Physics-Informed Neural Networks-based Optimization for Gas-Powered Energy Systems

Diego Armando Pérez Rosero

Universidad Nacional de Colombia
Faculty of Electrical, Electronics and Computing Engineering
Department of Engineering and Architecture
Manizales, Colombia
2024

Physics-Informed Neural Networks-based Optimization for Gas-Powered Energy Systems

Diego Armando Pérez Rosero

Thesis presented as a partial requirement for obtaining the degree of:
Maestría en Ingeniería-Automatización Industrial

Advisor:
Prof. Dr. Andrés Marino Álvarez Meza
Co-advisor:
Prof. Dr. César Germán Castellanos Domínguez

Research Line:
Artificial Intelligence
Research Group:
Control and Digital Signal Processing Group (GCPDS)

Universidad Nacional de Colombia
Faculty of Electrical, Electronics and Computing Engineering
Department of Engineering and Architecture
2024

Optimización basada en Redes Neuronales Informadas por la Física para Sistemas de Energía a Gas

Diego Armando Pérez Rosero

Tesis presentada como requisito parcial para optar por el título de:
Maestría en Ingeniería-Automatización Industrial

Director(a):
Prof. Dr. Andrés Marino Álvarez Meza
Codirector(a):
Prof. Dr. César Germán Castellanos Domínguez

Línea de investigación:
Inteligencia Artificial
Grupo de investigación:
Grupo de control y procesamiento digital de señales (GCPDS)

Universidad Nacional de Colombia
Facultad de ingeniería eléctrica, electrónica y computación
Departamento de ingeniería y arquitectura
2024

"The concern for man and his destiny must always be the primary interest of any technical effort. Never forget this between your diagrams and equations"

Albert Einstein

"Believe you can, and you are halfway there."

Theodore Roosevelt

Declaración

Me permito afirmar que he realizado ésta tesis de manera autónoma y con la única ayuda de los medios permitidos y no diferentes a los mencionados el presente texto. Todos los pasajes que se han tomado de manera textual o figurativa de textos publicados y no publicados, los he reconocido en el presente trabajo. Ninguna parte del presente trabajo se ha empleado en ningún otro tipo de tesis.

Manizales., 07-23-2024

Diego Armando Pérez Rosero

Acknowledgements

I extend my deepest gratitude to my parents, whose steadfast belief and unwavering support have been the cornerstone of my academic and personal endeavors. Additionally, I extend special thanks to my siblings, whose companionship, empathy, and enduring encouragement have been an irreplaceable pillar of strength and inspiration.

I am also indebted to two distinguished professors, Andrés Marino Álvarez Meza and Germán Castellanos Domínguez, for their invaluable mentorship during my research journey. Their insights and guidance have been crucial to my scholarly growth.

My appreciation goes to the members of the Signal Processing and Recognition Group(SPRG), at the Manizales branch of the Universidad Nacional de Colombia a for their thought-provoking reflections and educational discussions, which have greatly contributed to my academic journey.

Finally, I recognize that this research would not have been possible without the support given by the project: *"Desarrollo de una herramienta para la planeación a largo plazo de la operación del sistema de transporte de gas natural de Colombia'"*, supported by MINCIENCIAS (Code 111085269982).

Diego Armando Pérez Rosero
2024

Abbreviations

AD Automatic Differentiation 15, 16, 21, 24, 25, 27, 28, 32, 33, 67

AI Artificial Intelligence 6, 7, 18

ANN Artificial Neural Networks iii, v, 15, 16, 21, 42, 46, 59, 60, 67

BC Branch and Cut 13, 18

CREG Comisión de Regulación de Energía y Gas iii, v, 8

CVXPY Convex Optimization in Python vii, xii, 23, 32, 34

DNN Deep Neural Networks 18

EXP Exponential Programming 13, 18

GCPDS Grupo de Control y Procesamiento Digital de Señales 6, 7

GNN Graph Neural Networks 16

IP Inner Point 14, 18

LP Linear Programming 14, 18

MAPE Mean Absolute Percentage Error vii, ix, 39–41, 43, 59, 63

MIP Mixed-Integer Programming 13, 17, 18

NLP Non-Lineal Programming 17, 18

NN Neural Networks 29, 44

NOPT Nonlinear Optimization iii, v–vii, x, 8–11, 13–16, 20–23, 34, 38–40, 42, 43, 45, 58, 59, 67

PCP power cone programming 13, 18

PINN Physics-Informed Neural Network viii, 11, 12, 16, 18, 20, 45, 52, 55, 67

QP Quadratic Programming 13, 18

RPINN Regularized Physics-Informed Neural Network iii, v, vii–x, xii, 16, 20–23, 34–40, 42–45, 53, 59–62, 67

SDP Semidefinite Programming 13, 18

SOCP Second-Order Cone Programming 13, 18

UPME Unidad de Planeación Minero Energética 8

Resumen

La transición global hacia un paradigma energético sostenible destaca tanto la persistencia de los combustibles fósiles como el auge de las fuentes de energía renovable. En este contexto, la adopción de tecnologías basadas en turbinas de gas surge como una opción viable que minimiza el impacto ambiental al reducir las emisiones y los costos operativos. Aunque se espera que la demanda de combustibles fósiles para la generación de electricidad aumente hasta 2030, paralelamente a este evento, se prevé una inversión considerable en energía renovable, especialmente en los sectores solar y eólico. América Latina, con una notable capacidad hidroeléctrica de 340,332 MW, está a la vanguardia de los esfuerzos de mitigación de emisiones de carbono. Sin embargo, la región enfrenta desafíos, incluida la dependencia de combustibles importados y la fragilidad de su infraestructura energética.

En Colombia, la variabilidad climática impulsa la diversificación energética. La Comisión de Regulación de Energía y Gas (CREG) tiene como objetivo optimizar el transporte de gas, consciente de la limitación temporal de las reservas actuales, estimadas en solo siete años. El modelo lineal utilizado tradicionalmente por CREG, aunque efectivo para ciertos propósitos, no considera factores críticos como las variaciones de presión y el papel de las estaciones de compresión, revelando la necesidad de un modelo de optimización no lineal (NOPT). Sin embargo, los desafíos de NOPT incluyen la gestión de datos variables y mediciones ruidosas, lo que puede llevar a soluciones erróneas o subóptimas.

Este documento presenta un nuevo enfoque: la Red Neuronal Informada por la Física Regularizada (RPINN), diseñada para abordar tanto tareas de optimización supervisada como no supervisada. RPINN combina funciones de activación personalizadas y penalizaciones de regularización dentro de una arquitectura de red neuronal artificial (ANN), lo que le permite manejar la variabilidad de los datos y entradas inexactas. Incorpora principios físicos en el diseño de la red, calculando variables de optimización a partir de los pesos de la red y las características aprendidas. Además, emplea técnicas de diferenciación automática para optimizar la escalabilidad del sistema y reducir el tiempo de cálculo mediante la retropropagación por lotes.

Los resultados experimentales demuestran que RPINN es competitivo con los solucionadores de última generación para tareas de optimización supervisada y no supervisada, mostrando robustez frente a mediciones ruidosas. Este avance lo posiciona como una solución prometedora para entornos con información fluctuante, como se evidencia en los modelos de mezcla uniforme y sistemas alimentados por gas. La base ANN de RPINN no solo asegura flexibilidad y escalabilidad, sino que también destaca su potencial como una metodología robusta y efectiva en comparación con la tradicional.

Palabras clave: Optimización, Sistemas de Gas, Redes Neuronales, Programación No Lineal, Convergencia, Tiempo de Cómputo,Redes Informadas Físicamente.

Abstract

The global transition towards a sustainable energy paradigm highlights both the persistence of fossil fuels and the rise of renewable energy sources. In this context, the adoption of gas turbine-based technologies emerges as a viable option that minimizes environmental impact by reducing emissions and operational costs. While the demand for fossil fuels for electricity generation is expected to increase until 2030, parallel to this event a considerable investment in renewable energy, especially in the solar and wind sectors, is expected. Latin America, with a notable hydroelectric capacity of 340,332 MW, is at the forefront of carbon emission mitigation efforts. However, the region faces challenges, including dependence on imported fuels and the fragility of its energy infrastructure.

In Colombia, climatic variability drives energy diversification. The Energy and Gas Regulatory Commission (CREG) aims to optimize gas transportation, aware of the temporary limitation of current reserves, estimated at only seven years. The linear model traditionally used by CREG, though effective for certain purposes, does not consider critical factors such as pressure variations and the role of compression stations, revealing the need for a nonlinear optimization (NOPT) model. However, NOPT challenges include managing variable data and noisy measurements, which can lead to erroneous or suboptimal solutions.

This document presents a new approach: the Regularized Physics-Informed Neural Network (RPINN), designed to address both supervised and unsupervised optimization tasks. RPINN combines custom activation functions and regularization penalties within an artificial neural network (ANN) architecture, enabling it to handle data variability and inaccurate inputs. It incorporates physical principles into the network design, calculating optimization variables from network weights and learned features. Additionally, it employs automatic differentiation techniques to optimize system scalability and reduce computation time through batch-based backpropagation.

Experimental results demonstrate that RPINN is competitive with state-of-the-art solvers for both supervised and unsupervised optimization tasks, exhibiting robustness against noisy measurements. This advancement positions it as a promising solution for environments with fluctuating information, as evidenced in uniform mixture models and gas-fed systems. The ANN foundation of RPINN not only ensures flexibility and scalability but also highlights its potential as a robust and effective methodology when compared to traditional approaches.

Keywords: Optimization, Gas Systems, Neural Networks, Nonlinear Programming, Convergence, Computation Time, Physic-Informed Networks.

List of Figures

1-1 Colombia's energy mix in perspective: The graph contrasts Colombia's heavy reliance on hydropower with global and South American energy trends. . . . 2

1-2 Colombian Power Generation Model: This figure illustrates the overall pattern of Colombia's integrated gas-electricity energy system, highlighting its crucial role in energy sustainability. 4

1-3 Optimization Framework: This diagram illustrates the strategy employed by the Energy and Gas Regulatory Commission (CREG) to optimize gas system operations, where the optimal flows are denoted by F^*. 5

1-4 Classical Optimization Model: This figure depicts the general framework of a traditional optimization method aimed at finding a solution to a problem, emphasizing the systematic approach from the initial point to the optimal outcome. 5

1-5 Deep Learning Optimization Model: This figure illustrates the general framework of a deep learning model applied to an optimization problem, highlighting the process from data input to the iterative search for an optimal solution. . 6

1-6 Problem Overview: This diagram illustrates the primary factors affecting nonlinear optimization challenges in gas pipeline network design. Factors leading to non-convergence are indicated in blue, while those contributing to extended computational time are marked in red. 9

1-7 Convergence issues in nonlinear optimization: The figure summarizes the factors contributing to the non-convergence in NOPT problems. 11

1-8 Time complexity challenges for nonlinear optimization: The figure presents a summary of factors leading to extended computation times in nonlinear optimization. 12

1-9 Mathematical programming-based methods. Red: main drawbacks. Blue: reference. 13

1-10 Gradient-based methods. Red: main drawbacks. Blue: reference. 14

1-11 Decomposition-based methods. Red: main drawbacks. Blue: reference. . . . 15

1-12 Search-based methods. Red: main drawbacks. Blue: reference. 15

1-13 Summary of key deep learning techniques for nonlinear problem optimization, with drawbacks highlighted in red. 17

3-1 Thesis Contributions 21

4-1 Function Differentiation 24

4-2 Forward propagation: The graph is illustrated with the architecture of the example seen in Eq. 4-5, with the arithmetic operations specific to the studied multivariate function 28

4-3 Backward propagation: The graph is illustrated with the architecture of the example seen in Eq. 4-5, with the arithmetic operations specific to the studied multivariate function 28

4-4 Schematic representation of a neural network: The diagram displays the general components of a network, consisting of an input, dense/fully connected hidden layers, and the output as estimated by the model. 29

4-5 Simplified representation of the network structure, assuming uniform features within each layer and the components. 30

4-6 Classical optimization pipeline for NOPT 33

4-7 CVXPY main scheme. NN-based models are employed within an backpropagation scheme introducing the CVXPY layer. 34

4-8 Regularized physics-informed neural network for data-driven nonlinear constrained optimization main sketch 35

4-9 Adult Fly Population Model Diagram: The figure displays the random behavior in the fly population growth over time, graphically summarizing the dynamics of Equation 4-18. 36

4-10 Uniform mixture model optimization. Left: weighted uniform probabilities. Right: visual representation of the mixing results 37

4-11 RPINN pipeline for the uniform mixture model-based NOPT 38

4-12 RPINN pipeline for the adult fly population model-based NOPT 39

4-13 Blowfly population model MAPE results. Left: output error. Right: weights error 41

4-14 Uniform mixture model MAPE results. Left: output error. Right: weights error. (N): noisy-free. (-1),(3), and (5) stand for the SNR value 41

4-15 RPINN uniform mixture model-based NOPT results. First row: SNR= -1. Second row: SNR= 3. Third row: noise-free. Left: output prediction. Right: weight distribution. Green: target. Red: noisy target. Black: RPINN. Blue: SCS 42

4-16 RPINN vs. Solver: The graph compares solution times for the test data between the classical technique (CLARABEL and SCS, in blue) and our strategy (RPINN, in green). On the left, training times are shown, while on the right, prediction times are displayed. The first and second rows correspond to the adult fly population model and the uniform mixture model, respectively. 43

5-1 8-node network: The diagram depicts nodes as dots. Arrows symbolize gas pipelines, denoting the flow direction or connectivity between nodes. Trapezoidal shapes represent compressors. 48

