the data but also guided by the underlying physical laws. This approach can be seen as a specialized form of regularization, where the model is penalized if its predictions do not satisfy the physical constraints. The overall loss function can be expressed as:

$$\mathcal{J}(\Theta) = \mathcal{J}_{\text{data}}(\Theta) + \mathcal{J}_{\text{balance}}(\Theta) + \mathcal{J}_{\text{weymouth}}(\Theta), \quad (4.1)$$

with $\mathcal{J}_{\text{data}}(\Theta)$ representing the traditional data-driven loss, $\mathcal{J}_{\text{balance}}(\Theta)$ is the loss associated with the gas balance constraint, and $\mathcal{J}_{\text{weymouth}}(\Theta)$ is the loss associated with the Weymouth equation constraint.

The gas balance constraint is enforced at each node through the following loss function:

$$\mathcal{J}_{\text{balance}} = \mathbf{T} \cdot \hat{\mathbf{f}}_e - \mathbf{d} + \hat{\mathbf{f}}_n. \quad (4.2)$$

In here, $\hat{\mathbf{f}}_e \in \mathbb{R}^{N_e}$ is the predicted edge flow, $\hat{\mathbf{f}}_n \in \mathbb{R}^{N_v}$ is the predicted nodal injection, $\mathbf{d} \in \mathbb{R}^{N_v}$ is the actual demand, and $\mathbf{T} \in \mathbb{R}^{N_e \times N_v}$ is the signed incidence matrix.

The Weymouth constraint relates the gas flow to the difference of squared pressures between connected nodes through the following this loss function:

$$\mathcal{J}_{\text{weymouth}} = \mathbf{M}_{\mathcal{P}} \left(\hat{\mathbf{f}}_e^{\circ 2} - \mathbf{K} \circ (\mathbf{T} \cdot \hat{\boldsymbol{\pi}}^{\circ 2}) \right), \quad (4.3)$$

begin $\hat{\boldsymbol{\pi}} \in \mathbb{R}^{N_v}$ the predicted pressure at each node, $\hat{\mathbf{f}}_e^{\circ 2}$ is the element-wise square of edge flows, , the vector $\mathbf{K} \in \mathbb{R}^{N_e}$ contains the Weymouth constants for each edge and is zero for compressors, while $\mathbf{M}_{\mathcal{P}} \in \mathbb{R}^{N_e \times N_e}$ is a diagonal matrix that selects only the pipeline edges $p \in \mathcal{P}$. Together, these physics-informed components guide the neural

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

network toward solutions that not only fit the data but also obey the operational and physical laws of gas transport.

For practical implementation, we build upon the experimental setup outlined in Section 2.2, maintaining the same general approach while incorporating new elements that account for the physics of the natural gas system represented in Equations (4.2) and (4.3) following the scheme shown in Figure 4.1. The samples are generated using the nonlinear natural gas network optimization model from Chapter 3. In this process, a power-interconnected system was considered; however, since this study focuses on the gas system, the power system remained constant without variation. As in the previous setup, noise was introduced into the base values of two gas networks: a small-scale test network of 8 nodes and the more extensive Colombian natural gas transportation system. The noise levels, ranging from 5% to 25%, simulate various operating conditions, providing diverse training data.

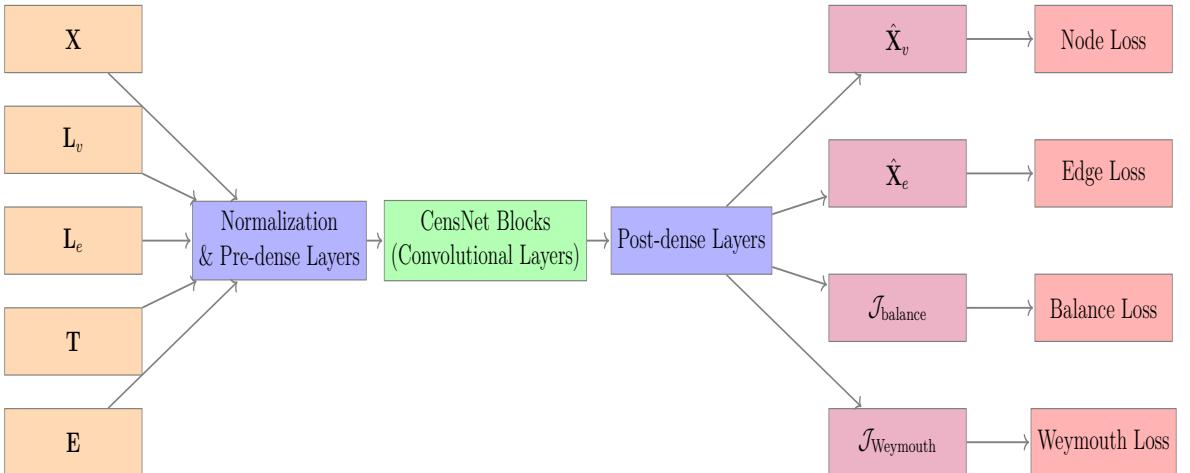


Figure 4.1: General outline of the CensNet-based model with physics guidance.

4.2 Results

In this section, we present the results of the proposed model, which now incorporates physical constraints from the natural gas system. The focus remains on the relationship between the predicted outputs and the actual observed values, evaluating the model's performance across the 8-node test network and the Colombian natural gas

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

transportation system. By incorporating physics-based constraints, the goal is to assess the model's ability to predict critical parameters under various operational conditions while ensuring that the physical laws governing gas flow are respected.

4.2.1 Case Study I: 8-node Network

In this chapter, we begin with experiments that account for both node and edge losses, as it was found that considering only the node loss did not produce adequate results. The best parameters identified for this experiment were N channels = 25, N layers = 4, and N dense = 11. These settings yielded a total loss of 6.8, with a node loss of 2.8 and an edge loss of 4.02.

The results corresponding to the nodes, shown in Figure 4.2a, exhibit a similar behavior to that observed in Figure 2.6a, demonstrating that the model accurately captures the injection pattern at the nodes. The correlation between the actual and predicted values is also strong, as indicated by an R^2 of 0.983. In turn, edge flows show some variation, as seen in fig. 4.2b, mainly when predicting the flows through the first pipeline connected to the injection field, where slight deviations from the actual flow values were observed in the blue dots. However, the model performed well overall, achieving an R^2 of 0.983 for the edge flows. While the first pipeline presents some prediction challenges, the accuracy in predicting flows across the remaining the pipelines remains high, demonstrating the model's ability to handle the complexity of gas transportation in this nonlinear system.

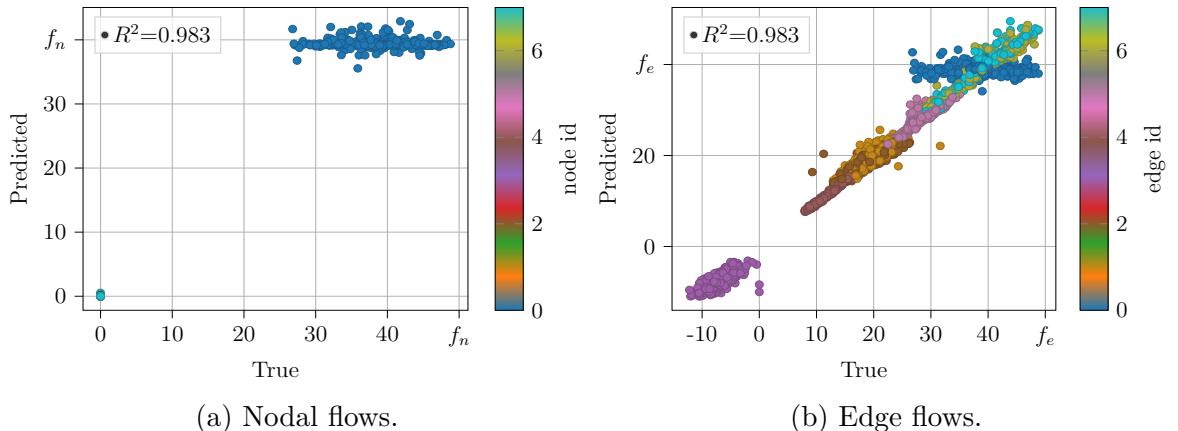


Figure 4.2: Model results using the losses associated with nodal and edge flow predictions in the 8-node network.

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

The second part of this experiment involves the additional loss associated with the gas balance, building upon the previous setup that considered both node and edge losses. The hyperparameter optimization yielded the best parameters: $N \text{ channels} = 61$, $N \text{ layers} = 2$, and $N \text{ dense} = 2$. These settings resulted in a total loss of 10.041, with a node loss of 2.9, an edge loss of 6.4, and a balance loss of 0.78. The prediction behavior at the nodes, as shown in Figure 4.3a, remained consistent with the results obtained in the previous experiment, where the balance loss was not included. The model accurately captured the gas injection pattern, with an R^2 of 0.983 for node flow predictions, identical to the earlier case. Similarly, the prediction of edge flows, shown in Figure 4.3b, followed the same general trend as before, although a slight decrease in accuracy was observed, reflected by an R^2 of 0.973. While this represents a minor reduction in performance compared to the previous experiment, the model still demonstrated a strong ability to predict gas flows through the edges, maintaining a high level of accuracy.

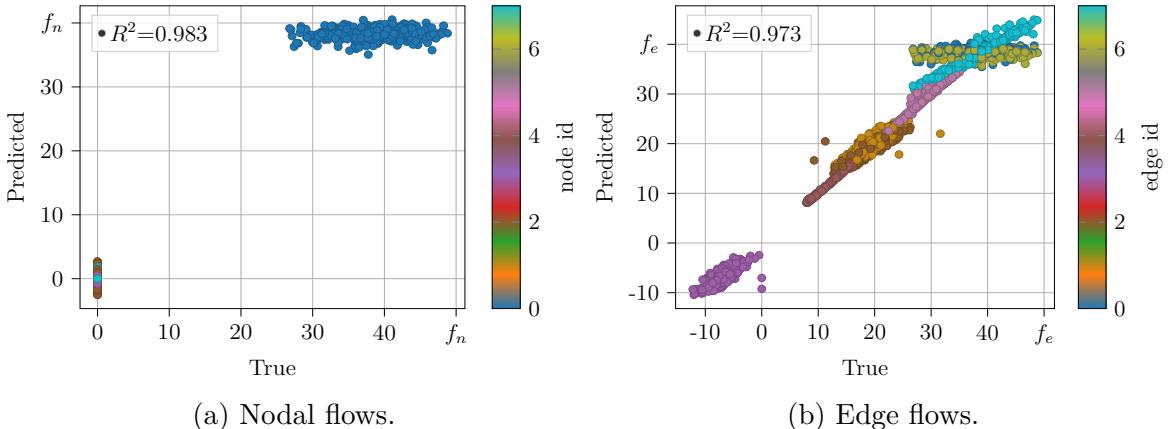


Figure 4.3: Model results using the losses associated with nodal and edge flow predictions along with the gas balance loss in the 8-node network.