5-2 15-node network: The diagram illustrates the expansion from the 8-node network, forming a synthetic network of 15 nodes. 48

5-3 Colombian Network: The diagram displays the geographical distribution of Colombia's network along with its elements. 49

5-4 Loss Functions: The diagram summarizes the loss functions used to train the PINN, which are aimed at ensuring compliance with the restrictions of the balance of the nodes, Weymouth and the limitations of the compression ratio. 52

5-5 Activation functions: The diagram illustrates the suggested functions that convert an input x into $\bar{x}$ to comply with the supply, unsupplied gas constraints, and the physical limitations of the variables. 54

5-6 Enhanced PINN for NLP-based Gas System Optimization: The figure displays the overall scheme of the proposed approach to tackle the problem discussed in this chapter, highlighting in green the two strategies designed to deal with the constraints of the NLP problem in gas transportation. 55

5-7 RPINN pipeline for the gas-powered system-based NOPT 55

5-8 Colombian Network: The diagram shows in green the elements used to validate our proposal in the network of Colombia with their respective geographical distribution 57

5-9 The chart presents the results of performance evaluation in three critical aspects. From left to right, the first image shows the performance of the three loss functions when evaluating the balance of nodes. The second image shows the result of the Weymouth equation, and the third, the costs of the solutions. 58

5-10 The figure illustrates the behavior of dynamic activation functions in a simulated scenario of oscillating demand. In the figure on the left, the behavior of the injection flow is observed when the source capacity is exceeded, and how unsupplied demand ceases to be zero in this context. On the other hand, to the right, the sum of supply flows with and without the proposed dynamic function is presented. 59

5-11 Gas-powered system regularized loss illustration. Left: node balance and Weymouth penalties based on conventional Huber-loss. Middle: Compression factor limit constraint using our Huber-based enhancement. (see Eq. 5-15). Right: Gas-powered system custom penalty evolution (Blue: Weymouth equality constraint; Orange: compression ratio limit constraint; Green: Compression factor constraint). 60

5-12 Gas-powered system objective cost and constraint compliance MAPE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function) between RPINN and IPOPT. 61

5-13 Gas-powered system objective cost and constraint compliance MAE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function) between RPINN and IPOPT. 62

5-14 Gas-powered system bound constraint MAPE results. The star symbol on this graph denotes the defined limits for each of the sources, compressors, pipelines, and pressures as well as their behavior. The number on the x-axis indicates the node to which the information belongs. MMSCFD: Million standard cubic feet per day. psia: pounds per square inch absolute. 63

5-15 The graph compares the samples that exceed 1×10^0 in our strategy. The data that remain below the network's capacity are represented in blue. In orange, the data where 50% of the samples exceed the capacity are shown, and in green, those that completely surpass it. On the left, the results of the node balance are presented, while on the right, the corresponding ones to the Weymouth equation. 64

5-16 Illustration representing the occurrence of convergence failures in each gas network, with numerical values indicating the nodes of each network.The data that remain below the network's capacity are represented in blue. In orange, the data where 50% of the samples exceed the capacity are shown, and in green, those that completely surpass it. 64

5-17 Gas-powered system objective cost and constraint compliance MAE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function). 65

5-18 RPINN vs. IPOPT computational cost results. The graph compares solution times for the test data between the classical technique (IPOPT, in blue) and our strategy (RPINN, in green). On the left, the training times are shown, while on the right, the prediction times are displayed. 65

5-19 RPINN vs. IPOPT computational cost results. The graph compares solution times for the test data between the classical technique (IPOPT, in blue) and our strategy (RPINN, in green). 66

List of Tables

1-1 State-of-the-art solvers for optimization. (*) Except mixed-integer SDP. (**) Features available with the licensed version only 18

4-1 RPINN details for the uniform mixture model-based NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Param. #: number of trainable parameters. Total # of parameters: 30 39

4-2 RPINN details for the adult fly population model-based NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Param. #: number of trainable parameters. Total # of parameters: 3330 39

5-1 RPINN architecture details for the gas-powered system NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Source switching, unsupply gas switching, custom dense, and bounded dense stand for specific switching, limited, and scaled layers, as explained in Section 5.2. Param. #: number of trainable parameters. Total # of parameters: 12707 (49.67K6B) 56

Content

Acknowledgements **ii**

Abbreviations **i**

Resumen **iii**

Abstract **v**

List of Figures **ix**

List of Tables **x**

Content **xiii**

1 Preliminaries **1**

1.1 Motivation 1

1.2 Problem Statement 8

1.2.1 Convergence Issues in Nonlinear Optimization 9

1.2.2 Time Complexity Challenges for Nonlinear Optimization 10

1.2.3 Research Question 11

1.3 State of the Art 13

1.3.1 Classic Models 13

1.3.2 Deep Learning-based Methods 15

1.3.3 Relevant Optimization Solvers 17

1.3.4 Summary 18

2 Aims **19**

2.1 Aims 19

2.1.1 General Aim 19

2.1.2 Specific Aims 19

3 Outline and Contributions **20**
3.1 Supervised Regularized Physics-Informed Neural Network to Support Nonlinear Constrained Optimization . . . 20
3.2 Unsupervised Regularized Physics-Informed Neural Network to Support Nonlinear Constrained Optimization . . . 20
3.3 Thesis structure . . . 22

4 Supervised Regularized Physics - Informed Neural Network to Support Nonlinear Constrained Optimization **23**
4.1 Theoretical Background . . . 23
4.1.1 Automatic Differentiation . . . 23
4.1.2 Backpropagation Algorithm for Neural Networks Training . . . 29
4.1.3 Nonlinear Optimization Fundamentals (NOPT) . . . 32
4.1.4 Convex Optimization in Python CVXPY . . . 32
4.1.5 Regularized Physics-Informed Neural Network (RPINN) . . . 34
4.2 Tested Scenarios for NOPT using RPINN . . . 35
4.3 Experimental Set-Up . . . 38
4.3.1 Deep Learning Architectures . . . 38
4.3.2 Training Details and Method Comparison . . . 39
4.4 Results and Discussion . . . 40
4.4.1 Adult Fly Population Model Results . . . 40
4.4.2 Uniform Mixture Model Results . . . 40
4.4.3 Computational Cost Results . . . 42
4.5 Summary . . . 43

5 Unsupervised Regularized Physics - Informed Neural Network to Support Nonlinear Constrained Optimization **45**
5.1 Theoretical Background . . . 45
5.1.1 Activation Functions . . . 46
5.1.2 Loss Function . . . 47
5.1.3 Gas Distribution Networks . . . 48
5.1.4 Mathematical Model . . . 49
5.2 Tested Scenarios for NOPT using RPINN . . . 50
5.2.1 Activation Function for Modeling Gas Supply . . . 53
5.2.2 Activation Function for Modeling Unsupplied Gas . . . 53
5.3 Experimental Set-Up . . . 54
5.3.1 Deep Learning Architectures . . . 54
5.3.2 Training Details and Method Comparison . . . 56
5.3.3 Synthetic Gas Networks . . . 57

5.4 Results and Discussion . . . 58
5.4.1 Loss Function Selection . . . 58
5.4.2 Dynamic activation functions . . . 58
5.4.3 Unsupervised Constrained Optimization Results . . . 59
5.4.4 Convergence . . . 61
5.4.5 Computational Cost Results . . . 62
5.5 Summary . . . 67

6 Final Remarks **68**
6.1 Conclusions . . . 68
6.2 Future Work . . . 69
6.3 Academic Products . . . 69
6.3.1 Academic Discussion . . . 69
6.3.2 Software and Repositories . . . 69

References **70**

1 Preliminaries

1.1 Motivation

The global shift towards more sustainable and efficient energy production methods is notably reflected in the 30.44% increase in electricity generation from natural gas. This increase is attributed to its high efficiency, reduced carbon emissions, and the lower costs associated with the rise of fracking [*Vásquez Cordano & Zellou*, 2020]. This transition is further complemented by the growth of renewable energy sources, which are becoming essential components of the energy mix, as illustrated in Figure **1-1** [*Ritchie & Rosado*, 2020]. Concurrently, there is a reduction in reliance on fossil fuels and air pollution, driven by the development of technologies such as photovoltaic and concentrated solar power [*Chen et al.*, 2020a]. In this context, the integration of gas turbine technology emerges as a key strategy to reduce both emissions and operational costs [*Hasanzadeh et al.*, 2021].

At the legislative level, initiatives such as the Inflation Reduction Act in the U.S., the Fit for 55 package in the EU, and the Climate Change Act in Australia reflect governmental commitment to the energy transition [*International Energy Agency (IEA)*, 2023, *Van de Graaf & Colgan*, 2016]. However, a significant portion of future energy investments may still be directed towards traditional generation methods. Nonetheless, a rise in the demand for fossil fuels by 2030 is anticipated, alongside increasing investment in renewable sources, particularly solar and wind [*International Energy Agency (IEA)*, 2022, *McKinsey & Company*, 2023].

In Latin America, despite the predominance of fossil fuels in electricity generation, extensive use of hydropower positions the region as a leader in low carbon emissions [*Pérez-Cedeño et al.*, 2022], as shown in Figure **1-1**. However, dependence on imported fossil sources is intensified by factors such as carbon inefficiency and the interdependence of electrical and natural gas systems, increasing the risk of cascading failures [*Chen et al.*, 2023]. In Argentina, the transportation sector significantly contributes to CO2 emissions and energy consumption [*Muñoz et al.*, 2022], highlighting broader regional challenges. While the transition to more sustainable options is vital, it faces institutional and infrastructure obstacles, lagging the region behind global advances and underscoring the urgent need for innovative investments

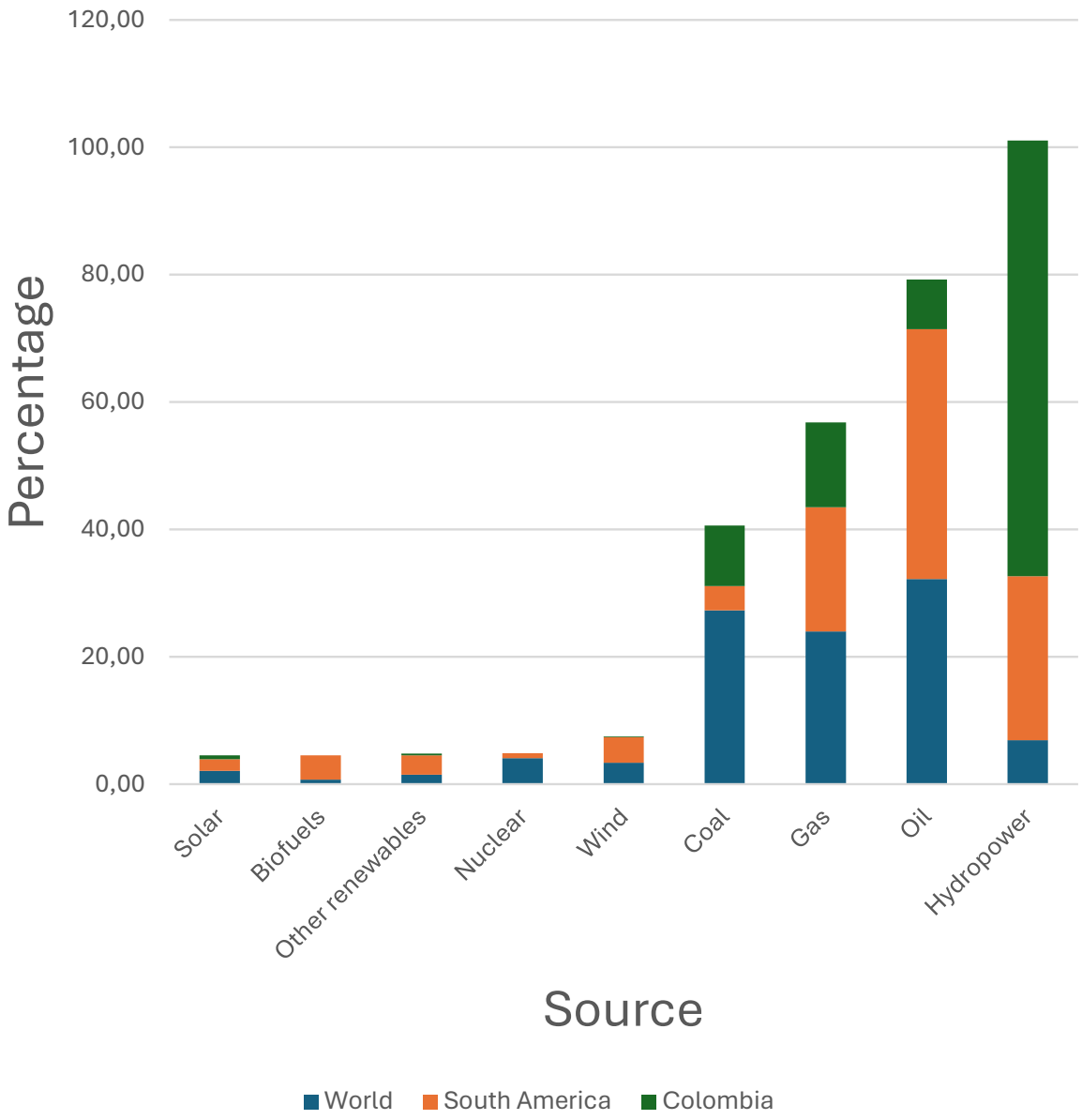

Figure 1-1: Colombia's energy mix in perspective: The graph contrasts Colombia's heavy reliance on hydropower with global and South American energy trends.

and strategies to enhance energy efficiency [*Harrington*, 2018]. The growing vulnerability of energy systems also underscores the critical importance of strengthening infrastructure and diversifying sources, imperative for the long-term resilience and sustainability of the region [*Li et al.*, 2022, *Beyza et al.*, 2020].