In the following part of this experiment, we incorporated losses associated with node and edge flows, the gas balance, and the Weymouth equation. The hyperparameter optimization for this setup yielded the following best parameters: $N \text{ channels} = 17$, $N \text{ layers} = 1$, and $N \text{ dense} = 4$. These settings resulted in a total loss of 20.670, with the individual losses being a node loss of 3, an edge loss of 11.35, a balance loss of 2.72, and a Weymouth equation loss of 3.59. As shown in Figure 4.4a, the behavior of the node flow predictions remained consistent with the previous experiments, with an R^2

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

of 0.983. The model continued to accurately capture the gas injection patterns at the nodes. However, the prediction accuracy for edge flows showed a notable deterioration, as seen in Figure 4.4b. The R^2 value for edge flow predictions dropped to 0.952. This decrease in performance is primarily due to the difficulties encountered in predicting flows along edges 1, 2, 6, and 7. Edges 1 and 2 correspond to pipelines that are part of a closed path in the network, while edges 6 and 7 correspond to compressors. These complexities in the network configuration likely contributed to the reduction in predictive accuracy for these specific edges.

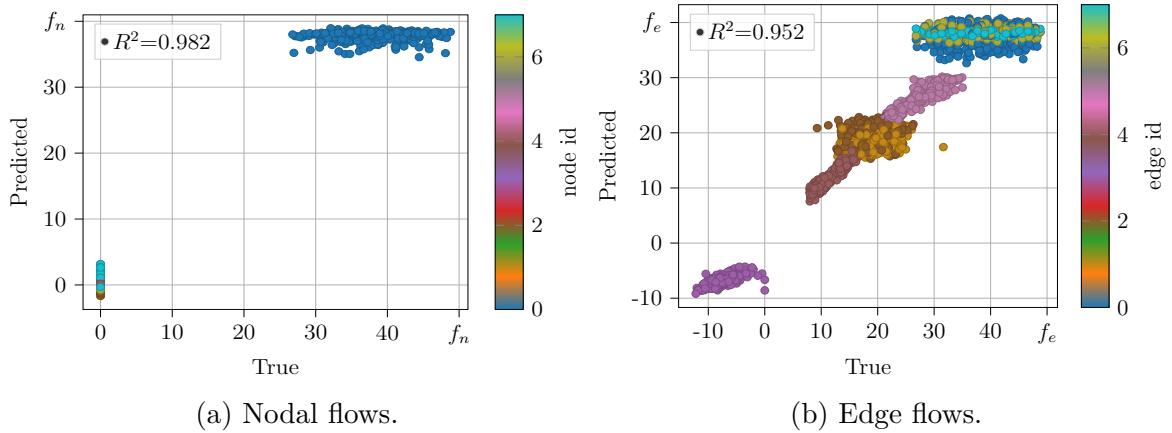


Figure 4.4: Model results using the losses associated with nodal and edge flow predictions along with the gas balance and Weymouth losses in the 8-node network.

In the subsequent experiment, the losses associated with node flows and the physical equations—namely, the gas balance and the Weymouth equation—were considered. The hyperparameter optimization process resulted in the best parameters being $N \text{ channels} = 18$, $N \text{ layers} = 1$, and $N \text{ dense} = 5$. These settings led to a total loss of 10.270, with a node loss of 3.976, a balance loss of 4.75, and a Weymouth equation loss of 1.55. The prediction at the nodes, shown in Figure 4.5a, remained largely consistent with previous experiments, though there was a slight decrease in accuracy, with the R^2 value dropping to 0.976. This minor reduction indicates that the model continues to perform well in predicting gas injection patterns at the nodes. However, the prediction accuracy for edge flows, as seen in fig. 4.5b, experienced another decline. The R^2 value dropped to 0.899, reflecting increased difficulties in predicting flows through the compressors and the pipeline connected to the injection field.

The results in Table 4.1 present the differences between the IPOPT benchmark and

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

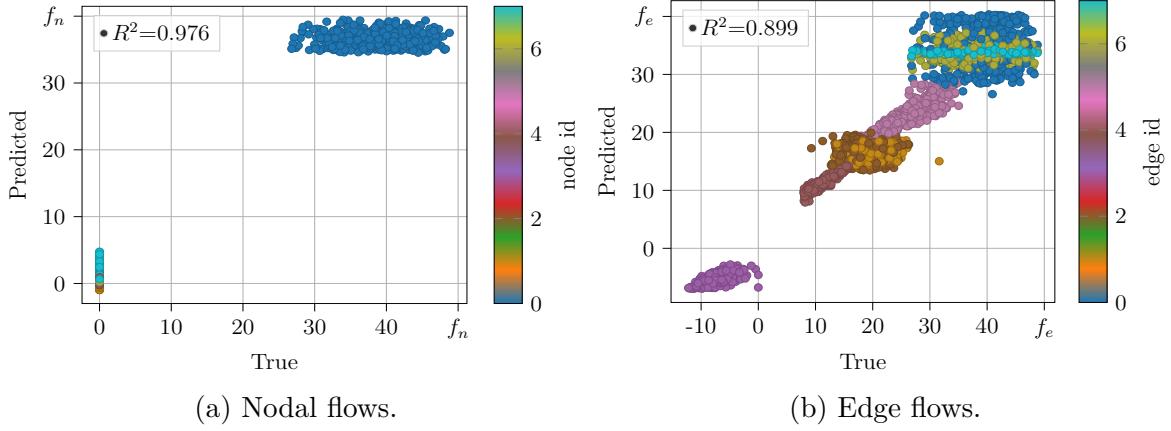


Figure 4.5: Model results using the loss associated with nodal flow predictions along with the gas balance and Weymouth losses in the 8-node network.

the CensNet-based models trained with different combinations of loss functions. The model trained with only the nodal loss (CensNet (N)) achieves nearly identical nodal flow predictions compared to the benchmark, but shows a large discrepancy in edge flow values. This confirms that training without edge-level supervision leads to significant errors in edge behavior, despite accurate nodal outputs.

When the edge loss is added (CensNet (N+E)), the model achieves a substantial reduction in edge flow error, improving from 22.76 to 0.22. This improvement is accompanied by a decrease in balance error, suggesting that learning edge-level patterns enhances global consistency across the network. Adding the balance loss (CensNet (N+E+B)) leads to further improvements in balance prediction, reducing the error to just -0.028 with small changes in nodal and edge differences, confirming that this constraint contributes to overall physical consistency.

The inclusion of the Weymouth loss (CensNet (N+E+W)) increases the edge flow error slightly and leads to a modest rise in balance error. However, the nodal prediction remains accurate, showing that incorporating physical knowledge through the Weymouth constraint introduces trade-offs without severely degrading model accuracy. Finally, combining all losses (CensNet (N+E+B+W)) yields a balanced model with low errors across all metrics and maintains high consistency with the benchmark.

All CensNet-based models significantly outperform the IPOPT method in terms of inference time, offering over an order of magnitude faster predictions with minimal variability. Among the tested configurations, **CensNet (N+E+B)** achieves the best

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

Method	Node Error	Edge Error	Balance Error	Time
CensNet(N)	0 ± 17.99	22.76 ± 15.43	-0.01 ± 17.45	0.86 ± 0.5
CensNet(N+E)	-0.11 ± 18.27	0.22 ± 21.65	-0.12 ± 1.7	0.85 ± 0.5
CensNet(N+E+B)	-0.02 ± 18	-0.02 ± 21.3	-0.03 ± 0.9	0.85 ± 0.5
CensNet(N+E+W)	-0.07 ± 17.56	2.6 ± 20.03	-0.070 ± 2.21	0.86 ± 0.5
CensNet(N+E+B+W)	0.05 ± 17.91	0.25 ± 21.14	0.05 ± 1.69	0.85 ± 0.5

Table 4.1: Comparison of the differences in mean and standard deviation values for nodal flows (Node Error), edge flows (Edge Error), nodal balance (Balance Error), and prediction time (Time) between the optimization-based benchmark (IPOPT) and the CensNet-based models. Each value reflects the difference between IPOPT and the respective model.

overall performance, with the smallest errors in edge flow and balance, while preserving nodal accuracy and computational efficiency.

A stochastic analysis was conducted to assess the robustness of the trained CensNet-based model and its ability to generalize under uncertainty. This analysis evaluates the model’s consistency and robustness within the distribution of the training data, particularly in generating outputs from synthetic input samples, thereby quantifying its reliability within this learned space. Initially, a kernel density estimate (KDE) was fitted to the input training data used for the CensNet model. From this estimated distribution, a set of synthetic input samples, denoted as X_{sample} , was generated. These synthetic inputs were then propagated forward through the trained network, taking into account all loss components, to yield a corresponding set of output predictions, denoted as y_{sample} . In parallel, the original training inputs were propagated through the network to obtain output predictions, denoted as \bar{y}_{train} . A second KDE was then fitted using the training output data (hereafter referred to as y_{train}). Two log-likelihoods were calculated based on this second KDE: one using y_{sample} and the other using \bar{y}_{train} . The log-likelihood for the synthetic outputs, y_{sample} , was found to be $-6,696,247.56$, while the log-likelihood for the training outputs, \bar{y}_{train} , was $-6,657,534.62$. The closeness of these two values suggests that the synthetic outputs generated from the input KDE closely resemble the distribution of the training outputs.

To evaluate the goodness-of-fit between the distributions of y_{sample} and \bar{y}_{train} , a Kolmogorov–Smirnov (K–S) test was conducted using three alternatives: two-sided, less, and more significant. The test results were as follows: a two-sided test yielded a statistic of 0.01833 with a p-value of 0.81482 (statistic location -0.34018 ; the ‘less’

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

alternative produced a statistic of 0.01167 with a p-value of 0.72137 (statistic location 0.30738; and the 'greater' alternative returned a statistic of 0.01833 with a p-value of 0.44640 (statistic location -0.34018). These statistical measures provide an indication that the synthetic outputs generated via the KDE-based sampling are consistent with the distribution of outputs observed during training, demonstrating the model's ability to maintain consistency under uncertainty within the learned data distribution. After completing the stochastic analysis of the variables associated with the optimization model, the next step involves leveraging the KDE fitted to the training data. This fitted KDE was used to generate a new set of input samples that represent the learned probability distribution of the optimization model's inputs. These samples were then forward-propagated through the trained neural network model.

Following this analysis, we next study the uncertainty inherent in the network by approximating the joint probability density functions (PDFs) for the input variables. For training, each node was assigned five inputs, corresponding to the lower limit of gas injection capacity, the upper limit of gas injection capacity, the demanded flow, the lower limit of pressure, and the upper limit of pressure. Since the lower limits are permanently fixed at zero, these two variables were excluded from the analysis. Although the training data were initially sampled from a uniform distribution, only the successfully converged samples during the optimization process were retained. Consequently, the resulting PDFs do not exhibit a uniform distribution; instead, they display a variety of distribution shapes.

Figure 4.6 presents the first group of joint distributions identified in this study, characterized by the presence of two distinct modes, which likely correspond to the two main operating regimes of the system. This group includes joint patterns involving the injection node (N_0), the demand nodes (N_6 and N_7), and several intermediate or pass-through nodes (N_2 , N_4 , and N_5). The observed relationships reflect the physical and topological structure of the network, particularly the connections among injection, compression, and demand points. A relationship is observed between the upper-pressure limits at nodes N_2 and N_4 , which aligns with their placement within a sequential compression segment of the network. Node N_2 serves as the output of the first compressor and the input to the second, while node N_4 lies on a closed-loop path downstream of this compression chain. The correlation between their pressure bounds likely reflects how the viable operating range at one point in the compression sequence affects neighboring nodes within the loop. Similarly, the pressure limit correlation between node N_5 , also part of the closed path, and the downstream demand node N_6 , suggests that feasible pressure ranges propagate along connected elements of the network, especially

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

where compression and demand interact. A subgroup of distributions reveals strong associations between the maximum demand at node N_7 (\bar{f}_u) and upstream variables. Specifically, N_7 's maximum flow shows dependence on the maximum injection flow at node N_0 (\bar{f}_w), and on the maximum pressures at nodes N_0 , N_1 (compressor 1 input), and N_5 . These patterns suggest that the flow capacity at the far end of the system (N_7) is conditioned by both supply and compression states at the upstream nodes. Also, a direct relationship is identified between the pressure limits of node N_7 and node N_4 . Given their physical connection, this correlation likely reflects the way pressure constraints propagate between adjacent sections of the network.