In Colombia, as in many Latin American countries, hydroelectric power has historically been the primary energy source, significantly complemented by fossil fuels (see Figure **1-1**). This configuration has generated 1330 Exajoules with almost zero CO2 emissions, demonstrating

a strong commitment to sustainability and renewable sources [*Pérez-Cedeño et al.*, 2022]. Recently, however, there has been an increasing diversification in the energy matrix, with notable growth in the adoption of renewable energies such as wind and solar [*González et al.*, 2021]. Entities such as the Ministerio de Energía y Minas, Unidad de Planeación Minero Energética(UPME), and Comisión de Regulación de Energía y Minas (CREG) play key roles in overseeing and managing the energy industry [Ene, 2023]. These efforts reflect the country's focus on innovation and the integration of advanced technologies to enhance energy efficiency and reliability [*Harrington*, 2018].

However, the sustainable development model of Kuznets presents challenges for Colombia, such as the paradoxical increase in CO2 emissions with economic growth, exacerbating environmental impact [*Sosa & Navarro*, 2020]. This underscores the importance of balancing development and sustainability, which is crucial for advancing towards cleaner manufacturing practices and achieving the country's sustainable development goals [*Gonzalez & Peña-Vinces*, 2023].

Colombia's reliance on water resources, exacerbated by phenomena such as El Niño, which induces droughts and raises temperatures, presents a significant risk, given that approximately 70% of its energy comes from hydroelectric sources [Cor, 2023]. The unpredictability during the dry season, even with optimal reservoir levels, and delays in wind and solar projects due to environmental issues and community consultations, highlight the urgency of diversifying energy sources [REA, 2023]. This situation underscores the importance of exploring alternative energy sources, such as the use of hydrocarbons in thermoelectric plants, where natural gas stands out for its lower environmental impact and affordability [*Kuhns & Shaw*, 2018, *Greyson*, 2021].

Despite a 40% increase in gasoline prices since October 2022, natural gas has seen a more moderate increase of 27.39% up to April 2023 [*Unigas*, 2023], maintaining it as a viable option. Integrated gas-electricity energy systems, illustrated in Figure **1-2**, are essential for sustainable energy, allowing for the incorporation of renewable sources and efficient resource allocation [*Greyson*, 2021]. Research highlights the importance of meticulous management of these networks and the need for precise flow calculations to improve safety and reduce costs [*Fallahi & Maghouli*, 2020, *Chen et al.*, 2022, *Sun et al.*, 2022, *Yu et al.*, 2023]. The Ministerio de Energía y Minas, along with UPME, is committed to the safe, sustainable, and efficient supply and transportation of natural gas, implementing regulations to improve and expand transportation infrastructure and establishing strict criteria for investment projects [*UPME*, 2021b, *UPME*, 2021a, *Ministerio de Minas y Energía*, 2021, *Ministerio de Minas y Energía*, 2020].

The UPME has achieved significant annual savings, recording 346.04 terajoules per year thanks to the strategies implemented in the Programa de Evaluación Industrial (PEVI) [*UPME*, 2019]. This achievement is attributed to a 25.14% reduction in natural gas consumption and a 63.11% decrease in electricity usage [*UPME*, 2022a]. Additionally, UPME has established standards and procedures that expand the range of goods and services eligible for tax incentives, following international technical standards [*UPME*, 2022c].

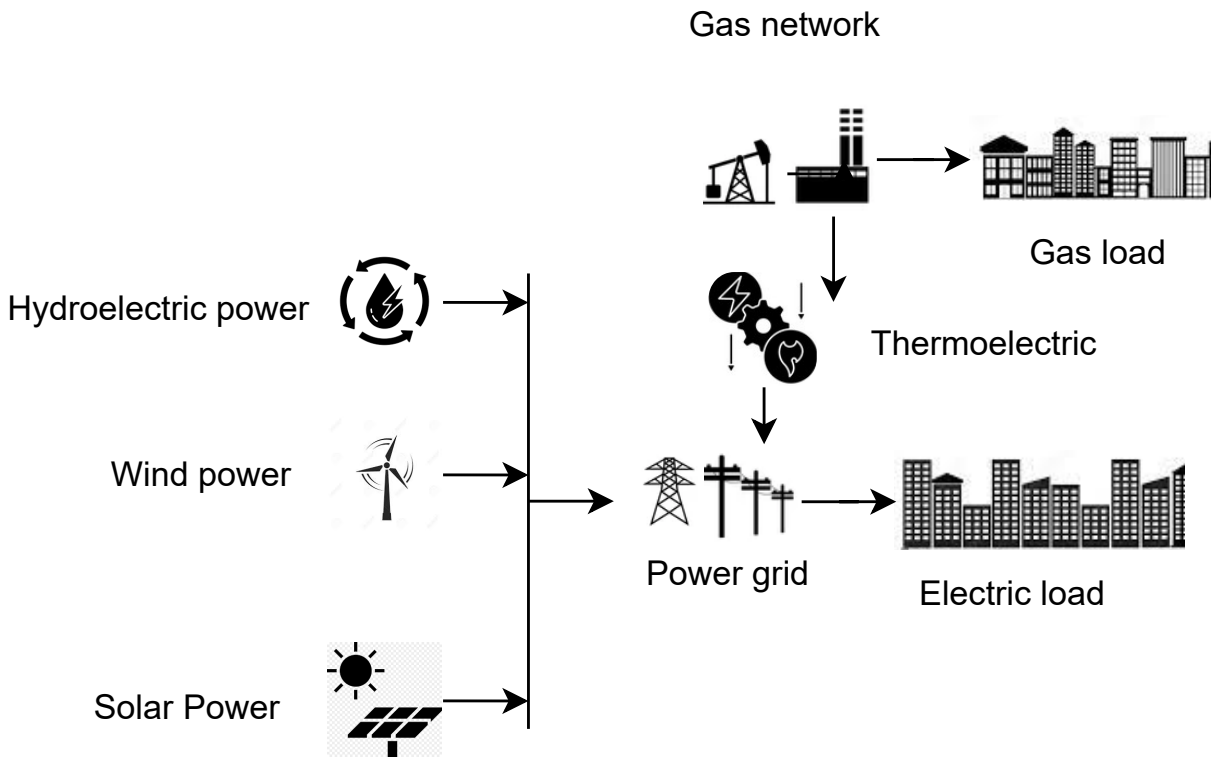

Figure 1-2: Colombian Power Generation Model: This figure illustrates the overall pattern of Colombia's integrated gas-electricity energy system, highlighting its crucial role in energy sustainability.

Meanwhile, the CREG plays a crucial role in regulating renewable energies, including small and large-scale self-generation, as well as integrating wind and solar energy into the wholesale energy market. This entity is also responsible for defining the conditions for integration into the Sistema Interconectado Nacional(SIN) [*CREG*, 2022b, CRE, 2023].

These initiatives have significantly improved the efficiency and sustainability of Colombia's energy framework. In particular, CREG also focuses on optimizing the flow of natural gas, considering its non-renewable nature and a sustainability projection of reserves for 7 years [*ANH*, 2022]. The Linear Programming model illustrated in Figure **1-3** has been crucial in achieving these objectives, allowing significant cost savings and offering considerable economic benefits. It is estimated that its application could reduce annual expenses by 700 billion pesos, increase consumer surplus by 6.2 trillion pesos, and reduce domestic gas prices in Colombia by 14%, highlighting its economic impact on the country [*CREG*, 2022a]. All these actions align with Colombia's commitment to transforming towards a more sustainable and eco-friendly energy system [*UPME*, 2022b].

However,in the context of gas distribution networks, traditional mathematical models often overlook critical factors such as pressure variations and the role of compression stations [*CREG*, 2022a]. This omission is significant since the pressure difference is the primary driver

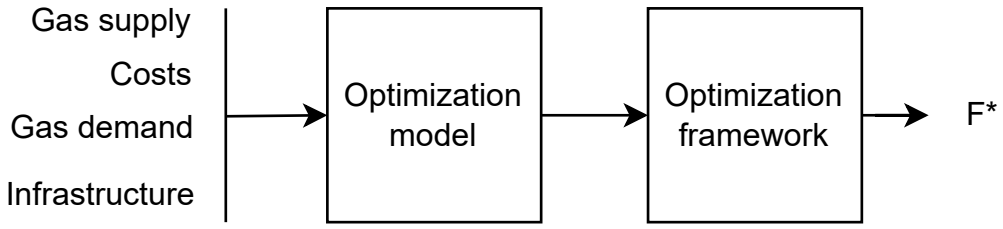

Figure 1-3: Optimization Framework: This diagram illustrates the strategy employed by the Energy and Gas Regulatory Commission (CREG) to optimize gas system operations, where the optimal flows are denoted by F^*.

of gas flow in pipelines [*Schreider et al.*, 2015]. This gap highlights the need for a more comprehensive optimization model that incorporates these essential elements to achieve a more accurate simulation of real-world gas system conditions [*Chen et al.*, 2020b, *Yu et al.*, 2021]. Incorporating flow dynamics and pressure conditions, however, introduces considerable complexity. While traditional optimization methods are fundamental for cost evaluation and its implications across various areas, they face significant challenges, including practical inaccuracies, susceptibility to suboptimal solutions, and difficulties in handling complex and nonlinear systems [*Hu & Zhang*, 2023]. Figure **1-4** illustrates how conventional optimization methods address these challenges.

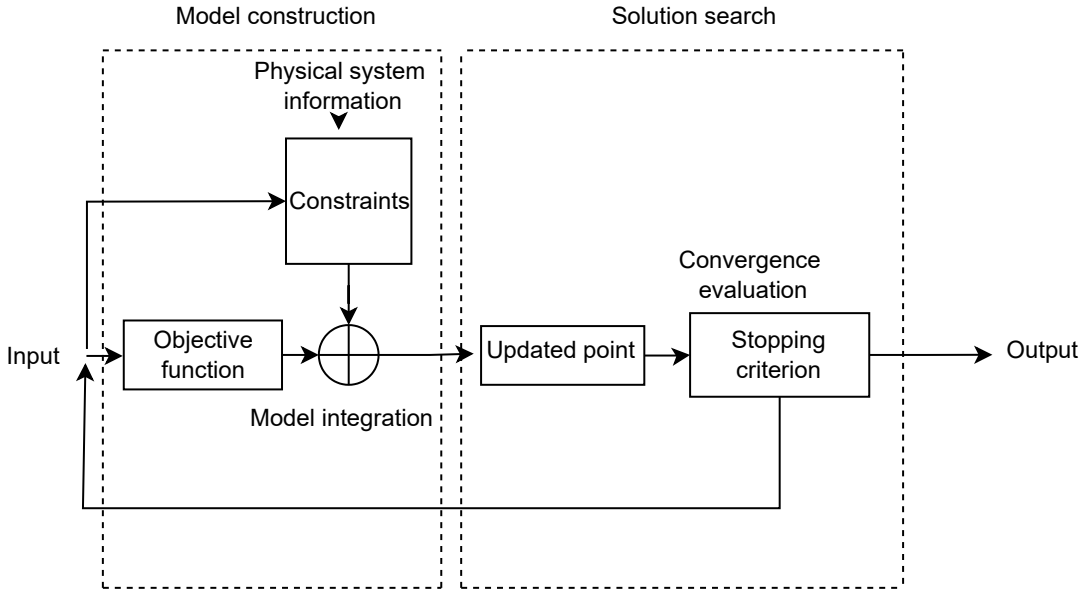

Figure 1-4: Classical Optimization Model: This figure depicts the general framework of a traditional optimization method aimed at finding a solution to a problem, emphasizing the systematic approach from the initial point to the optimal outcome.

Given these obstacles, the importance of machine learning and other Artificial Intelligence (AI)-based alternatives is progressively being acknowledged. These technologies leverage past data to enhance efficiency and decision-making processes [*Park et al.*, 2023]. Neural networks are particularly effective in these situations: they can accurately discover complex connections in high-dimensional spaces, resulting in solutions that are faster and more precise [*Chevalier & Chatzivasileiadis*, 2023]. Their ability to provide quick and accurate predictions of certain occurrences is well-known for improving operational efficiency[*Nellikkath & Chatzivasileiadis*, 2022a].

Moreover, these networks have demonstrated expertise in effectively handling operational unpredictability and security issues in electrical systems, specifically regarding the incorporation of renewable energy sources [*Zhang et al.*, 2023]. Their advantage arises from combining physical insights with machine learning to provide novel and secure approaches in scenarios with incomplete rules and ambiguous boundaries in energy systems, particularly when faced with limited computing resources. Figure **1-5** illustrates how deep learning models tackle an optimization problem, offering a practical depiction of these concepts.