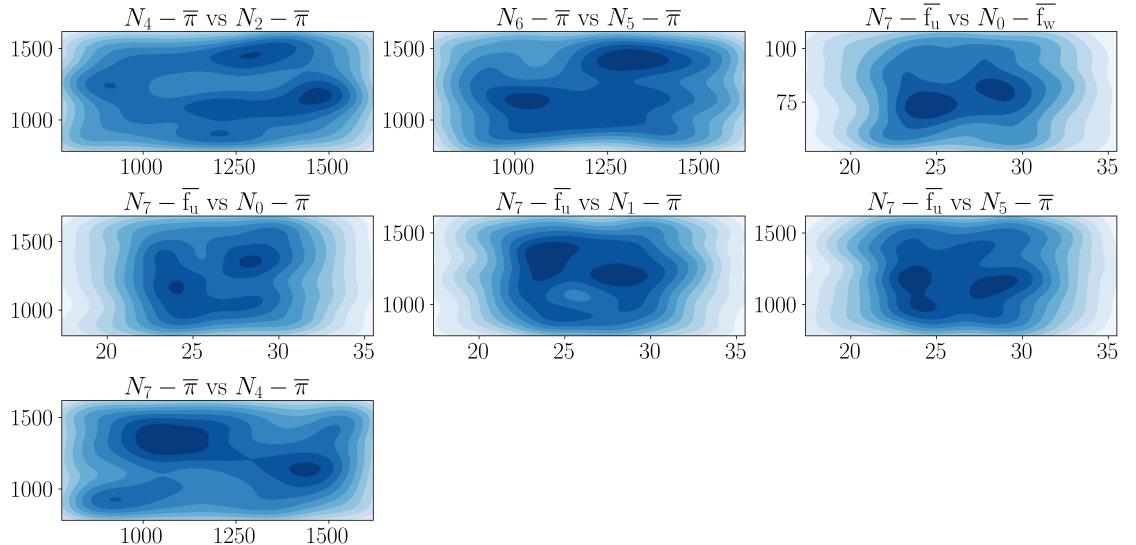


Figure 4.6: Bimodal joint PDFs identified in the training inputs of the 8-node gas network.

Using the network results obtained from the KDE fit, a new joint analysis was conducted, analogous to the previous one performed with the variables of the optimization model. Following the same structure presented in the first part of this analysis, we begin by highlighting the relationships that exhibit two distinct modes of operation according to their joint probability density functions. In this case, as shown in Figure fig. 4.7, only two variable pairs present this bimodal behavior when the analysis is performed using the inputs propagated through the trained neural network. These pairs are: the nodal pressure limits of N_4 and N_2 , and those of N_6 and N_5 . In both cases, the joint PDFs

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

suggest the existence of two operational regimes, potentially associated with alternative configurations or routing strategies within the network that are learned by the neural model.

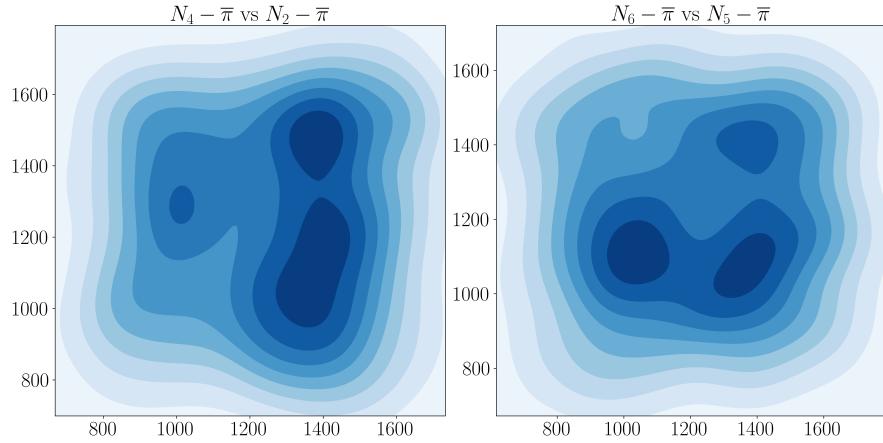


Figure 4.7: Bimodal joint PDFs from the KDE-sampled input data propagated through the CensNet model for the 8-node network.

Figure 4.8 presents a second group of joint distributions identified in the study, which are characterized by low probability concentrations at their extreme values and generally diffuse patterns with small high-probability regions. This behavior, visually represented by lighter tones across most of the joint probability density functions and occasional darker spots, suggests a weaker or more dispersed correlation structure between the involved variables. Several of these relationships involve the injection node N_0 and its connection to both upstream and downstream nodes in terms of maximum pressure ($\bar{\pi}$) and maximum injection flow (\bar{f}_w). For instance, the joint distributions between the maximum pressure at N_2 and N_0 , as well as between the maximum pressures at N_3 , N_4 , and N_6 with that of N_0 , reveal broadly scattered patterns with localized high-probability points. This indicates that while these variables are structurally related, their extreme values do not consistently co-occur. A similar pattern is observed in the distributions involving the maximum pressure at N_3 and the injection capacity at N_0 . Despite the proximity of these nodes within the system's trajectory, the weak concentration suggests a more flexible operational regime or buffering effect along the path. Further examples include the relationships between N_5 and both N_0 and N_2 , where again, the lack of sharp and extensive density at the extremes hints at variability in how pressure is maintained

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

or transmitted through these nodes. Lastly, the distributions between the maximum pressure at N_6 and both N_0 and N_3 continue this trend, reinforcing the understanding that, in these cases, the system does not exhibit a strong link in the upper operational ranges of the involved variables.

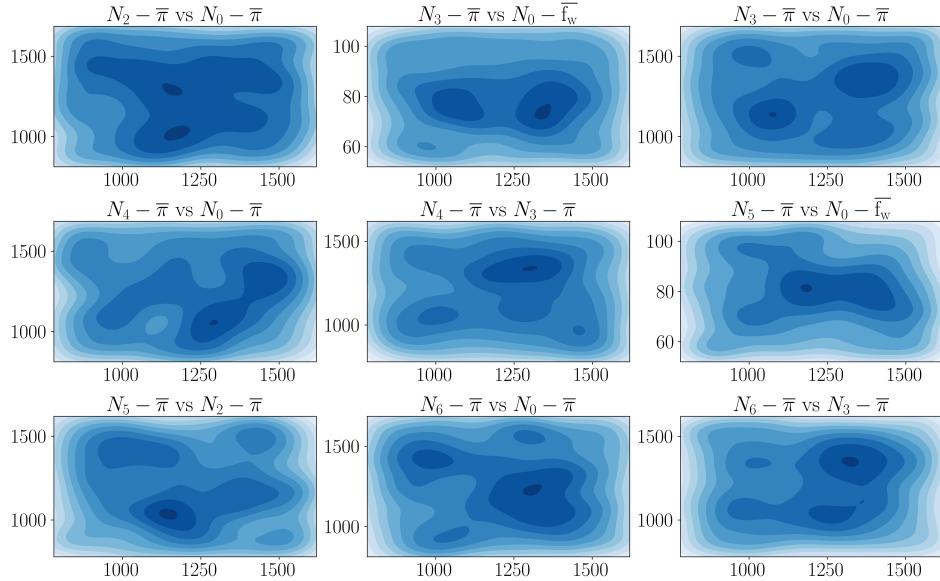


Figure 4.8: Joint PDFs between the inputs used in the 8-node network with a high probability zone concentrated in a small region.

Continuing with the distributions that exhibit a small high-probability region (Figure 4.9), three pairs of variables were identified in this case whose joint PDFs display this characteristic. These pairs are: The Upper injection limit and the nodal pressure limits between N_0 and N_1 ; the pressure limits between N_1 and N_4 ; and the gas flow demand in N_6 and the nodal pressure limit in N_4 . These distributions are characterized by a concentrated and sharply defined region of high probability, suggesting that the neural network tends to favor very specific combinations of input values during prediction.

Figure 4.10 presents a third group of joint distributions distinguished by the presence of large dark zones in the joint probability density functions, indicating regions of high probability concentration. These distributions suggest strong dependencies between the involved variables. A clear example is the distribution involving the maximum

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

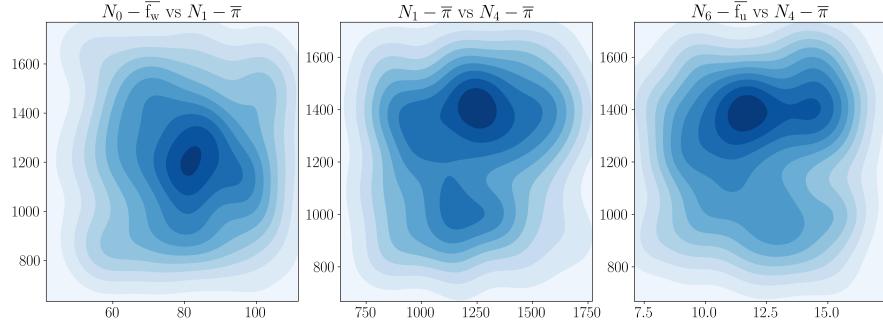


Figure 4.9: Joint PDFs between the inputs used in the 8-node network with a high probability zone concentrated in a small region in the KDE sampled variables.

injection flow at node N_0 and the maximum pressure at node N_1 , which corresponds to the input of the first compressor. The high-probability region observed here suggests a consistent relationship between the capacity to inject gas into the network and the pressure conditions at the compressor's intake, possibly reflecting operational constraints or control strategies that maintain a stable relationship between these variables. Likewise, the joint distribution between the maximum pressure at node N_1 and the pressure at node N_4 (a pass-through node connected downstream) also shows a dense concentration of probability. This implies a strong coupling between the input conditions of the first compressor and the downstream pressure, likely due to the system's physical configuration and flow continuity. Lastly, a similar pattern is observed in the joint distribution between the maximum demand at node N_6 and the pressure at node N_4 . The high-probability zone here may indicate a regulatory relationship, where variations in demand at N_6 are closely associated with pressure conditions at N_4 , possibly due to their direct physical connection and the influence of demand on local pressure levels. It is worth noting that, unlike in this case, no distributions with large, clearly defined areas of high probability (i.e., darker and more extensive areas) were observed in the joint PDFs obtained from the adjusted KDE.