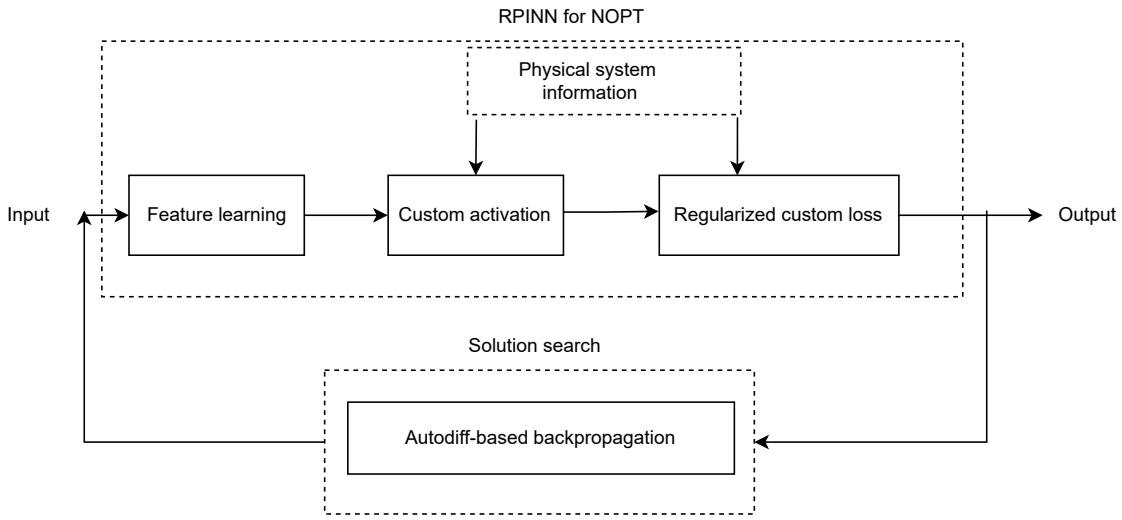

Figure 1-5: Deep Learning Optimization Model: This figure illustrates the general framework of a deep learning model applied to an optimization problem, highlighting the process from data input to the iterative search for an optimal solution.

In a local context, the Grupo de Control y Procesamiento Digital de Señales GCPDS at Universidad Nacional de Colombia has been developing AI models to implement advanced technologies in various research domains, aiming to support or enhance current solutions. As evidence of the group's scientific achievements, multiple research and innovation projects are highlighted, underpinned by Minciencias, the Dirección Nacional de Investigaciones de Manizales, and the Vicerrectoría de Investigaciones at the Universidad Nacional de Colombia:

- Desarrollo de un sistema automático de análisis de volumetría cerebral como apoyo en

la evaluación clínica de recién nacidos con asfixia perinatal (2019-actual). Financiado por Minciencias.

- Desarrollo de una herramienta de seguimiento de aguja y localización de nervios en ecografía para la práctica de anestesia regional: aplicación al tratamiento de dolor agudo traumático y prevención del dolor neuropático crónico (2019-actual). Financiado por Minciencias

- Caracterización morfológica de estructuras cerebrales por técnicas de imagen para el tratamiento mediante implantación quirúrgica de neuroestimuladores en la enfermedad de Parkinson (2019-actual). Financiado por Minciencias

- Desarrollo de un sistema de apoyo al diagnóstico no invasivo de pacientes con Epilepsia fármaco-resistente asociada a displasias corticales cerebrales: método costo-efectivo basado en procesamiento de imágenes de resonancia magnética (2017-2019). Financiado por Minciencias.

Recently, the GCPDS has expanded its research boundaries by forming a partnership with the Universidad Tecnológica de Pereira. This collaboration is dedicated to creating an innovative tool as part of the project 'Desarrollo de una herramienta para la planeación a largo plazo de la operación del sistema de transporte de gas natural de Colombia' (Code 111085269982), supported by MINCIENCIAS. The tool explores the utilization of AI techniques that incorporate physical knowledge, aiming to enhance and refine optimization models. Specifically, these models are tailored to predict optimal flows in gas networks utilized for electricity generation in Colombia.

In light of the foregoing, this master's thesis stands at the critical intersection of technological innovation and environmental mandates, amidst a global shift towards more sustainable and efficient energy sources. It emphasizes the imperative need to adapt energy systems to renewable energies and the advancement of gas-based technologies. With a specific focus on Latin America, particularly Colombia, the thesis aims to navigate towards a more varied and environmentally friendly energy matrix. This transition faces unique challenges, such as the strong dependence on hydroelectric power, the urgent need to reduce carbon dioxide emissions, and the crucial task of strengthening energy infrastructure to ensure its resilience and long-term viability. In this context, the adoption of open-source tools is key, fostering collaboration, transparency, and adaptability in technological solutions, and driving innovation and knowledge to meet environmental and energy challenges. Therefore, the research presented here proposes the development of AI based approaches to enhance the management and operation of gas systems, focusing on improving efficiency and sustainability in the face of these challenges, underscoring the importance and necessity of integrating open-source tools in this process.

1.2 Problem Statement

Effective integration of gas and energy systems necessitates sophisticated optimization models that not only reduce expenses and facilitate decision-making but also accurately capture complexities. State agencies such as UPME and CREG employ a linear optimization model for managing flow and expenses [*CREG*, 2022a, *CREG*, 2022c]. However, this model overlooks essential elements like pressure fluctuations in the gas network and the function of compression stations, which are critical as they propel the flow in pipelines [*Schreider et al.*, 2015]. Consequently, there is a pressing requirement for an enhanced optimization model that encompasses these components and precisely reflects the actual conditions of the network [*Chen et al.*, 2020b, *Yu et al.*, 2021]. Nevertheless, the incorporation of pressure and flow dynamics scenarios introduces considerable complexity, resulting in a Nonlinear Optimization NOPT problem due to the nature of the equations [*Farrokhifar et al.*, 2020, *Klatzer et al.*, 2022, *Wu et al.*, 2018].

Addressing nonlinear challenges presents significant obstacles, particularly in the context of non-convex functions and integer variables, which are common in complex and extensive problems. These functions hinder the application of traditional optimization methods due to their multiple constraints and assumptions [*Arya et al.*, 2022]. Additional complications include exceeding the maximum number of allowed iterations, the absence of feasible solutions, and the inherent difficulties of unbounded or non-convex dilemmas [*Mai & Mortari*, 2022]. Non-convex inequality constraints, in turn, can lead to infeasible and suboptimal results, significantly increasing the difficulty of handling these problems [*Mhanna & Mancarella*, 2021]. Moreover, certain methodologies exhibit a slow convergence rate and a tendency to stall in local maxima, limiting their effectiveness in scenarios where efficiency and reliability are crucial [*Chang et al.*, 2022].

In the complex context of NOPT, the design of gas pipeline networks presents notable challenges, including the management of uncertainties in model parameters, handling complex equations, and selecting optimal installation locations. These challenges are magnified by the need to simultaneously address both continuous and discrete aspects in the design and operation of these networks [*Arya*, 2022]. It is crucial to recognize that optimization models for natural gas networks, given their high nonlinearity, have traditionally been limited to smaller scales. This limitation underscores the importance of adapting NOPT to larger and more complex infrastructures, an adaptation that must aim to minimize execution times and simplify the inherent complexity of these problems [*Wang et al.*, 2023]. This involves mitigating nonlinearities, addressing computationally demanding NP-hard challenges, and developing more effective approximate methods [*Robuschi et al.*, 2021]. It is important to note that the use of these approximations can increase processing time and require specialized software, thereby increasing the risk of errors in estimating solutions [*Montoya et al.*, 2020a].

Given the imperative to enhance optimization models for gas and energy systems, this study addresses two key limitations: i) the issue of non-convergence in nonlinear models [*Schreider*

et al., 2015, *Arya et al.*, 2022], and ii) the extended computation times necessitated by iterative techniques [Kim *et al.*, 2023]. These subjects are thoroughly investigated in the subsequent sections. Figure **1-6** presents an overview of these factors and their consequences in nonlinear optimization challenges, particularly impacting convergence and computational processing time.

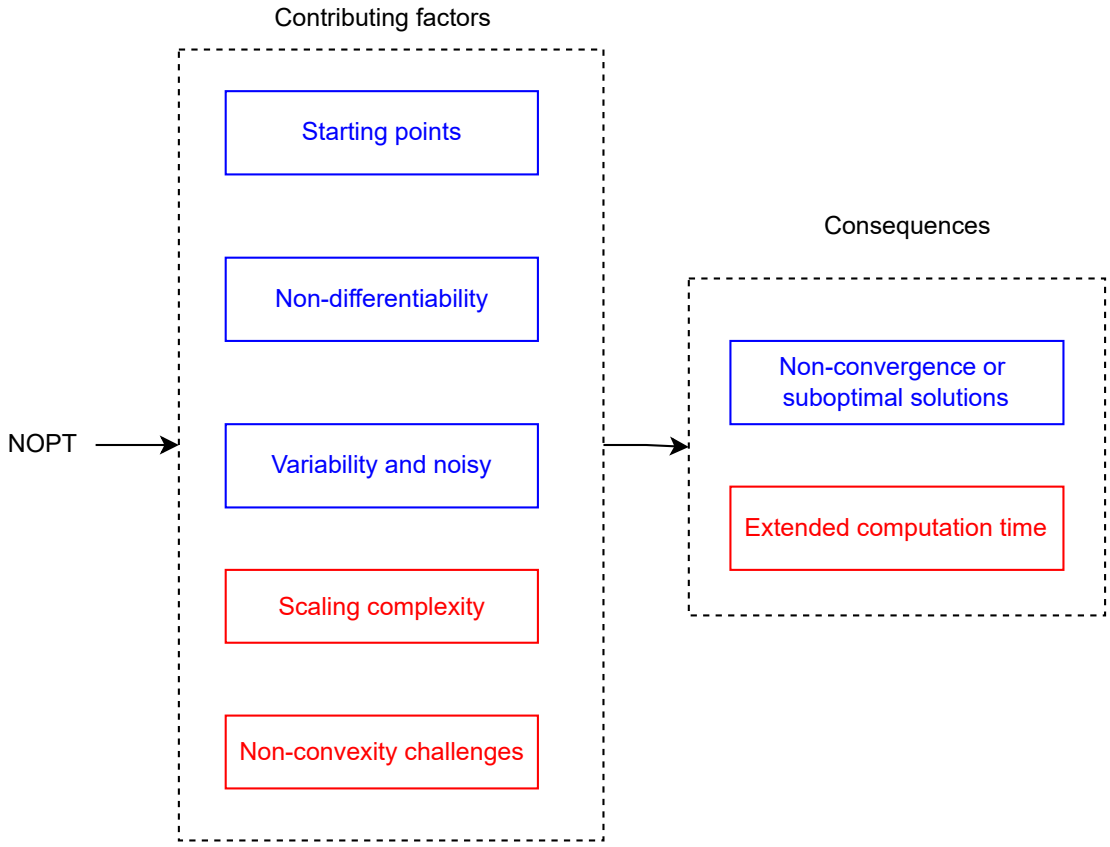

Figure 1-6: Problem Overview: This diagram illustrates the primary factors affecting nonlinear optimization challenges in gas pipeline network design. Factors leading to non-convergence are indicated in blue, while those contributing to extended computational time are marked in red.

1.2.1 Convergence Issues in Nonlinear Optimization

NOPT problems face significant challenges due to convexity and a tendency to converge towards suboptimal solutions [*Schreider et al.*, 2015]. These challenges arise from factors such as the choice of starting points, non-differentiability, and the variability of data and

noisy input measurements. Specifically, traditional iterative methods critically depend on starting points, also known as "warm starts" [*Huang et al.*, 2023], significantly influencing the accuracy and quality of solutions. Selecting an appropriate starting point is, therefore, essential for effectively solving optimization problems [*Han et al.*, 2024].

Additionally, conventional approaches require mathematical models to meet certain essential conditions, such as differentiability [*Cotter et al.*, 2019]. Some algorithms need second-order derivative information [*Gao et al.*, 2022], a requirement often impractical, compromising problem-solving accuracy. Traditional methods also face limitations due to their sensitivity to initial conditions and parameter selection [*Han et al.*, 2024], factors that can considerably affect algorithm efficiency, disrupting the optimization process and making it susceptible to external disturbances [*Kohjitani et al.*, 2022]. Figure **1-7** illustrates the factors contributing to non-convergence in nonlinear optimization problems.

Given the importance of efficient optimization in critical sectors like energy, where strategic decisions heavily rely on the reliability of the models used, the relevance of these challenges is increasingly significant, as highlighted in recent discussions on the impact of optimization in the energy sector [*CEOWORLD*, 2022]. Furthermore, the Energy Information Administration (EIA) has indicated that effective distribution and management are crucial for the sustainable use of natural gas [*EIA*, 2022]. In Colombia, improving gas pipeline networks emerges as a critical issue; inadequate models not only compromise supply sustainability and reliability but also the efficiency of supply chain management [*Energy5*, 2024]. This context underscores the critical need to develop robust and accurate optimization models specifically tailored to the needs of the Colombian gas network.

1.2.2 Time Complexity Challenges for Nonlinear Optimization

The "No Free Lunch" theorem, formulated in [*Wolpert & Macready*, 1997], highlights the intrinsic challenges in NOPT. This principle reveals that any performance improvement of an algorithm for certain types of problems usually results in reduced effectiveness for others, underscoring the absence of a universal solution. In this context, while traditional methods are accurate, they face progressively longer processing times as the problem dimension increases [*Böttcher et al.*, 2023]. Additionally, NOPT problems are particularly challenging due to their intense demand for computational power and time, restricting their practical applicability [*Owerko et al.*, 2020]. These challenges, especially in NOPT problems, pose significant barriers due to convergence difficulties and extensive calculation durations [*Nellikkath & Chatzivasileiadis*, 2022b]. The need for innovation is palpable, urging the creation of efficient and specialized methods to effectively handle the specific complexities of these challenges. Figure **1-8** synthesizes the main factors contributing to prolonged computation times in nonlinear optimization.