To continue with the stochastic analysis, here we discuss the joint probability density functions (PDFs) that exhibit noteworthy behaviors, obtained from the output variables of the optimization model when it was supplied with the various scenarios generated through the previously described sampling process. It is important to note that, at this stage, each output variable was normalized with respect to the maximum capacity used in the corresponding scenario. That is, the values of the variables shown in this section

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

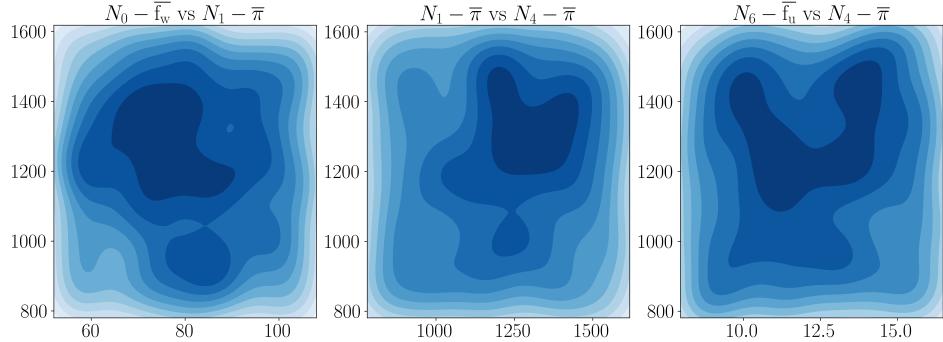


Figure 4.10: Joint PDFs between the inputs used in the 8-node network, which have a considerably wide mode with respect to the entire distribution.

are expressed relative to the maximum value that the variable could reach under that specific configuration.

Figure 4.11 displays a set of joint probability density functions in which the flow through pipeline p_3 , the component that closes the loop in the network, takes on negative values across all scenarios. These values typically fall within 10% to 30% of the normalized maximum, and are consistently observed along the X-axis of each subplot. The Y-axis in these figures corresponds to various system outputs, including normalized gas injection flow from the injection field, nodal pressures, and flows through other pipelines in the network. The observed negative values can be attributed to the assumed direction of flow in the optimization model, which turned out to be opposite to the one required for optimal operation. The model resolves this discrepancy by assigning negative values to the flow in p_3 . Although this pipeline is not necessary for system operability, it plays a role in cost minimization, which explains its limited but consistent utilization.

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

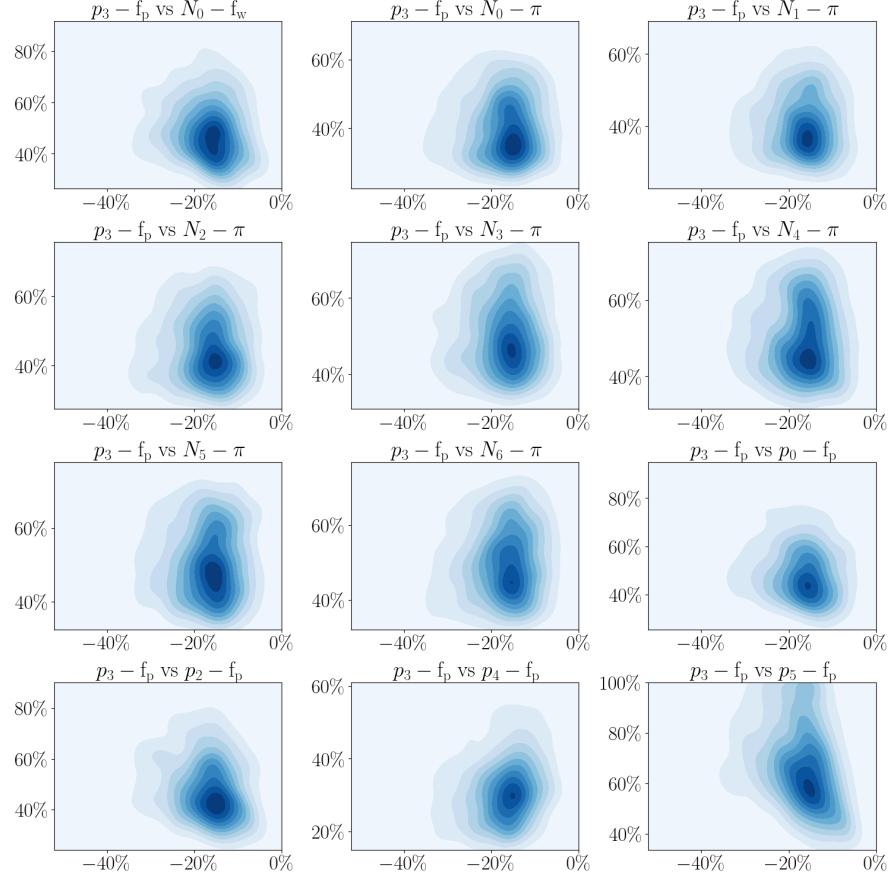


Figure 4.11: Joint PDFs between the outputs used in the 8-node network, which have a negative flow.

Figure 4.12 presents a second set of joint PDFs based on KDE-sampled data, which reinforces the patterns discussed previously. Again, the flow through pipeline p_3 (X-axis) consistently takes on negative values, indicating a reversal in the direction assumed during model formulation. This trend reflects the network's ability to adapt under optimal conditions and highlights the model's alignment with physical and economic system constraints. The robustness of this behavior is supported by a range of Y-axis variables, including the total injection at node N_0 and the flows through pipelines p_0 , p_1 , p_2 , p_4 , and p_5 , which show relationships with the reversed flow in p_3 .

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

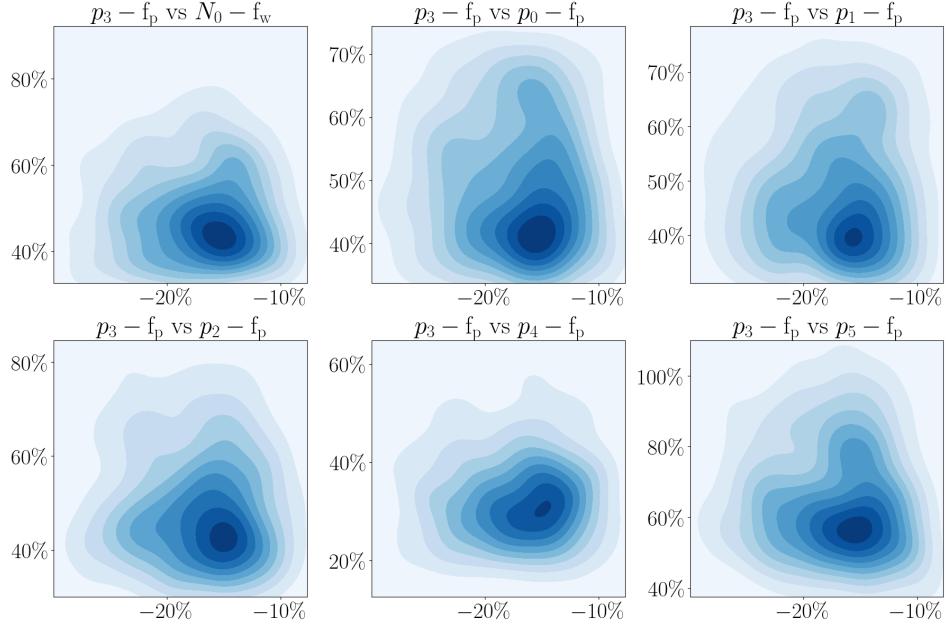


Figure 4.12: Joint PDFs between the outputs used in the 8-node network, which have a negative flow in the KDE sampled variables.

Figure 4.13 presents the joint distributions that exhibit a bivariate gamma distribution-like shape. In this case, the X-axis corresponds to the normalized flow transported through pipeline p_1 , which connects nodes N_3 and N_4 and forms part of the system's closed-loop trajectory. The Y-axis in each of the four subplots represents the normalized nodal pressure at nodes N_3 , N_4 , N_6 , and N_7 , respectively. Nodes N_3 and N_4 are part of a compact subsystem of pass-through nodes that define the core loop of the network. N_6 and N_7 , while functioning as demand nodes, are connected downstream of this region, with N_7 specifically being supplied via node N_4 . This spatial arrangement suggests that the observed distributions highlight the interdependence between the transported flow through this central pipeline and the pressure conditions at nearby or downstream nodes. In particular, although N_6 is slightly more distant, its role as a demand node may explain its inclusion in this group of correlated behaviors.

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

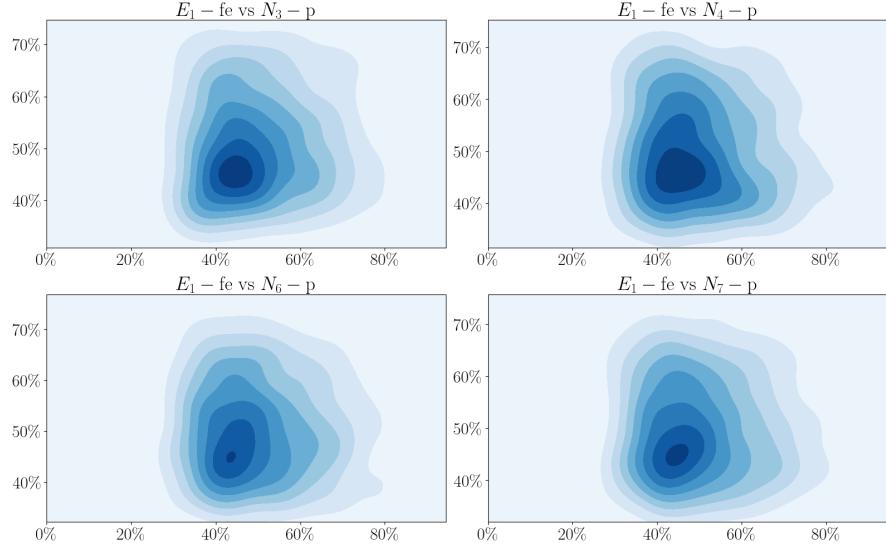


Figure 4.13: Joint PDFs between the outputs used in the 8-node network have a behavior quite similar to that of a bivariate gamma distribution.

Figure 4.14 displays a third group of joint probability density functions that exhibit another pattern. In these distributions, both variables tend to take values within a range—approximately between 25% and 70% of their respective maximum capacities. Despite this wide span, the distributions do not resemble uniform patterns. Instead, they appear to exhibit a clear concentration of higher probability in a more restricted interval, specifically when both variables fall between 30% and 40% of their normalized limits. This behavior is particularly evident in the joint PDFs that relate nodal pressures across different yet structurally connected regions of the system. One set of these distributions involves the pressure at the injection node (N_0) and the pressures observed at downstream nodes, including the compressor-connected nodes (N_1 , N_2 , and N_3), as well as the nodes forming the closed loop (N_3 , N_4 , and N_5) and the final demand nodes (N_6 and N_7). These distributions suggest a degree of coordination in the pressure values that extends from the injection point through the compression stages and into the rest of the system. The tendency of these variables to cluster within a narrower operational range may indicate that the system's feasible configurations favor pressure levels that are not too close to the lower or upper bounds, but rather stabilized in a middle region that ensures both efficiency and reliability. This is especially relevant in subsystems like the closed-loop segment formed by nodes N_3 , N_4 , and N_5 , which play a critical role

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

in redistributing flow toward the demand points. The same applies to the pressures at the terminal load nodes (N_6 and N_7), which depend on adequate upstream conditions to meet demand.

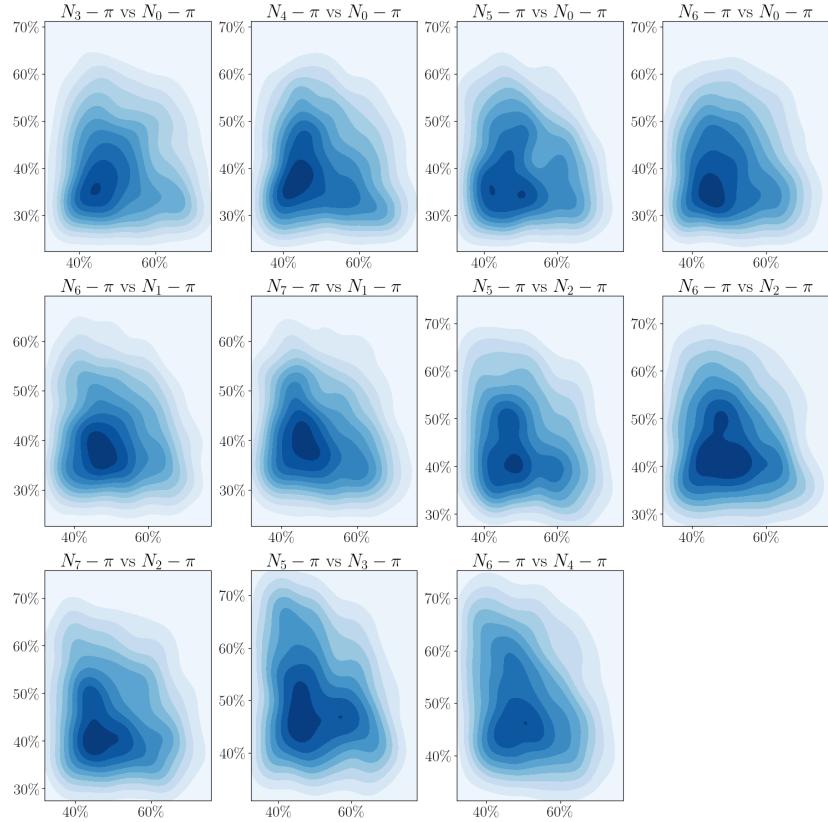


Figure 4.14: Joint probability density functions (PDFs) between nodal pressure outputs in the 8-node network, highlighting a concentration of values within mid-range operating levels (30%–40% of normalized limits)

It is also relevant to examine the joint probability distributions between input and output variables (Figure 4.15), particularly those in which the input variables exhibit significantly broader dispersion compared to the outputs. This behavior can be observed in a group of joint PDFs where the outputs—specifically, the flow through pipeline p_4 —tend to remain concentrated within the 20% to 40% range of their normalized capacities, while the corresponding input variables vary much more widely. It is im-

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

portant to highlight that, in this case, normalization was only applied to the output variables. The inputs, being defined as upper bounds or operational limits (such as the maximum allowable injection flow or pressure limits), were not normalized. The most illustrative examples of this pattern involve joint distributions between the upper bound of the injection flow at node N_0 and the flow in p_4 , as well as between the upper nodal pressure bounds at nodes N_0 through N_7 and the same pipeline output. Pipeline p_4 corresponds to the final segment delivering gas to one of the system's demand nodes. The limited range of variation in this output, despite the broad dispersion in input conditions, suggests that the optimizer consistently resolves the system by prioritizing stable and relatively low utilization of this pipeline.

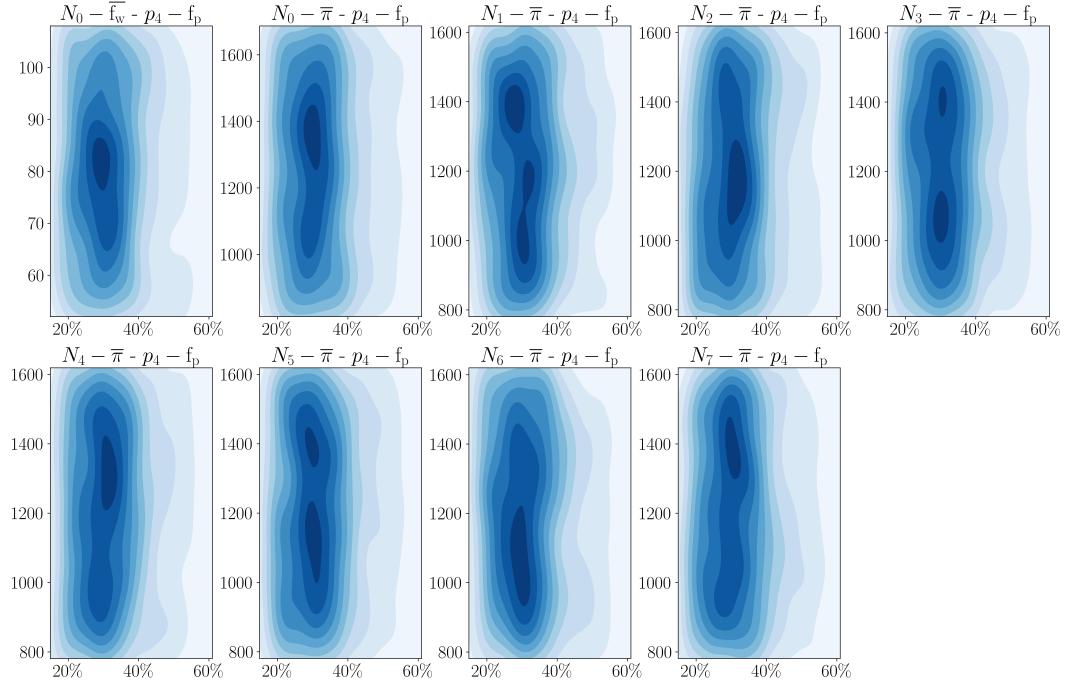


Figure 4.15: Joint PDFs between the inputs and outputs used in the 8-node network, which present a wide dispersion along second variable.

As shown in Figure 4.16, there is a group of joint probability distributions where the input variables span a wide range of values, while the corresponding output, specifically, the flow through pipeline p_4 , remains narrowly concentrated between 20% and 40% of

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

its normalized capacity. This behavior mirrors what was previously discussed, suggesting that, even under the learned data distribution, the network consistently predicts a relatively low and stable utilization of this particular pipeline. The pairs exhibiting this characteristic include the upper bound on injection flow at node N_0 and the upper bounds on nodal pressures at nodes N_0 through N_7 , all compared against the predicted flow through pipeline p_4 . As in the previous case, normalization was only applied to the output variables, while the inputs were left in their original scale. The similarity in joint distribution patterns between the optimization model and the neural network predictions indicates that the trained model effectively internalized a similar resolution strategy, favoring stable usage of pipeline p_4 regardless of the variability in input conditions.

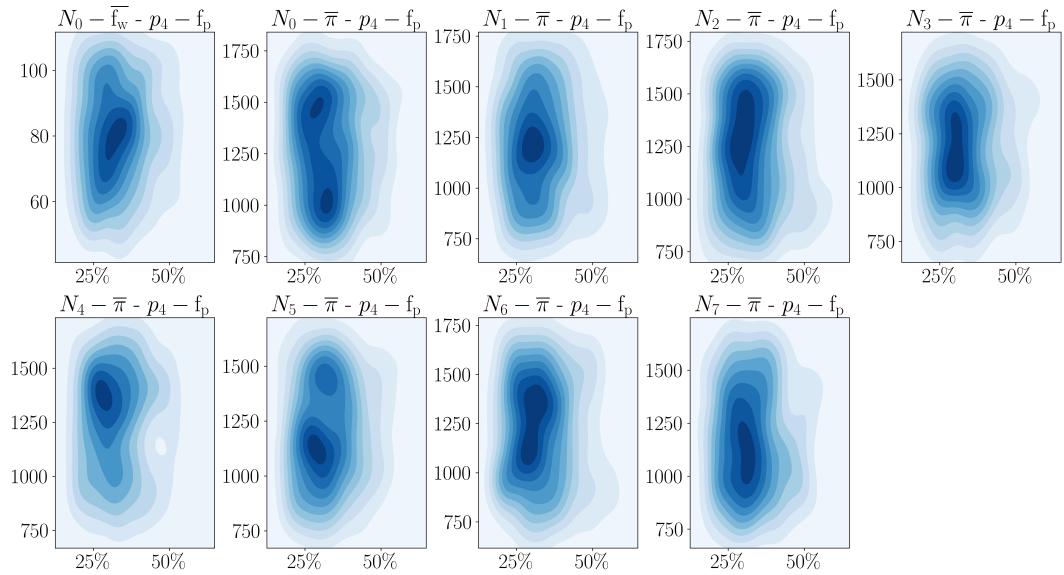


Figure 4.16: Joint PDFs between the inputs and outputs used in the 8-node network, which present a wide dispersion along second variable.

Another remarkable group of joint PDFs arises from the relationship between input and output variables that exhibit a seemingly linear behavior, shown in Figure 4.17. In these distributions, an increase in the input variable tends to be associated with an increase in the output variable, suggesting a positive correlation. This pattern is especially evident in the pairs formed by the upper bound of the demand flow at nodes

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

N_6 and N_7 , and the flow transported through pipelines p_4 , p_5 , p_6 , and p_7 . These pipelines are directly or indirectly responsible for supplying gas to the demand nodes, meaning that as the upper flow limits at the consumption points increase, the optimizer adjusts the flows along these pipelines accordingly, typically utilizing a larger fraction of their capacity. However, it is important to note an exception within this group: the pair formed by the upper flow bound at node N_7 and the flow through pipeline p_3 presents a negative linear relationship. In this case, an increase in the upper demand limit at N_7 is associated with a decrease in the use of pipeline p_3 . This inverse behavior can be explained by the system's looped configuration, where p_3 contributes to a closed trajectory. Under higher demand at N_7 , the optimizer might favor alternative paths that more directly satisfy the load, such as pipelines p_5 through p_7 , reducing the role of p_3 in transporting gas.

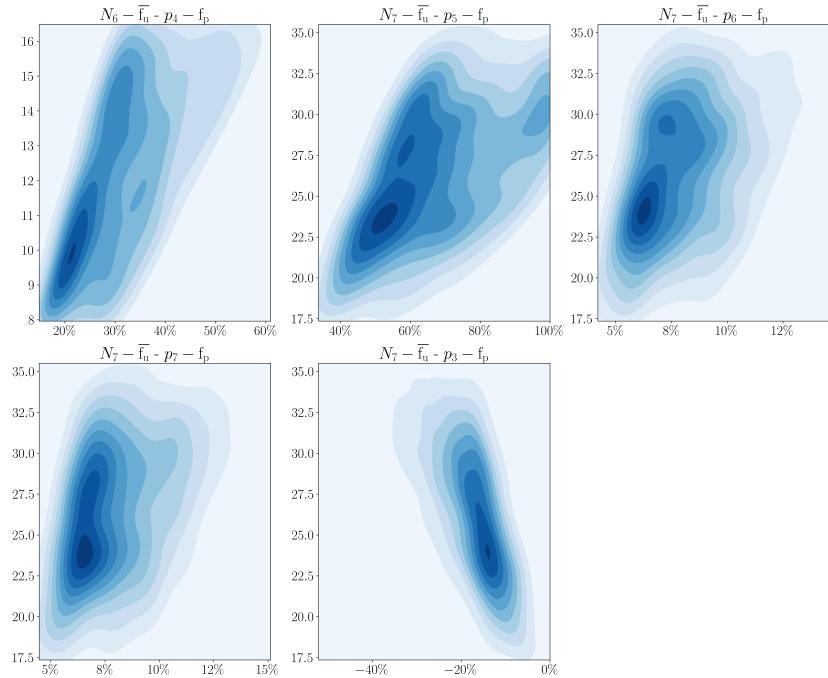


Figure 4.17: Joint PDFs between the inputs and outputs used in the 8-node network, which present a linear behavior, although with different dispersions among them.