In the Colombian context, addressing these challenges in NOPT is crucial for advancing

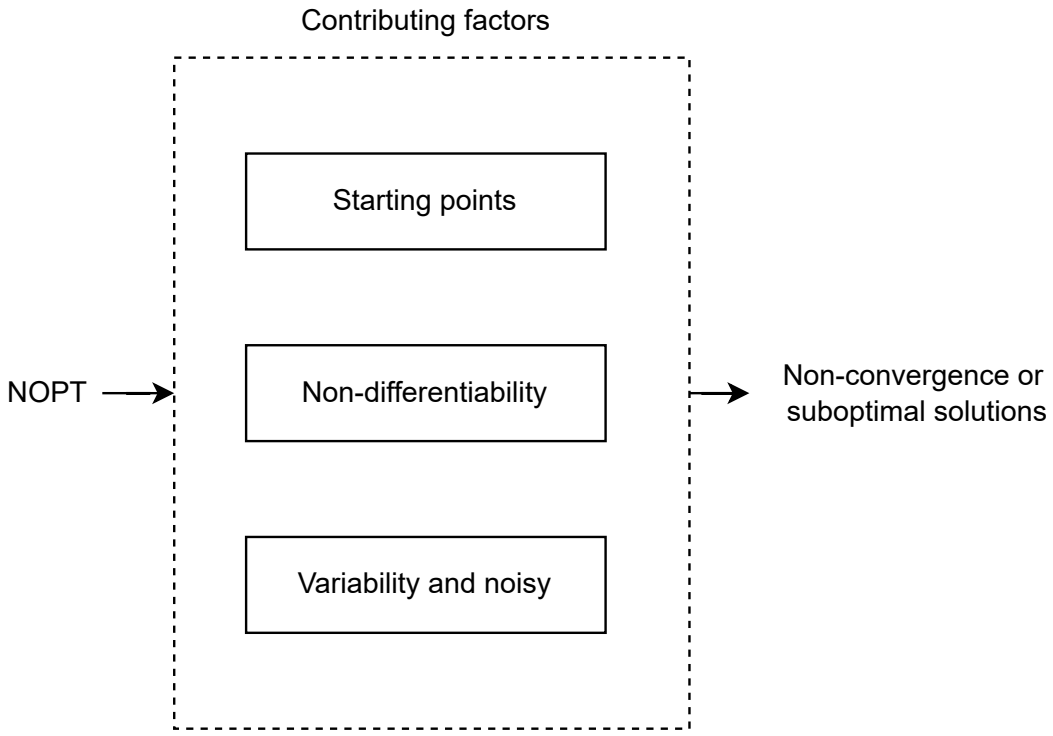

Figure 1-7: Convergence issues in nonlinear optimization: The figure summarizes the factors contributing to the non-convergence in NOPT problems.

the global energy sector. Prolonged processing times constitute a significant obstacle to the development and implementation of renewable energy solutions, essential for Colombia's energy transition goals [Uni, 2023]. These delays not only hinder progress towards sustainability targets but also complicate the integration of these technologies into the national grid, highlighting the importance of developing more efficient and optimized energy management strategies [*Impacto TIC*, 2023]. Overcoming these obstacles is vital for strengthening the country's energy resilience and improving operational efficiency, emphasizing the need to refine optimization processes as Colombia moves towards a more sustainable energy system [*Energy Master*, 2023].

1.2.3 Research Question

The review of relevant literature reveals an urgent need for innovation in developing optimization strategies that integrate conventional methods with advanced artificial intelligence techniques, utilizing PINN. The aim of these approaches is to enhance the modeling of gas pipeline networks, considering nonlinear constraints such as flow dynamics and pressure

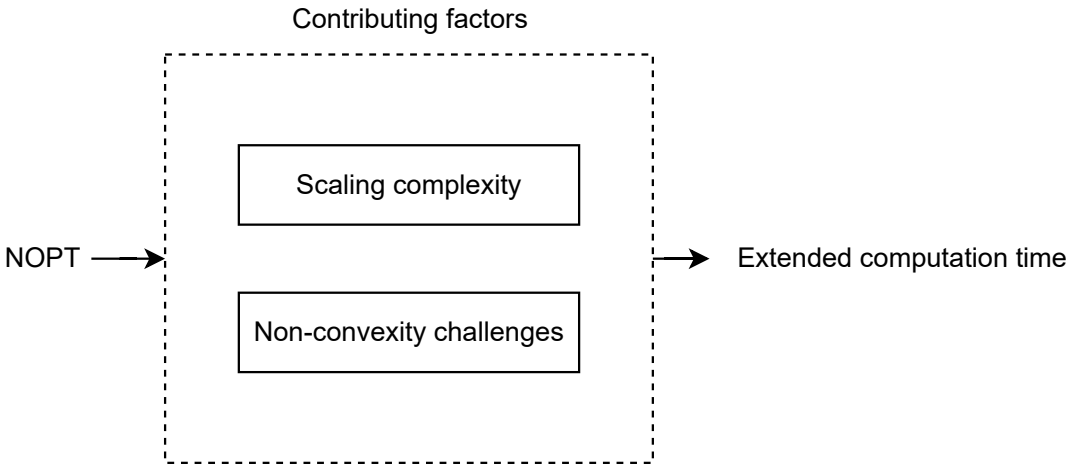

Figure 1-8: Time complexity challenges for nonlinear optimization: The figure presents a summary of factors leading to extended computation times in nonlinear optimization.

conditions, which are prominent in nonlinear and non-convex optimization challenges. Despite these advancements, significant obstacles related to convergence and computational efficiency persist, crucial factors in decision-making for the management and distribution of natural gas. These challenges are particularly critical in Colombia, where the integration of renewable energies and the efficiency of the national gas network are priorities.

This landscape poses an essential research question: How can an AI-based strategy, employing PINN, be developed to integrate nonlinear constraints, optimize convergence, and reduce computation times, thus achieving effective and economical solutions in the optimization of gas energy systems, ensuring accurate estimation of flows and pressures?

1.3 State of the Art

1.3.1 Classic Models

Some studies have shown that mathematical programming has become a crucial tool in numerical optimization. A notable example is the analysis by [*Mhanna & Mancarella*, 2021], which employs a sequential linear programming algorithm to address nonlinearities and discontinuities. In this context, the simplex method proves essential, being a classic technique effective for solving linear programming problems through iterative adjustments of solutions within a feasible set [*Applegate et al.*, 2021]. Similarly, the study by [*Sadat & Sahraei-Ardakani*, 2021] explores a solution via quadratic programming (QP). Mixed-integer programming (MIP), on the other hand, is an optimization strategy that uses both integer and continuous variables. It is widely used to solve difficult problems [*Zhao et al.*, 2020], focusing on how the branch-and-cut (BC) algorithm can be employed to find the best solution [*Vo et al.*, 2023]. Furthermore, second-order cone programming (SOCP) facilitates effective solutions for problems involving linear and quadratic constraints [*Sun et al.*, 2020]. New studies, like [*Lin et al.*, 2022b], look into semidefinite programming (SDP), and the work in [*Chowdhury & Kamalasadan*, 2021] uses convexification techniques. Likewise, exponential programming (EXP), which models NOPT objectives and constraints through exponential functions [*Asgharieh Ahari & Kocuk*, 2023]. Additionally, power cone programming (PCP) is considered for modeling product and square relationships [*Kumar & Rahaman*, 2020]. Yet, these classical methods face challenges such as scalability, computation time, convergence, and practical precision, underscoring their inherent complexity and limitations. Furthermore, the use of relaxations or approximations affects the optimization accuracy [*Lin et al.*, 2022a].The summary of this discussion can be seen in Figure **1-9**

Figure 1-9: Mathematical programming-based methods. Red: main drawbacks. Blue: reference.

On the other hand, gradient methods' efficiency in identifying optimal solutions highlight their relevance for practical optimization tasks. The work in [*Abubakar & Kumam*, 2019] uses the Dai-Liao conjugate gradient method and hyperplane projections for global convergence to solve nonlinear equations. In addition, [*Chen et al.*, 2021] faces the non-convex issue based on a set of starting points. Moreover, nonlinear decomposition using linear programming (LP) and gradient descent was also proposed [*Mahapatra & Rajan*, 2020]. Further, the work in [*Karimi et al.*, 2019] examined the Newton-based search to deal with convergence issues in poorly conditioned systems. Also, the semi-sweeping Newton technique was applied for optimization in Hilbert spaces [*Mannel & Rund*, 2021]. For noisy problems, the authors in [*Pinheiro et al.*, 2022] use piecewise polynomial interpolation and box reformulations, along with an interior-point (IP) method. Authors in [*Delgado et al.*, 2022] tackle similar problems with integrated penalty techniques. Overall, gradient methods are effective at solving NOPT tasks, but they have a challenging time convergent and are expensive to run in noisy and nonlinear situations [*Liu et al.*, 2021]. Besides, it can be challenging to choose the best learning rate, and they run the risk of finding local minima [*Haji & Abdulazeez*, 2021]. As seen in [*Ibrahim & Hossain*, 2021], it is also important to make sure that at least first-degree differentiation continuity is maintained when using techniques like the conjugate gradient, the IP, and the Newton-based approach. The summary of this discussion can be seen in Figure **1-10**

Figure 1-10: Gradient-based methods. Red: main drawbacks. Blue: reference.

In turn, decomposition methods have become essential tools for addressing the inherent complexities of large NOPT. These approaches break complex optimization into manageable subproblems, thus facilitating efficient resolution. One relevant example is the Benders decomposition, which solves models that are not convex by breaking them up into linear tasks [*Gao & Li*, 2021]. Similarly, in [*Attari & Torkayesh*, 2018] a mixed programming is used to tackle significant uncertainties and simplified assumptions. Another option is the dual decomposition [*Shi & Hong*, 2020] for adding penalties in non-convex optimizations. Falsone in [*Falsone et al.*, 2020] has also implemented this technique to distribute the process among multiple agents. Dantzig-Wolfe decomposition has been used to handle the increase in the number of variables and the resulting complexity [*Wang & Zhou*, 2023]. While decomposition methods have demonstrated their effectiveness, they may struggle to converge in noisy, nonlinear, and non-convex situations because they use linear approximations that

are obtained through truncated relaxation [*Füllner & Rebennack*, 2022].

Figure 1-11: Decomposition-based methods. Red: main drawbacks. Blue: reference.

Conversely, metaheuristics continue to evolve as key tools for efficiently balancing exploration and exploitation in various contexts. The work in [*Gao et al.*, 2019] has enriched the evolutionary algorithm with chaotic local search, noting the need for further research on the effectiveness of chaotic maps. Then, [*Wang et al.*, 2020b] has proposed improvements to overcome the limitations of the gravity search algorithm, such as low search efficiency and premature convergence. Comparative studies show that classical algorithms can be more effective if they know everything about the environment, even though metaheuristic approaches are competitive [*Ab Wahab et al.*, 2020]. In addition, metaheuristics can help with gradient-based methods, but they still have issues, such as local solutions [*Naderi et al.*, 2021]. Besides, they experience slow convergence, high temporal complexity, and failures in highly complex or noisy contexts [*Lee et al.*, 2023].

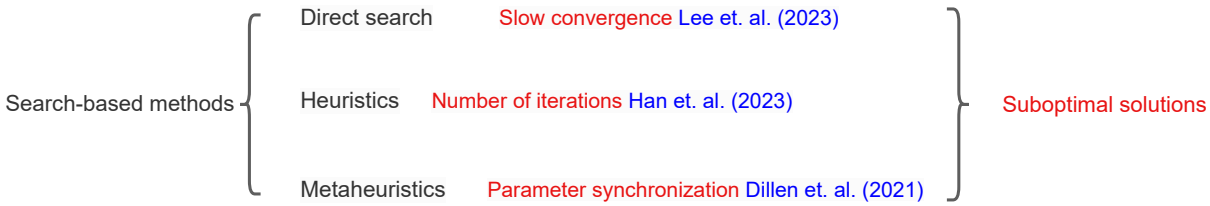

Figure 1-12: Search-based methods. Red: main drawbacks. Blue: reference.

1.3.2 Deep Learning-based Methods

Recently, ANNs have positioned themselves as fundamental tools in optimization by incorporating deep learning techniques, effectively addressing the complexity and non-linearities of various problems. Conventional ANNs employ supervised learning to tackle nonlinear and stochastic problems through regression tasks. To this end, historical data or solutions precomputed by specialized NOPT tools are used to train these networks [*Pan et al.*, 2021]. This approach enables ANNs to learn complex patterns and make accurate predictions even under significant uncertainty [*Jamii et al.*, 2022]. Typically, ANN-based approaches utilize AD, a computational method for efficiently and accurately evaluating function derivatives.

Instead of numerical or symbolic differentiation, which can have issues with accuracy and require a lot of computing power, AD breaks functions down into simple operations whose derivatives are known and uses the chain rule consistently [*Baydin et al.*, 2018]. Thereby, AD ensures machine-level accuracy in derivative calculations and simplifies the determination of loss function gradients in relation to network parameters, enabling the use of gradient-based search with back-propagation. The work in [*Baker*, 2020] combines quasi-Newton methods and ANNs for NOPT. Furthermore, the authors in [*Pan et al.*, 2021] utilize deep learning to solve optimal flow problems. Similarly, the work in [*Zhou et al.*, 2023] introduces an integrated training technique that, while effective, requires larger neural networks and presents challenges in generalization. Concurrently, [*Liang & Zhao*, 2023] uses elastic layers and incremental training as optimization-based solvers. Furthermore, the method by [*Falconer & Mones*, 2023] combines convex relaxation with graph neural networks.