A final group of joint distributions worth highlighting is shown in ???. Those involve input and output variables that exhibit a linear relationship, similar to the patterns

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

observed in the optimization model. In particular, two of the joint PDFs show a positive correlation, where increases in the input variable are associated with proportional increases in the output. This behavior is seen in the pairs formed by the upper bound of the demand flow at node N_6 and the predicted flow in pipeline p_4 , as well as between the same bound at node N_7 and the predicted flow in pipeline p_5 . These relationships suggest that the network has captured a coherent system response, in which more flexible demand conditions at the consumption nodes lead to higher utilization of the pipelines supplying them, closely resembling the optimizer's behavior. A third pair, formed by the upper demand bound at node N_7 and the flow in pipeline p_3 , also shows a strong linear pattern, though in this case, the correlation is negative. This result mirrors the exception found in the optimization model, and can again be attributed to the presence of a loop in the network's structure. Under increasing demand at N_7 , the network tends to shift the flow away from pipeline p_3 , likely in favor of more direct paths that connect to the demand node. This further confirms that the trained model replicates the general operating tendencies of the optimization scheme.

This stochastic analysis conducted in the study provided a view of how input uncertainties propagate through the natural gas network and shape the behavior of output variables. By examining the joint probability distributions among inputs, among outputs, and between inputs and outputs, it was possible to uncover structural dependencies, dominant operating regimes, and regions of stability or variability within the system. The presence of strong correlations in certain areas, such as between upstream injection pressures and downstream demands, as well as the identification of operational preferences, like the consistent underutilization of specific pipelines, demonstrate the model's ability to capture operational behavior.

4.2.2 Case Study II: 63-node Network (Colombia)

This section addresses the second case study, focusing on the Colombian natural gas network. As in the previous cases, this analysis explores various configurations of loss function to evaluate the predictive performance of the CensNet-based model. The first experiment examines the model's predictive capabilities when incorporating node and edge losses.

This experiment used optimized hyperparameters, with $N \text{ channels} = 21$, $N \text{ layers} = 5$, and $N \text{ dense} = 4$, which were selected to enhance the model's performance. The experiment yielded a total loss of 267.6, encompassing node and edge losses, along with a calculated balance loss. Specifically, the node loss reached 17.5, while the edge loss was considerably higher at 250.1. Additionally, a balance loss of 338.7 was recorded.

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

Notably, the balance loss was calculated to assess network consistency but was not incorporated into the model's cost function during training; instead, it serves as an independent evaluation metric.

The CensNet-based model's predictive accuracy in this experiment was quantified using R^2 metrics, and the results are shown in Figure 4.18. The nodal predictions exhibited high accuracy, with an R^2 score of 0.993 in Figure 4.18a, indicating that the model closely approximates the observed nodal flow values. Similarly, the edge predictions achieved an R^2 score of 0.963, demonstrating robust performance in predicting edge flows. This last value can be seen in Figure 4.18b.

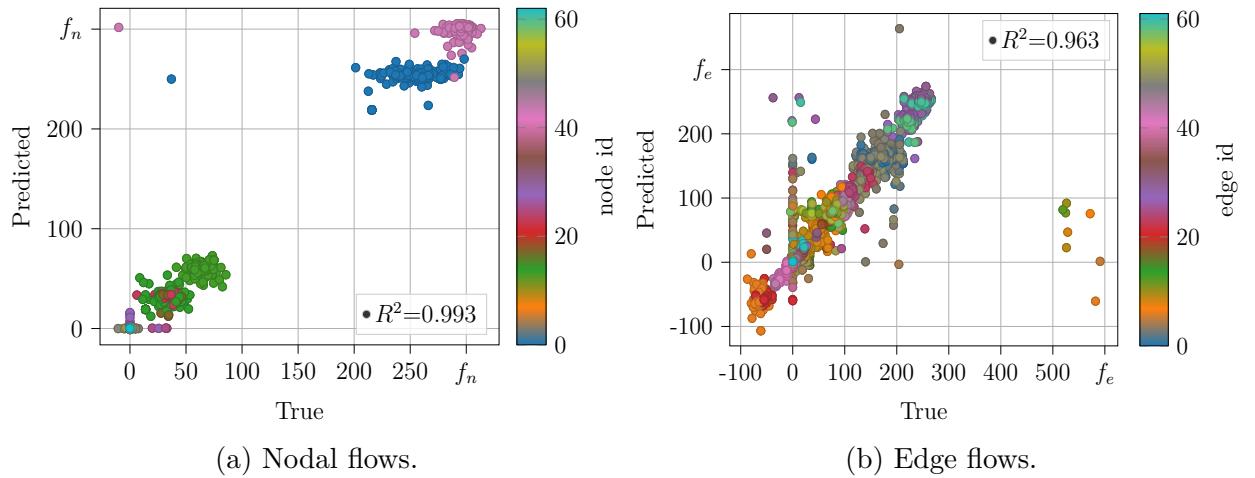


Figure 4.18: Model results using the losses associated with nodal and edge flows in the Colombian case network.

The second experiment evaluated the CensNet-based model, focusing on losses associated with nodes and balance. Using the optimized hyperparameters $N \text{ channels} = 49$, $N \text{ layers} = 5$, and $N \text{ dense} = 2$, the model yielded a total loss of 24.9. This loss value includes a node loss of 17.02, a calculated edge loss of 2,912.2, and a balance loss of 7,9. In this experiment, only the node and balance losses were included in the model's cost function, while the edge loss was computed independently to assess prediction accuracy for edge flows.

The predictive accuracy, evaluated using R^2 metrics, yielded $R^2 = 0.993$ for nodal predictions and $R^2 = 0.57$ for edge predictions, as shown in Figure 4.19a and Figure 4.19b. The scatterplot in Figure 4.19a illustrates the accuracy of nodal predictions, with predicted values aligning closely to the true values. In contrast, the scatterplot

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

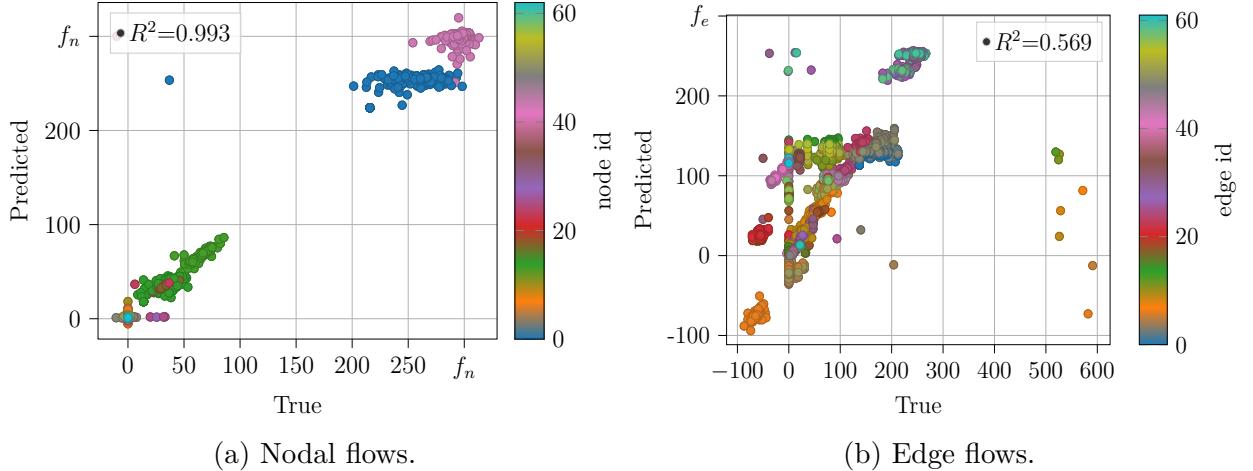


Figure 4.19: Model results using the losses associated with nodal and edge flow predictions along with the gas balance and Weymouth losses in the Colombian case network.

in fig. 4.19b reveals less consistency in edge predictions, likely due to the exclusion of edge loss from the cost function. Although a general linear trend is observable between predicted and true values in this scatterplot, a notable number of inaccurately predicted values diminish the overall R^2 , impacting the edge loss and prediction accuracy.

The third experiment evaluated the CensNet-based model's predictive capabilities by incorporating losses associated with nodes, edges, and balance. With the optimized hyperparameters $N \text{ channels} = 20$, $N \text{ layers} = 5$, and $N \text{ dense} = 2$, the model achieved a total loss of 347.7, including a node loss of 37.7, an edge loss of 264.2, and a balance loss of 45.8. This experiment integrated all three losses into the model's cost function, allowing a more comprehensive assessment of its predictive performance.

The predictive accuracy was quantified by R^2 values of 0.984 and 0.961 for nodes and edges, respectively, as shown in Figure 4.20a and Figure 4.20b. In these scatterplots, the predictions demonstrate a clearer alignment with the true values, with both graphs illustrating shapes more similar to straight lines, which indicates improved model performance in capturing the underlying patterns of nodal and edge flows.

This third experiment marks the final test conducted in this study. Attempts to include the Weymouth loss were hindered by significant computational complexity, preventing the acquisition of reliable results for the model when considering the Weymouth function loss. Therefore, further tests were not feasible within the current scope.

The results in Table 4.2 show the differences between the IPOPT benchmark and

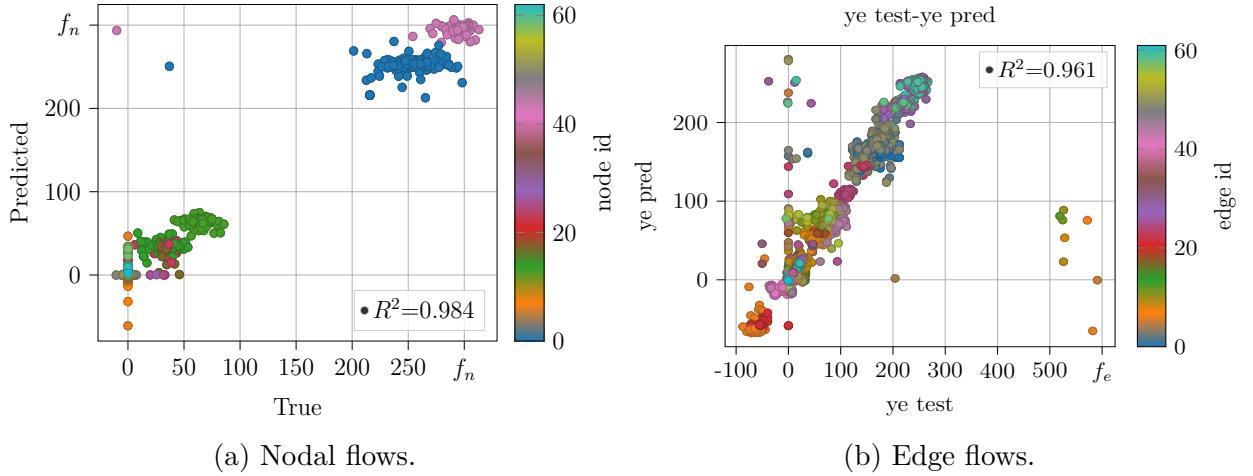


Figure 4.20: Model results using the losses associated with nodal and edge flow predictions along with the gas balance loss in the Colombian case network.