Besides, PINN has recently emerged as a powerful optimization tool. These training approaches have proven effective in various NOPT applications, integrating relevant physical principles within ANNs [*Pan et al.*, 2023]. The KKT criteria are applied to formulate constraints that are incorporated into an ANN's cost function during supervised training [*Nellikkath & Chatzivasileiadis*, 2022c]. In [*Misyris et al.*, 2020], a PINN framework is detailed that imposes penalties for constraint violations in the loss function. The study in [*Misyris et al.*, 2021] proposes a loss function that combines errors from differential and algebraic states with normative equation violations. Additionally, a novel strategy has been proposed to include constraints in unsupervised training using an RK-based technique [*Huang & Wang*, 2023]. Nevertheless, complete approaches based on ANNs and RPINNs face challenges such as optimality degradation. In response, advanced alternatives like [*Habib & Yildirim*, 2022] have emerged, integrating system constraints into the cost function and applying penalties for violations. Furthermore, [*Yang et al.*, 2021] introduces an algorithm to address nonlinear problems modeled by partial differential equations with noisy data through Bayesian physics-informed neural networks (B-RPINNs). Additionally, [*Schiassi et al.*, 2021] proposes a parametric differential equation-based approach holding functional connections to enhance the robustness and accuracy of PINNs. In turn, [*Raynaud et al.*, 2022] presents a truncated Fourier decomposition, termed Modal-PINNs, to optimize the reconstruction of periodic signals. However, these alternatives often lack adequate precision, generalization capability, and scalability [*Nellikkath & Chatzivasileiadis*, 2022a]. Finally, supervised data is usually required, complicating their application in various NOPT scenarios. In a related field, [*Böttcher et al.*, 2023, *Owerko et al.*, 2020, *Owerko et al.*, 2022] have explored the use of Graph Neural Networks (GNN) to optimize power flow, although they encountered limitations in terms of accuracy, generalization, and scalability.The prior discussion regarding the time complexity challenges for nonlinear optimization issues is summarized in Figure **1-13**.

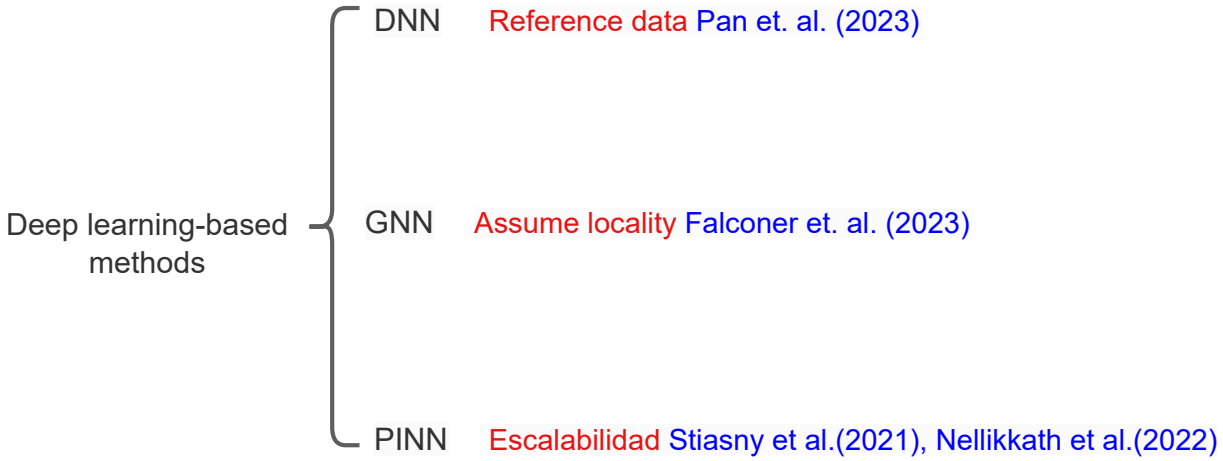

Figure 1-13: Summary of key deep learning techniques for nonlinear problem optimization, with drawbacks highlighted in red.

1.3.3 Relevant Optimization Solvers

Of note, most of the available optimization solvers are based on the classical approaches mentioned above. Among them, Clarabel stands out for its versatility in optimizing a wide variety of problems. However, it still faces significant challenges in areas such as MIP [*Goulart & Chen*, 2024]. Gurobi is renowned for its proficiency in MIP due to its extensive range of techniques, including simplex and IP methods. However, because it is proprietary software, it might not be able to be used in situations that require license flexibility [Gur, 2024]. Mosek is efficient concerning the IP approach, but its support for MIP is relatively limited, and its aptitude for NLP remains under debate, which could be a hindrance for developers who prefer open-source solutions [Mos, 2024]. Xpress specializes in solving MIP, offers conditional support for NLP, but is a closed-license alternative [FIC, 2024]. In turn, SCS, leveraging its open-source status, promotes adaptability and collaborative development, although its limitations in NLP reduce its effectiveness in certain optimization areas [*O'Donoghue*, 2021]. IPOPT excels at solving NLP problems, and its open access allows for flexibility [Ipo, 2024].

Now, in this multifaceted optimization environment, the integration of tools such as MATPOWER, GEKKO, and CVXPY significantly expands the available options. MATPOWER is essential for solving energy system issues and supports solvers like Gurobi and IPOPT for linear, mixed-integer, and nonlinear programming [*Zimmerman & Murillo-Sánchez*, 2020, *Wang et al.*, 2007, *García-Marín et al.*, 2022]. GEKKO specializes in dynamic systems and nonlinear models, offering a holistic and open-source Python platform [*Beal et al.*, 2018, *Mugel et al.*, 2022]. CVXPY is an open-source modeling language for convex optimization problems embedded in Python. It allows you to express your problems naturally, mirroring the

mathematical formulation rather than conforming to the restrictive standard form required by solvers [*Diamond & Boyd*, 2016, *Agrawal & Boyd*, 2020]. Table **1-1** summarizes the mentioned solvers.

Table 1-1: State-of-the-art solvers for optimization. (*) Except mixed-integer SDP. (**) Features available with the licensed version only

Solver	LP	QP	SOCP	SDP	EXP	PCP	MIP	NLP	Strategy	Open source	Software
Clarabel [*Goulart & Chen*, 2024]	✓	✓	✓	✓	✓	x	x	x	IP	✓	CVXPY
Gurobi [Gur, 2024]	✓	✓	✓	x	x	x	✓	x	IP, Simplex, BC	x	MATPOWER, CVXPY
Mosek [Mos, 2024]	✓	✓	✓	✓	✓	✓	✓*	x	IP	x	MATPOWER, CVXPY
Xpress [FIC, 2024]	✓	✓	✓	x	x	x	✓	✓**	IP, Simplex, BC	x	CVXPY
SCS [*O'Donoghue*, 2021, *O'Donoghue et al.*, 2016]	✓	✓	✓	✓	✓	✓	x	x	IP	✓	CVXPY
IPOPT [Ipo, 2024]	✓	✓	✓	✓	✓	✓	✓	✓	IP	✓	MATPOWER, GEKKO

1.3.4 Summary

In the discussion regarding classic algorithms, deep learning, and optimization solvers, the significance of IP-based techniques like IPOPT, utilized in the energy sector, is emphasized. This algorithm is renowned for its robustness in solving a broad spectrum of optimization problems [*Böttcher et al.*, 2023, *Diehl*, 2019]. Nonetheless, it encounters substantial challenges in the context of large networks, such as slow convergence and high computational costs, which restrict its applicability in extensive systems [*Owerko et al.*, 2020]. Furthermore, IPOPT demonstrates a superlinear time complexity in executing operations at each optimization step, rendering it inefficient for large-scale systems [*Falconer & Mones*, 2023], and sometimes fails to converge, as evidenced in tests where a reference solution was not achieved for all examples [*Owerko et al.*, 2022].

These challenges underscore the significance of exploring alternative approaches to optimization, such as algorithms based on AI. In this context, the use of PINN, which embed the constraints of optimization problems directly into the training phase, effectively addresses the complexities and nonlinearities inherent in these systems. This is accomplished through the application of automatic differentiation in the context of batch backpropagation, thereby facilitating parallelization and adaptation to systems with a larger number of variables. However, this approach also poses challenges, including the need for extensive data preparation and the development of custom loss functions for each specific situation, as well as difficulties in generalization and in determining the optimal number of parameters for integration into various electrical systems [*Pan et al.*, 2021, *Zhou et al.*, 2023, *Liang & Zhao*, 2023]. Consequently, significant improvements in these areas are essential. For instance, the solution proposed by [*Pan et al.*, 2023] requires specific training data and must ensure the feasibility of its solutions. Meanwhile, the method of [*Falconer & Mones*, 2023] is limited by its high computational cost compared to a basic DNN and by the limited scalability of exact solutions.

2 Aims

2.1 Aims

2.1.1 General Aim

Develop a nonlinear optimization approach based on physics-informed neural networks that addresses the unique challenges of nonlinear constraints in Colombia's gas energy system. This approach preserves the relationship between flows and pressures, allowing us to manage convergence instability and computational time issues.

2.1.2 Specific Aims

- Develop an optimization strategy based on neural networks that adapts to different levels of noise and specific constraints, ensuring results comparable to conventional optimization techniques in terms of compliance and robustness.
- Develop a customized nonlinear optimization scheme based on neural networks for gas energy systems that considers physical limits and non-differentiable functions, focusing on modeling gas supply behavior and compliance with system constraints.
- Develop a custom neural network-based nonlinear optimization strategy for Colombia's gas energy system, modeling the behavior of supply and demand flows, and pressures, while incorporating physical limits and non-differentiable functions to ensure efficiency over time and convergence.

3 Outline and Contributions

In this section, we provide a concise overview of the main contributions presented in this thesis, focusing on a regularized PINN framework as an optimization tool for Nonlinear Optimization (NOPT) scenarios driven by both supervised and unsupervised data, see Figure **3-1**.

3.1 Supervised Regularized Physics-Informed Neural Network to Support Nonlinear Constrained Optimization

To develop a new regularized PINN framework, termed Regularized Physics-Informed Neural Networ RPINN, for solving NOPT in supervised scenarios, it is crucial to address data variability and noisy input measurements. To this end, two specific problems are investigated.The first case has a practical nature and focuses on chaotic ecological patterns used to formulate an NOPT problem. This problem is tackled using a deep neural network (DNN) specifically designed for this purpose, serving as the foundation for the development of RPINN. The second case is based on a uniform mixture distribution, leading to a constrained NOPT problem. In this scenario, a limited regularization penalty based on the l1 function is applied to ensure compliance with the constraints, allowing us to validate our approach in solving optimization problems. The effectiveness of this methodology was verified using synthetic datasets affected by noise, enabling the evaluation of RPINN's capability to overcome these challenges.

3.2 Unsupervised Regularized Physics-Informed Neural Network to Support Nonlinear Constrained Optimization

To apply RPINN in unsupervised scenarios, we address three key issues of NOPT. First, we integrate nonlinear constraints into the network architecture, adhering to the physical

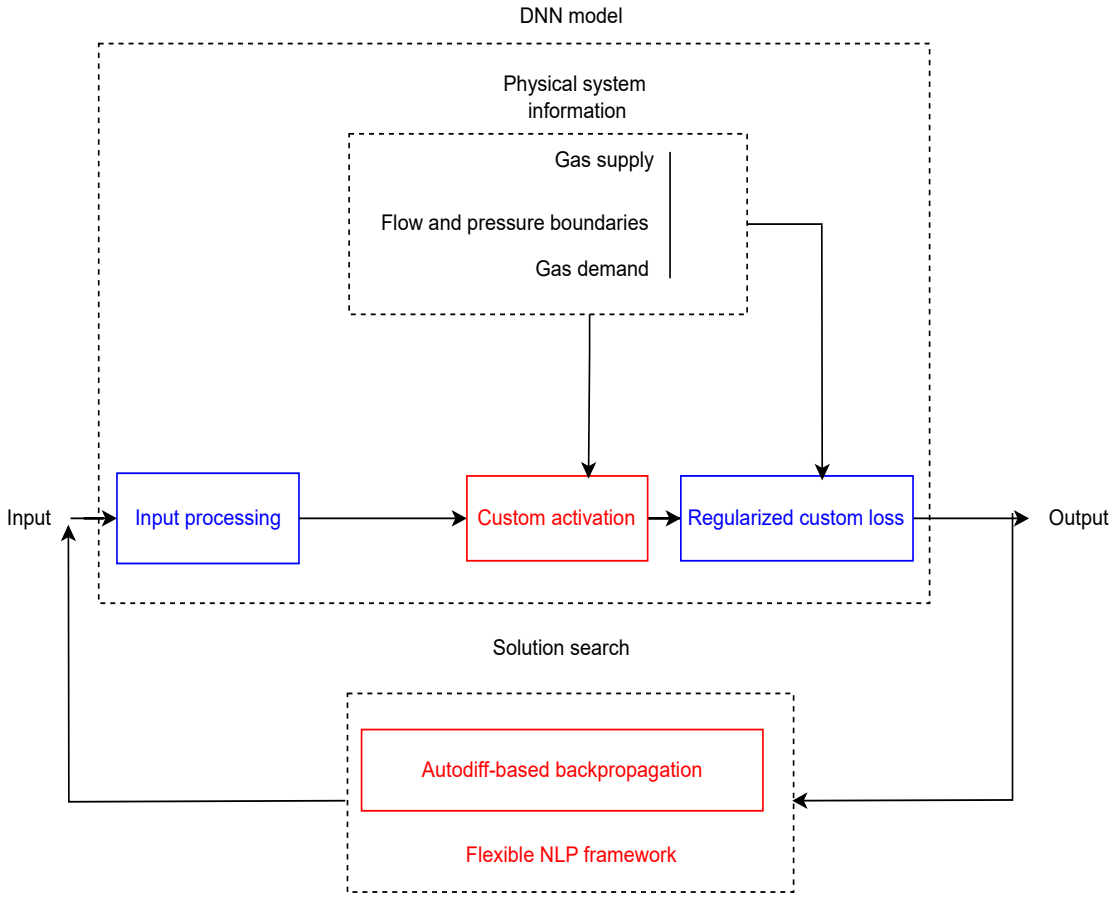

Figure 3-1: Schematic diagram illustrating the proposed solutions for convergence and time complexity addressed in this thesis to solve problems of non-linear optimization in gas-powered energy systems. The blue boxes represent the solutions to the convergence problem, while the red boards represent the solutions to the time problem.

principles of the model through customized activations based on the sigmoid curve and piecewise functions. This ensures boundary constraints, models the behavior of injected flows, and addresses unmet demand patterns in gas distribution networks. We also employ Huber-based regularization penalties within an ANN scheme to integrate fundamental constraints, such as gas flow balance, Weymouth equality, and compression factor limits, directly into the training process.