Method	Node Error	Edge Error	Balance Error	Time
CensNet(N)	0.10 ± 69.24	0.65 ± 116.75	0.10 ± 39.19	13.50 ± 2.86
CensNet(N+E)	-1.15 ± 69.14	-24.94 ± 110.98	-1.15 ± 29.84	13.50 ± 2.86
CensNet(N+E+B)	-1.01 ± 68.96	0.32 ± 116.28	-1.01 ± 30.68	13.50 ± 2.86

Table 4.2: Differences in mean and standard deviation between IPOPT and CensNet-based models for nodal flows (Node Error), edge flows (Edge Error), nodal balance (Balance Error), and prediction time (Time) in the Colombian natural gas system. Each value represents the difference IPOPT - CensNet.

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

the CensNet models applied to the Colombian natural gas network. As with the synthetic case, all CensNet configurations offer a significant advantage in prediction time, maintaining fast inference across all loss settings.

When trained with only the nodal loss (CensNet (N)), the model achieves low nodal error and a moderate edge error. However, the balance error remains relatively high, suggesting that nodal supervision alone is insufficient to enforce global consistency. Incorporating the edge loss (CensNet (N+E)) unexpectedly worsens the edge error. Further, it increases the balance error, indicating that without additional physical constraints, the model may overfit local patterns at the expense of system-wide coherence.

Adding the balance loss (CensNet (N+E+B)) improves edge predictions significantly, reducing the edge error from -24.94 to 0.32 , while also recovering a more accurate balance prediction. Although the nodal error remains similar, this configuration yields the best trade-off between nodal, edge, and balance accuracy.

Among the evaluated settings, **CensNet (N+E+B)** delivers the most physically consistent results, closely approximating the IPOPT benchmark while preserving the CensNet model's computational advantage.

An uncertainty analysis similar to the previous case was performed for the Colombian natural gas transportation network. Using the same methodology, a second kernel density estimate (KDE) was fitted to the training output data, and synthetic outputs were generated. The log-likelihood computed from the synthetic outputs was $-104,413,419.34$, while that from the training outputs was $-104,037,047.58$. To further assess the similarity between the two distributions, a Kolmogorov–Smirnov (K–S) test was conducted under three different alternatives: two-sided, less, and greater. The results were as follows: for the two-sided alternative, the test statistic was 0.00527 with a p-value of 0.95556 (statistic location 0.52109); for the less alternative, the test statistic was 0.00527 with a p-value of 0.59061 (statistic location 0.52109); and for the greater alternative, the test statistic was 0.00324 with a p-value of 0.81777 (statistic location -0.73443). These findings indicate a high degree of similarity between the synthetic and training outputs, as the low test statistics and high p-values suggest that the two distributions are statistically indistinguishable. The slight differences in log-likelihood values further confirm that the synthetic data effectively capture the essential characteristics of the training data, thereby validating the robustness of the stochastic sampling approach for this network.

4.3 Discussion

This chapter has explored the predictive capabilities of a CensNet-based model for gas network simulation under various loss function configurations. The results emphasize the importance of designing loss functions that integrate physical and operational constraints in order to improve the model’s ability to generalize and maintain consistency across different components of the network.

The initial experiments in the 8-node gas system confirmed that nodal loss alone is sufficient for capturing nodal flow patterns, particularly under nonlinear conditions. However, when edge or balance losses are added, the model gains the ability to replicate more complex network dynamics. This trade-off between precision in nodal predictions and general system consistency reveals the flexibility of the CensNet model and its responsiveness to different training objectives. Incorporating edge loss helps improve flow predictions, while balance constraints contribute to system-wide consistency, even if they occasionally introduce minor trade-offs in prediction accuracy.

The inclusion of the Weymouth equation loss proved to be more challenging due to its nonlinear nature. Its integration caused a noticeable decrease in edge prediction accuracy, particularly in regions with compressors and closed loops, highlighting the difficulty of aligning learned representations with complex physical equations. These findings underline the need for careful calibration when physical constraints are introduced into data-driven models.

A comparative analysis of the CensNet model across linear and nonlinear cases further emphasizes the benefits of incorporating physical knowledge. The results from the nonlinear case showed improved balance accuracy, suggesting that physically-informed constraints can enhance the model’s ability to simulate realistic operating conditions. While nodal and edge predictions remained consistent across chapters, the balance constraint had a particularly beneficial effect, reducing variability and error without compromising computational efficiency.

For the 8-node case, the study was complemented with an in-depth probabilistic analysis of the model’s outputs. This involved examining the joint distributions of selected variables and revealed patterns of interdependence between operational variables, such as nodal pressures and flows through pipelines or compressors. These findings confirmed that the trained model not only learned the mapping between inputs and outputs but also preserved structural dependencies across the network.

When applied to the Colombian natural gas network, the CensNet model displayed consistent behavior across different combinations of losses. Including nodal and edge losses yielded strong performance in nodal predictions, with minimal deviation in bal-

CHAPTER 4. STOCHASTIC MODELING OF NATURAL GAS FLOWS USING PHYSICS-GUIDED NEURAL NETWORKS

ance, suggesting that such configurations are effective in applications where nodal flow is prioritized. Introducing balance constraints improved system consistency while maintaining strong overall performance.

To assess generalization under uncertainty, a stochastic framework was applied to both networks. Synthetic input samples were generated using kernel density estimation (KDE) and propagated through the trained model. The outputs were then compared with those from the training data using log-likelihood estimates and Kolmogorov–Smirnov (K–S) tests under three alternatives: two-sided, less, and greater. In both cases, the close match between the distributions, reflected in similar log-likelihood values and non-significant K–S statistics, validated the consistency and robustness of the trained model under unseen conditions. These results support the idea that the model has internalized features of the data distribution and can provide reliable predictions beyond the training set.

By validating the ability of the CensNet model to approximate system behavior under uncertainty, and by showing how stochastic sampling can be used to evaluate the variability of decision-relevant outputs, this work enables the formulation of optimization strategies that move beyond deterministic dispatch. The integration of uncertainty quantification into the modeling process lays the groundwork for decision-making frameworks that are robust under data variability and can be aligned with the operational constraints of gas transport systems. As such, the developments presented here represent a step toward the implementation of a stochastic gas flow optimization strategy for real-world applications.

Chapter 5

Conclusions and future work

5.1 Conclusions

This thesis set out to develop an optimization tool for natural gas transportation networks by combining knowledge of the network topology, an appropriate approximation of the Weymouth equation, and stochastic optimization techniques. Each specific objective contributed to achieving this goal, and the progress toward each objective is presented across the thesis chapters.

The first specific objective was to design a Graph Neural Network (GNN)-based approach that integrates natural gas network topology to reduce computational time for operational estimation. This objective was explored in Chapters 2 and 4, where a GNN-based model was applied to predict decision variables of the natural gas system while significantly reducing computation times compared to traditional optimization techniques. Both chapters included comparisons of computation times between optimizer-based methods and the GNN-based approach, with the GNN consistently providing faster results. In Chapter 2, a GNN was trained to approximate a standard linear optimization model for a simplified gas system (without pressure considerations) and then tested on more extensive networks, such as the Colombian gas system. Chapter 4 extended this work by incorporating physical constraints, like gas balance and the Weymouth equation, into the loss function, resulting in a more physically accurate model with consistent gains in computational efficiency. Across all tests, the GNN model provided predictions much faster than optimizers, demonstrating that this approach successfully meets the objective of reducing computational time for operational estimation.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

The second specific objective was to develop an optimization model for natural gas transportation systems that incorporates the Weymouth equation to reduce approximation errors in gas flow calculations. Chapter 3 addressed this objective by developing a Mathematical Program with Complementarity Constraints (MPCC)-based optimization model that accurately represents the nonlinear Weymouth equation without requiring mixed-integer formulations. This approach leveraged binary-behaving continuous variables to capture the bidirectional nature of gas flows, enabling the model to represent operational decisions more precisely under complex and high-demand network conditions. The model was successfully tested on realistic systems, including the Colombian gas-power system, consistently achieving lower flow approximation errors compared to traditional methods. These results underscore the practical applicability of MPCC in gas transport optimization, especially for scenarios that demand reliable and accurate short-term scheduling.

The third specific objective was to develop a stochastic optimization strategy that quantifies uncertainties in gas system operation by sampling from the probability distributions of the constraints in the transportation problem. This objective was addressed in Chapter 4, where the GNN-based model—previously formulated in Chapter 2—was extended to incorporate physical constraints, enhancing its ability to model the complex behavior of natural gas networks. Once trained, the GNN model was capable of rapidly generating responses through forward propagation, enabling efficient evaluation of multiple scenarios without requiring repeated optimization. This rapid-response feature provides a practical mechanism for addressing uncertainty in gas system operation by allowing a wide range of scenarios to be tested quickly and systematically.

To support this strategy, stochastic analyses were conducted in Chapter 4 using kernel density estimates (KDEs) fitted to the training data. These KDEs were used to generate synthetic samples that emulate plausible input conditions. The GNN model’s outputs under these synthetic conditions were then compared to the distribution of training outputs using log-likelihood measures and Kolmogorov–Smirnov (K–S) tests. In both the 8-node and Colombian gas networks, the synthetic outputs were found to be statistically similar to those derived from the training set, with low test statistics and high p-values across multiple K–S test alternatives. These results confirmed that the trained models not only generalize well but also capture the underlying behavior of the system, validating the approach’s effectiveness in representing uncertainty.

While a stochastic optimization framework with probabilistic objective functions and constraints was not explicitly implemented, the capacity to efficiently simulate and evaluate numerous uncertain scenarios satisfies the core principles of stochastic opti-

mization. Thus, the proposed methodology achieves the intended objective by enabling a scalable and reliable strategy for uncertainty quantification in gas transportation networks.

5.2 Future Work

Future research can explore several directions to enhance the application and effectiveness of GNN-based models and MPCC formulations in natural gas and broader energy systems. One option is extending the GNN-based model to account for transient dynamics and operational uncertainty. This enhancement is particularly relevant in contexts involving the increasing integration of renewable energy sources, where variability in supply can significantly affect gas system operation.

Another direction is the development of stochastic optimization models that account for the variability introduced by renewable energy sources such as solar and wind. Incorporating stochastic elements into the MPCC formulation would enhance the framework's ability to maintain reliable operation under uncertain demand and supply conditions. This extension is particularly relevant given the increasing need for robust decision-making in energy systems subject to high levels of variability and uncertainty.

Additionally, improving the design of loss functions in GNN architectures remains an area of development. This work has shown that the inclusion of physical constraints, such as balance conditions and flow equations, enhances predictive accuracy. Future research could focus on implementing loss functions associated with the Weymouth equation that retain physical fidelity while reducing computational complexity, making them suitable for large-scale or real-time applications.