Second, we use the network weights and/or the learned features within a functional composition framework to determine the NOPT variables. Third, our ANN-based strategy employs AD training, which enhances system scalability and reduces computation time through batch-based backpropagation. The main advantage of our proposal lies in its stability against noisy input measurements, making it particularly valuable in contexts with fluctuating information. Additionally, because RPINN is based on ANN, it offers flexibility in terms of

network architecture.

The effectiveness of our approach has been validated through its incorporation into an RPINN, specifically designed to address NOPT problems. This method has been applied to both synthetic gas networks and the specific context of Colombia, allowing for the verification of constraint compliance, cost evaluation, reduction of processing time, and the convergence rate of solutions.

3.3 Thesis structure

The structure of this thesis is organized as follows: Chapter 4 details the development of a robust optimization approach named RPINN, specifically designed to tackle the challenges posed by NOPT problems. This approach lays the foundation for addressing more complex NOPT issues. Chapter 5 introduces an optimization strategy based on RPINN, designed to solve NOPT problems within gas network contexts. This chapter focuses on the unique constraints of these systems, aiming to ensure rapid and consistent efficiency and convergence. Finally, Chapter 6 summarizes the main findings of the research, explores future directions for further investigations, and provides a compilation of the academic achievements resulting from this work.

4 Supervised Regularized Physics - Informed Neural Network to Support Nonlinear Constrained Optimization

This section aims to establish the foundation for developing an optimization strategy based on RPINN, providing the essential theoretical framework for its implementation. The strategy begins by addressing a constrained NOPT problem derived from a synthetic uniform mixture problem, and then progresses to a sequential NOPT problem focused on chaotic ecological dynamics. The results are compared with traditional methods implemented using CVXPY. These problems are fundamental for tackling more complex optimization challenges. Finally, this section presents the results of experimental tests designed to highlight the strengths and weaknesses of our proposal, thereby offering crucial insights for extending the strategy to problems with more complex dynamics.

4.1 Theoretical Background

Addressing optimization challenges using RPINN typically requires the development of architectures tailored to the unique needs of the problem at hand. Therefore, this section includes the necessary theoretical background to understand the optimization problems presented in this segment, along with the theoretical foundation of our approach to meet the specific demands of each issue.

4.1.1 Automatic Differentiation

Investigating different strategies for differentiating functions uncovers four key methods. The initial approach, manual differentiation, employs a detailed and hands-on method utilizing fundamental principles to derive functions [*Baydin et al.*, 2018]. Symbolic differentiation, the second technique, mimics manual procedures on computers through the systematic

application of these principles[*Laue*, 2022]. Numerical differentiation, identified as the third method, is characterized by employing approximation methods for derivative calculation [*Meng et al.*, 2020b], as demonstrated in Eq. 4-2.

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ f'(x) &\approx \frac{f(x+h) - f(x)}{h} \end{aligned} \tag{4-1}$$

In this equation, f symbolizes the mathematical function to be derived, and h represents a minutely small increment nearing zero, utilized in the computation of the derivative. The notation $f'(x)$ signifies the derivative of the function f in relation to the independent variable.

Lastly, Automatic Differentiation AD, the fourth method, is highlighted as a groundbreaking development in computing derivatives accurately and efficiently through a sequence of arithmetic operations and basic functions [*Bolte & Pauwels*, 2021]. This approach combines the precision of symbolic differentiation with the practicality of numerical methods, rendering it an effective tool for diverse applications. Each method, with its distinct features and uses, is concisely depicted in Fig. **4-1**.

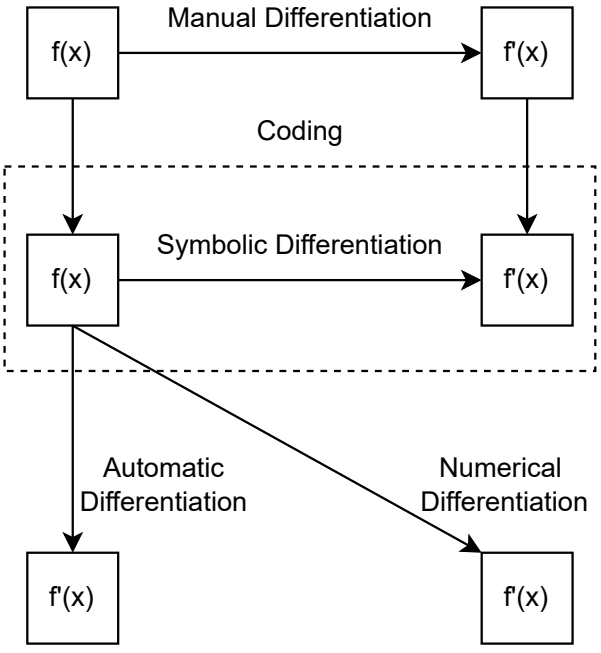

Figure 4-1: Function Differentiation: This figure presents and analyzes in detail the four fundamental techniques used to perform function differentiation.

In AD, the decomposition of differentials into composite functions through the chain rule is pivotal. Here's an outline of this method [*Paszke et al.*, 2017].

Given a function $y = f(g(h(x)))$, the AD-based decomposition can be viewed as:

$$\begin{aligned} y &= f(g(h(w_0))) \rightarrow w_0 = x \\ y &= f(g(w_1)) \rightarrow w_1 = h(\cdot) \\ y &= f(w_2) \rightarrow w_2 = g(\cdot) \\ y &= w_3 \rightarrow w_3 = f(\cdot) \end{aligned} \tag{4-2}$$

f, g, and h are individual mathematical functions that, when combined, create a composite function transforming x into y by successively applying the n function in the sequence, denoted by w_n.

The chain rule for the previous case would be:

$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial w_2}\frac{\partial w_2}{\partial w_1}\frac{\partial w_1}{\partial x} = \frac{\partial f(w_2)}{\partial w_2}\frac{\partial g(w_1)}{\partial w_1}\frac{\partial h(w_0)}{\partial x} \tag{4-3}$$

Eq. 4-3 illustrates how the rate of change of y relative to x can be expressed as the product of the derivatives of each intermediate variable w_n in its respective composition.

Practical Foundations: Dual Numbers

Dual numbers are a mathematical concept that extends real numbers in a similar way to how complex numbers extend real numbers. A dual number can be written in the form $x + \dot{x}\epsilon$, where x and $\dot{x}$ are real numbers, and ϵ is an additional element, that fulfils $\epsilon^2 \rightarrow 0$ and $\epsilon \neq 0$ [*Mora-Camino & Nunes Cosenza*, 2018].

Dual numbers have the following properties:

$$
\begin{aligned}
&1.\ (x, \dot{x}) = (x + \dot{x}\epsilon) \\
&2.\ (x, \dot{x}) + (y, \dot{y}) = (x + y, \dot{x} + \dot{y}) \\
&3.\ (x + \dot{x}) - (y + \dot{y}) = (x - y, \dot{x} - \dot{y}) \\
&4.\ k\,(x, \dot{x}) = (kx, k\dot{x}) \\
&5.\ (x, \dot{x})\,(y, \dot{y}) = (xy, \dot{x}y - x\dot{y}) \\
&6.\ \frac{(x, \dot{x})}{(y, \dot{y})} = \left(\frac{x}{y}, \frac{\dot{x}y - x\dot{y}}{y^2}\right) \\
&7.\ (x, \dot{x})^n = \left(x^n, nx^{n-1}\dot{x}\right) \\
&8.\ \sin(x, \dot{x}) = (\sin(x), \cos(x)\dot{x}) \\
&9.\ \cos(x, \dot{x}) = (\cos(x), -\sin(x)\dot{x}) \\
&10.\ e^{(x,\dot{x})} = (e^x, e^x\dot{x}) \\
&11.\ \log(x, \dot{x}) = \left(\log(x), \frac{\dot{x}}{x}\right); x \neq 0 \\
&12.\ \sqrt{(x, \dot{x})} = \left(\sqrt{x}, \frac{\dot{x}}{2\sqrt{x}}\right); x \neq 0
\end{aligned}
\tag{4-4}
$$

where, x and y are the real parts, $\dot{x}$ y $\dot{y}$, respectively, are the dual parts of x and y, k is a scalar, and n is an integer exponent used in the elevation of a power.

Application of Dual Numbers in Automatic Differentiation

When assessing functions through dual numbers, one can ascertain not only the value of the function but also its derivative at a particular point. Consider the multivariate function:

$$f(x_1, x_2) = x_1 x_2 + \sin(x_1) \tag{4-5}$$

This serves as an example to grasp the application of dual numbers in automatic differentiation (AD).

Original program	Dual program
$w_1 = x_1$	$\dot{w}_1 = \dot{x_1}$
$w_2 = x_2$	$\dot{w}_2 = \dot{x_2}$
$w_3 = w_1 \cdot w_2$	$\dot{w}_3 = w_1\dot{w}_2 + \dot{w}_1 w_2$
$w_4 = \sin(w_1)$	$\dot{w}_4 = \cos(w_1)\dot{w}_1$
$w_5 = w_3 + w_4$	$\tilde{\dot{w}}_5 = \dot{w}_3 + \dot{w}_4$

In this case, x_n symbolizes the independent variables and w_n the intermediate operations, while $\dot{x_n}$ and $\dot{w}_n$ represent their corresponding double counterparts.

Using the chain rule on the decomposition of the original problem we have:

$$\begin{aligned} \frac{\partial f(x_1, x_2))}{\partial x_2} &= \frac{\partial f}{\partial w_5}\frac{\partial w_5}{\partial w_3}\frac{\partial w_3}{\partial w_2}\frac{\partial w_2}{\partial x_2} \\ \frac{\partial f(x_1, x_2))}{\partial x_2} &= (1)(1)(w_1)(1) = w_1 = x_1 \end{aligned} \tag{4-6}$$

To apply dual numbers, we first need to know that for $\frac{\partial f}{\partial x_2}$ the components $\dot{w}_1 = 0$,$\dot{w}_2 = 1$, therefore:

$$\begin{aligned} \dot{w}_1 &= 0 \\ \dot{w}_2 &= 1 \\ \dot{w}_3 &= w_1\dot{w}_2 + \dot{w}_1 w_2 = w_1 + 0w_2 = w_1 \\ \dot{w}_4 &= \cos(w_1)\dot{w}_1 = \dot{w}_4 = \cos(0)0 = 0 \\ \dot{w}_5 &= \dot{w}_3 + \dot{w}_4 = w_1 + 0 = w_1 = x_1 \end{aligned} \tag{4-7}$$

Thus, the solution $w_5 = x_1$ matches Eq 4-6

Computational Graph

A computational graph can be used to represent the mathematical operations performed when decomposing functions, see sub-section 4.1.1. In this graph, nodes represent operations or values (such as inputs, outputs, and weights), while arches represent the relationship or data flow between these operations. For AD to perform its task efficiently, computational graphs divide the problem into smaller subproblems [*Haghighat et al.*, 2021].

To illustrate the use of graphs in AD,we use the function in Eq. 4-5 to construct the forward and backward graphs, see Fig. **4-2** and **4-3**.