Bibliography

- [1] M. S. Alam, S. R. Paramati, M. Shahbaz, and M. Bhattacharya, “Natural gas, trade and sustainable growth: empirical evidence from the top gas consumers of the developing world,” *Applied Economics*, vol. 49, pp. 635–649, 2 2017.
- [2] J.-H. Yuan, S. Zhou, T.-D. Peng, G.-H. Wang, and X.-M. Ou, “Petroleum substitution, greenhouse gas emissions reduction and environmental benefits from the development of natural gas vehicles in china,” *Petroleum Science*, vol. 15, no. 3, pp. 644–656, 2018.
- [3] X. Yin, K. Wen, Y. Wu, X. Han, Y. Mukhtar, and J. Gong, “A machine learning-based surrogate model for the rapid control of piping flow: Application to a natural gas flowmeter calibration system,” *Journal of Natural Gas Science and Engineering*, vol. 98, p. 104384, 2022.
- [4] M. Aydin, “Natural gas consumption and economic growth nexus for top 10 natural gas-consuming countries: A granger causality analysis in the frequency domain,” *Energy*, vol. 165, p. 179–186, 2018.
- [5] IEA, “Eia projections indicate global energy consumption increases through 2050, outpacing efficiency gains and driving continued emissions growth,” 2023.
- [6] Promigas, “Informe del sector gas natural 2021,” 2021.
- [7] S. Arango-Aramburo, S. W. Turner, K. Daenzer, J. P. Ríos-Ocampo, M. I. Hejazi, T. Kober, A. C. Álvarez Espinosa, G. D. Romero-Otalora, and B. van der Zwaan, “Climate impacts on hydropower in colombia: A multi-model assessment of power sector adaptation pathways,” *Energy Policy*, vol. 128, p. 179–188, May 2019.
- [8] C. Villa-Loaiza, I. Taype-Huaman, J. Benavides-Franco, G. Buenaventura-Vera, and J. Carabalí-Mosquera, “Does climate impact the relationship between the

BIBLIOGRAPHY

- energy price and the stock market? the colombian case,” *Applied Energy*, vol. 336, p. 120800, Apr 2023.
- [9] M. P. Ignacio Fariza, “La sequía desploma la generación hidroeléctrica hasta su nivel más bajo en tres décadas,” Aug 2022.
 - [10] X. Shan, W. Yu, J. Gong, W. Huang, K. Wen, H. Wang, S. Ren, D. Wang, Y. Shi, and C. Liu, “A methodology to evaluate gas supply reliability of natural gas pipeline network considering the effects of natural gas resources,” *Reliability Engineering & System Safety*, vol. 238, p. 109431, 2023.
 - [11] J. Restrepo-Trujillo, R. Moreno-Chuquen, F. Jiménez-García, W. Flores, and H. Chamorro, “Scenario analysis of an electric power system in colombia considering the el niño phenomenon and the inclusion of renewable energies,” *Energies*, vol. 15, p. 6690, Sep 2022.
 - [12] J. Restrepo-Trujillo, R. Moreno-Chuquen, and F. N. Jiménez-García, “Strategies of expansion for electric power systems based on hydroelectric plants in the context of climate change: Case of analysis of colombia,” *International Journal of Energy Economics and Policy*, vol. 10, no. 6, p. 66–74, 2020.
 - [13] E. Raheli, Q. Wu, M. Zhang, and C. Wen, “Optimal coordinated operation of integrated natural gas and electric power systems: A review of modeling and solution methods,” *Renewable and Sustainable Energy Reviews*, vol. 145, p. 111134, 2021.
 - [14] S. Chen, A. J. Conejo, R. Sioshansi, and Z. Wei, “Unit commitment with an enhanced natural gas-flow model,” *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3729–3738, 2019.
 - [15] G. Byeon and P. Van Hentenryck, “Unit commitment with gas network awareness,” *IEEE Transactions on Power Systems*, vol. 35, no. 2, p. 1327–1339, 2020.
 - [16] J. D. Morcillo, F. Angulo, and C. J. Franco, “Analyzing the hydroelectricity variability on power markets from a system dynamics and dynamic systems perspective: Seasonality and enso phenomenon,” May 2020.
 - [17] C. Villa-Loaiza, I. Taype-Huaman, J. Benavides-Franco, G. Buenaventura-Vera, and J. Carabalí-Mosquera, “Does climate impact the relationship between the energy price and the stock market? the colombian case,” *Applied Energy*, vol. 336, p. 120800, Apr 2023.

BIBLIOGRAPHY

- [18] L. A. Cuartas, A. P. Cunha, J. A. Alves, L. M. Parra, K. Deusdará-Leal, L. C. Costa, R. D. Molina, D. Amore, E. Broedel, M. E. Seluchi, and et al., “Recent hydrological droughts in brazil and their impact on hydropower generation,” *Water*, vol. 14, p. 601, Feb 2022.
- [19] R. Bessa, C. Moreira, B. Silva, and M. Matos, “Handling renewable energy variability and uncertainty in power system operation,” *Advances in Energy Systems*, p. 1–26, Feb 2019.
- [20] V. Sharifi, A. Abdollahi, M. Rashidinejad, E. Heydarian-Foroushani, and H. H. Alhelou, “Integrated electricity and natural gas demand response in flexibility-based generation maintenance scheduling,” *IEEE Access*, vol. 10, pp. 76021–76030, 2022.
- [21] L. Yang, X. Zhao, and Y. Xu, “A convex optimization and iterative solution based method for optimal power-gas flow considering power and gas losses,” *International Journal of Electrical Power & Energy Systems*, vol. 121, p. 106023, 2020.
- [22] T. Jiang, C. Yuan, R. Zhang, L. Bai, X. Li, H. Chen, and G. Li, “Exploiting flexibility of combined-cycle gas turbines in power system unit commitment with natural gas transmission constraints and reserve scheduling,” *International Journal of Electrical Power & Energy Systems*, vol. 125, p. 106460, 2021.
- [23] C. Wang, S. Wang, F. Liu, T. Bi, and T. Wang, “Risk-loss coordinated admissibility assessment of wind generation for integrated electric-gas systems,” *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4454–4465, 2020.
- [24] R. J. Trudeau, *Introduction to graph theory*. Dover Publications, 2015.
- [25] E. A. Bender and S. G. Williamson, *Section 2: Digraphs, Paths, and Subgraphs*. University of California at San Diego, 2010.
- [26] R. J. Wilson, *Introduction to graph theory*, vol. 1. Oliver & Boyd, 4 ed., 1972.
- [27] H. Hippert, C. Pedreira, and R. Souza, “Neural networks for short-term load forecasting: A review and evaluation,” *IEEE Transactions on Power Systems*, vol. 16, p. 44–55, Feb 2001.
- [28] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks;,” *International Journal of Forecasting*, vol. 14, p. 35–62, Mar 1998.

BIBLIOGRAPHY

- [29] Y. Jia, J. Wang, W. Shou, M. R. Hosseini, and Y. Bai, “Graph neural networks for construction applications,” *Automation in Construction*, vol. 154, p. 104984, Oct 2023.
- [30] Y. Jia, J. Wang, W. Shou, M. R. Hosseini, and Y. Bai, “Graph neural networks for construction applications,” *Automation in Construction*, vol. 154, p. 104984, Oct 2023.
- [31] A. Gupta, P. Matta, and B. Pant, “Graph neural network: Current state of art, challenges and applications,” *Materials Today: Proceedings*, vol. 46, p. 10927–10932, Mar 2021.
- [32] C. Liu, J. Wu, W. Liu, and W. Hu, “Enhancing graph neural networks by a high-quality aggregation of beneficial information,” *Neural Networks*, vol. 142, p. 20–33, Oct 2021.
- [33] W. L. Hamilton, “Graph representation learning,” *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 14, no. 3, pp. 1–159.
- [34] T. N. Kipf and M. Welling, “Semi-Supervised Classification with Graph Convolutional Networks,” in *Proceedings of the 5th International Conference on Learning Representations*, ICLR ’17, 2017.
- [35] S. Fu, S. Wang, W. Liu, B. Liu, B. Zhou, X. You, Q. Peng, and X.-Y. Jing, “Adaptive graph convolutional collaboration networks for semi-supervised classification,” *Information Sciences*, vol. 611, p. 262–276, Sep 2022.
- [36] L. Gong and Q. Cheng, “Exploiting edge features for graph neural networks,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9203–9211, 2019.
- [37] Z. Guo, Y. Zhang, Z. Teng, and W. Lu, “Densely connected graph convolutional networks for graph-to-sequence learning,” *Transactions of the Association for Computational Linguistics*, vol. 7, p. 297–312, Nov 2019.
- [38] X. Jiang, R. Zhu, P. Ji, and S. Li, “Co-embedding of nodes and edges with graph neural networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7075–7086, 2023.

BIBLIOGRAPHY

- [39] L. D. R. Beal, D. C. Hill, R. A. Martin, and J. D. Hedengren, “Gekko optimization suite,” Jul 2018.
- [40] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.
- [41] Y. Wang, Y. Wang, Y. Huang, J. Yang, Y. Ma, H. Yu, M. Zeng, F. Zhang, and Y. Zhang, “Operation optimization of regional integrated energy system based on the modeling of electricity-thermal-natural gas network,” *Applied Energy*, vol. 251, p. 113410, 2019.
- [42] Z. Yang, K. Xie, J. Yu, H. Zhong, N. Zhang, and Q. Xia, “A general formulation of linear power flow models: Basic theory and error analysis,” *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1315–1324, 2019.
- [43] L. Yang, Y. Xu, H. Sun, and X. Zhao, “Two-stage convexification-based optimal electricity-gas flow,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1465–1475, 2020.
- [44] Y. Zhang, S. Xie, and S. Shu, “Multi-stage robust optimization of a multi-energy coupled system considering multiple uncertainties,” *Energy*, vol. 238, p. 122041, Jan 2022.
- [45] B. Baumrucker, J. Renfro, and L. Biegler, “Mpec problem formulations and solution strategies with chemical engineering applications,” *Computers & Chemical Engineering*, vol. 32, no. 12, p. 2903–2913, 2008.
- [46] Y. Zhang, J. Wu, and L. Zhang, “First order necessary optimality conditions for mathematical programs with second-order cone complementarity constraints,” *Journal of Global Optimization*, vol. 63, no. 2, p. 253–279, 2015.
- [47] G. Bouza and G. Still, “Mathematical programs with complementarity constraints: Convergence properties of a smoothing method,” *Mathematics of Operations Research*, vol. 32, p. 467–483, May 2007.
- [48] S. Scholtes, “Convergence properties of a regularization scheme for mathematical programs with complementarity constraints,” *SIAM Journal on Optimization*, vol. 11, p. 918–936, Mar 2001.

BIBLIOGRAPHY

- [49] H. Scheel and S. Scholtes, “Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity,” *Mathematics of Operations Research*, vol. 25, p. 1–22, Feb 2000.
- [50] D. Ralph and S. J. Wright, “Some properties of regularization and penalization schemes for mpecs,” *Optimization Methods and Software*, vol. 19, no. 5, pp. 527–556, 2004.
- [51] C. Ordoudis, P. Pinson, and J. M. Morales, “An integrated market for electricity and natural gas systems with stochastic power producers,” *European Journal of Operational Research*, vol. 272, no. 2, pp. 642–654, 2019.
- [52] A. Schwele, C. Ordoudis, J. Kazempour, and P. Pinson, “Coordination of power and natural gas systems: Convexification approaches for linepack modeling,” in *2019 IEEE Milan PowerTech*, pp. 1–6, 2019.
- [53] S. García-Marín, W. González-Vanegas, and C. Murillo-Sánchez, “Mpng: A matpower-based tool for optimal power and natural gas flow analyses,” *IEEE Transactions on Power Systems*, pp. 1–9, 2022.
- [54] C. Wang, W. Wei, J. Wang, and T. Bi, “Convex optimization based adjustable robust dispatch for integrated electric-gas systems considering gas delivery priority,” *Applied Energy*, vol. 239, pp. 70–82, 2019.
- [55] Y. Zhao, M. Gong, J. Sun, C. Han, L. Jing, B. Li, and Z. Zhao, “A new hybrid optimization prediction strategy based on sh-informer for district heating system,” *Energy*, vol. 282, p. 129010, 2023.