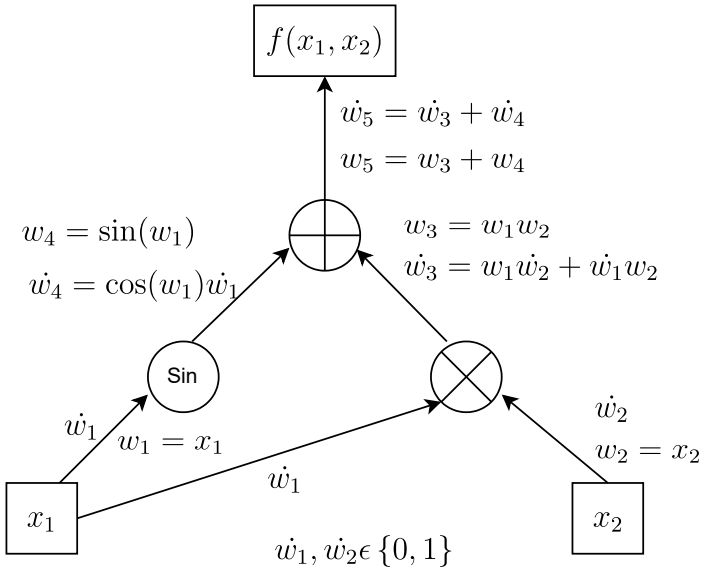

Figure 4-2: Forward propagation: The graph is illustrated with the architecture of the example seen in Eq. 4-5, with the arithmetic operations specific to the studied multivariate function

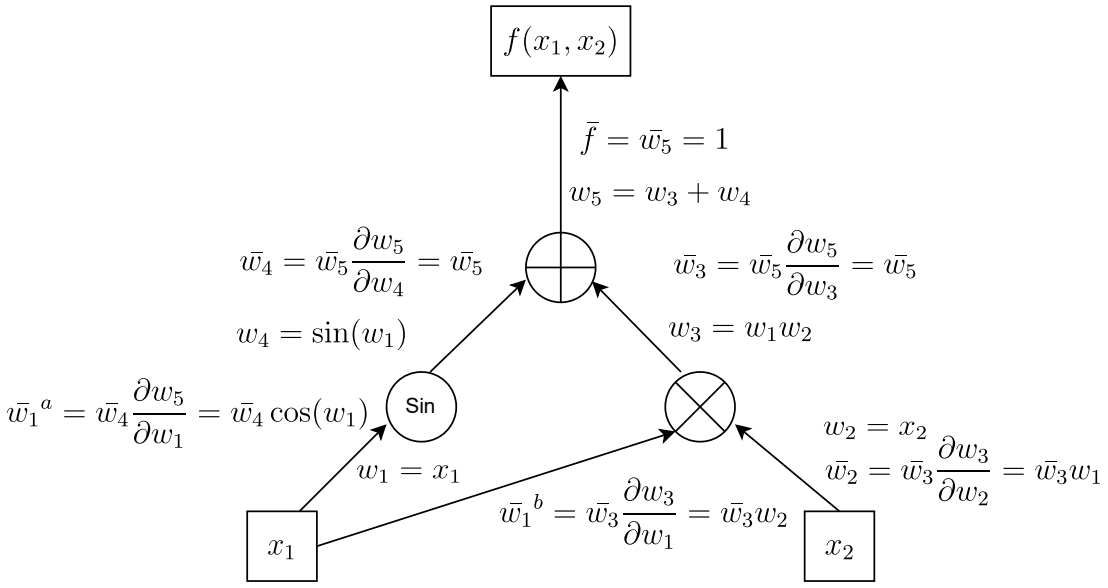

Figure 4-3: Backward propagation: The graph is illustrated with the architecture of the example seen in Eq. 4-5, with the arithmetic operations specific to the studied multivariate function

AD-based Jacobian

AD is pivotal in derivative computations, facilitating the targeted assembly of crucial Jacobian segments, circumventing the necessity for its complete computation or storage [*Baydin et al.*,

2018]. It's instrumental for functions $f : \mathbb{R}^n \to \mathbb{R}^m$, where the Jacobian is denoted by $J = F'(x) \in \mathbb{R}^{m \times n}$.

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_N} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \frac{\partial f_M}{\partial x_2} & \cdots & \frac{\partial f_M}{\partial x_N} \end{bmatrix} \tag{4-8}$$

Regarding the forward mode, the operation of interest is $J \cdot \vec{x}$, while in the reverse mode, focus shifts to the operation $\vec{y}^T \cdot J$. Here, $\vec{x}$ and $\vec{y}$ denote seed vectors. In forward mode, AD produces a Jacobian column, mirroring the gradient of f along $\vec{x}$. Conversely, in reverse mode, AD assembles a Jacobian row, encapsulating the derivatives of an output scalar function with respect to all input variables. This assembly is executed via the seed vector $\vec{y}$. Such procedures are crucial as they enable the efficient computation of gradients in the needed direction, $\vec{x}$ or $\vec{y}$, thereby streamlining the identification of pertinent derivatives for a specific problem.

4.1.2 Backpropagation Algorithm for Neural Networks Training

Retropropagation is crucial in neural networks, as it allows the model to adjust its weights by learning from its errors. During the direct passage, the network processes the input through sequential operations through its layers to produce an output [*Rumelhart et al.*, 1986]. Fig. **4-4** shows an example of a NN.

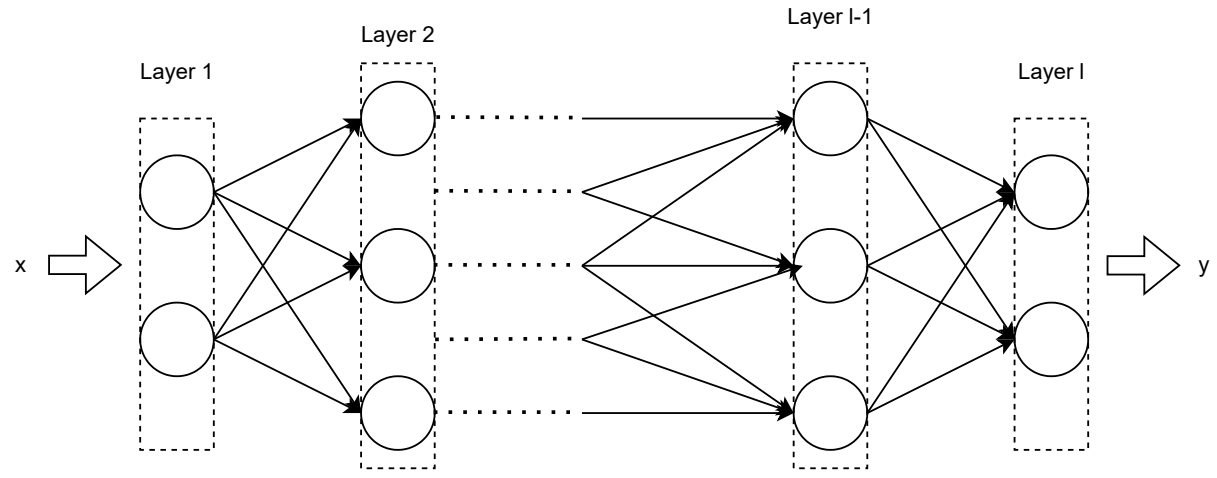

Figure 4-4: Schematic representation of a neural network: The diagram displays the general components of a network, consisting of an input, dense/fully connected hidden layers, and the output as estimated by the model.

To mathematically dissect the network's functionality, let's first assume that all units in a given layer exhibit akin traits. This assumption streamlines the network depicted in Fig. **4-4** into the configuration shown in Fig **4-5**.

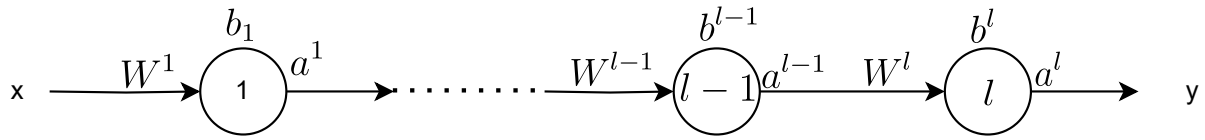

Figure 4-5: Simplified representation of the network structure, assuming uniform features within each layer and the components.

Within this framework, x and y symbolize the model's inputs and outputs, correspondingly. The weights interlinking layers $l-1$ and l are denoted by W^l, whereas b^l signifies the biases of layer l, with the superscript indicating the layer's sequential number.

The model produces an output, represented as $\hat{y} = a^l(z)$. In this expression, $z^l = W^l a^{l-1} + b^L$ denotes the entry for the layer l, affected by the bias (indicated by b^l). Subsequently, the a^l activation function is applied to this to get the output of the current layer.

During the model parameters optimization, it is crucial to examine the sensitivity of the cost function, represented by $C(y, \hat{y})$, to update W^l and b^l. This involves the calculation of the gradients $\frac{\partial C}{\partial W^l}$ and $\frac{\partial C}{\partial b^l}$. The analytical process consists of two essential stages: i) derivation at layer l; ii) derivation at layer $l-1$, both in relation to the parameters.

applying the chain rule, the derivatives with respect to the parameters of the l layer in the network would be:

$$\begin{aligned} \frac{\partial C}{\partial W^l} &= \frac{\partial C}{\partial a^l}\frac{\partial a^l}{\partial z^l}\frac{\partial z^l}{\partial W^l} \\ \frac{\partial C}{\partial b^l} &= \frac{\partial C}{\partial a^l}\frac{\partial a^l}{\partial z^l}\frac{\partial z^l}{\partial b^l} \end{aligned} \tag{4-9}$$

where:

$$\begin{aligned} &\frac{\partial C}{\partial a^l}\frac{\partial a^l}{\partial z^l} = \frac{\partial C}{\partial z^l} = \delta^l \\ &\frac{\partial z^l}{\partial W^l} = \frac{\partial (W^l a^{l-1} + b^l)}{\partial W^l} = a^{l-1} \\ &\frac{\partial z^l}{\partial b^l} = \frac{\partial (W^l a^{l-1} + b^l)}{\partial b^l} = 1 \end{aligned} \tag{4-10}$$

Substituting into Eq 4-9, we obtain:

$$\frac{\partial C}{\partial W^l} = \delta^l a^{l-1}$$
$$\frac{\partial C}{\partial b^l} = \delta^l \tag{4-11}$$

Again, based on the chain rule, the derivatives with respect to the parameters of the $l-1$ layer in the network would be:

$$\frac{\partial C}{\partial W^{l-1}} = \frac{\partial C}{\partial a^l}\frac{\partial a^l}{\partial z^l}\frac{\partial z^l}{\partial a^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}}\frac{\partial z^{l-1}}{\partial W^{l-1}}$$

$$\frac{\partial C}{\partial b^{l-1}} = \frac{\partial C}{\partial a^l}\frac{\partial a^l}{\partial z^l}\frac{\partial z^l}{\partial a^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}}\frac{\partial z^{l-1}}{\partial b^{l-1}}$$

Plugging the determined values into Eq 4-10, we get:

$$\frac{\partial C}{\partial W^{l-1}} = \delta^l \frac{\partial z^l}{\partial a^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}}\frac{\partial z^{l-1}}{\partial W^{l-1}}$$
$$\frac{\partial C}{\partial b^{l-1}} = \delta^l \frac{\partial z^l}{\partial a^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}}\frac{\partial z^{l-1}}{\partial b^{l-1}} \tag{4-12}$$

where:

$$\frac{\partial z^l}{\partial a^{l-1}} = \frac{\partial (W^l a^{l-1} + b^l)}{\partial a^{l-1}} = W^l$$

$$\frac{\partial z^{l-1}}{\partial W^{l-1}} = \frac{\partial (W^{l-1} a^{l-2} + b^{l-1})}{\partial W^{l-1}} = a^{l-2}$$

$$\frac{\partial z^{l-1}}{\partial b^{l-1}} = \frac{\partial (W^{l-1} a^{l-2} + b^{l-1})}{\partial b^{l-1}} = 1$$

$$\delta^l \frac{\partial z^l}{\partial a^{l-1}}\frac{\partial a^{l-1}}{\partial z^{l-1}} = W^l \delta^l \frac{\partial a^{l-1}}{\partial z^{l-1}} = \delta^{l-1}$$

Inputting into Eq 4-12 yields:

$$\frac{\partial C}{\partial W^{l-1}} = \delta^{l-1} a^{l-2}$$
$$\frac{\partial C}{\partial b^{l-1}} = \delta^{l-1} \tag{4-13}$$

In summary, backpropagation aims to compute the gradient of the loss function with respect to each network parameter, a crucial step for weight optimization [*Rumelhart et al.*, 1986].

AD streamlines this process by utilizing the computational graph to chart a reverse path from output to input. This reverse pathway enables the error to propagate throughout the network [*Baydin et al.*, 2018], thereby allowing parameter adjustments to minimize future prediction errors.

4.1.3 Nonlinear Optimization Fundamentals (NOPT)

Let $\mathbf{x} \in \mathbb{R}^P$ be a vector in P variables. The conventional NOPT problem can be summarized as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \varrho(\mathbf{x}) \\ \text{s.t.} \quad & \boldsymbol{\xi}_{\min} \leq \mathbf{x} \leq \boldsymbol{\xi}_{\max} \\ & h_L(\mathbf{x}) \leq \mathbf{0} \\ & h_N(\mathbf{x}) \leq \mathbf{0}, \end{aligned} \tag{4-14}$$

where the objective function $\varrho : \mathbb{R}^P \to \mathbb{R}$ is real-valued. Also, the bound constraints are shown by $\boldsymbol{\xi}_{\min}, \boldsymbol{\xi}_{\max} \in \mathbb{R}^P$. The linear and nonlinear constraints are described by $h_L : \mathbb{R}^P \to \mathbb{R}^{C_L}$ and $h_N : \mathbb{R}^P \to \mathbb{R}^{C_N}$, where $C_L \in \mathbb{N}$ and $C_N \in \mathbb{N}$.

Figure **4-6** depicts the main pipeline of the classical approaches for NOPT. First, it includes the physical system's parameters, constraints, limits, and the objective function to be optimized. Second, starting from an initial point, the optimization algorithm iterates until convergence. Of note, the number of iterations, the level of improvement, and the objective function thresholding are the relevant stopping criteria to return the final output.

4.1.4 Convex Optimization in Python CVXPY

CVXPY is a software library designed to solve convex programming problems [*Moehle et al.*, 2017]. It allows to efficiently build mathematical optimization models

where, thanks to the flexibility of the tool, various problems can be solved through the variety of solvers available, see Table **1-1**.

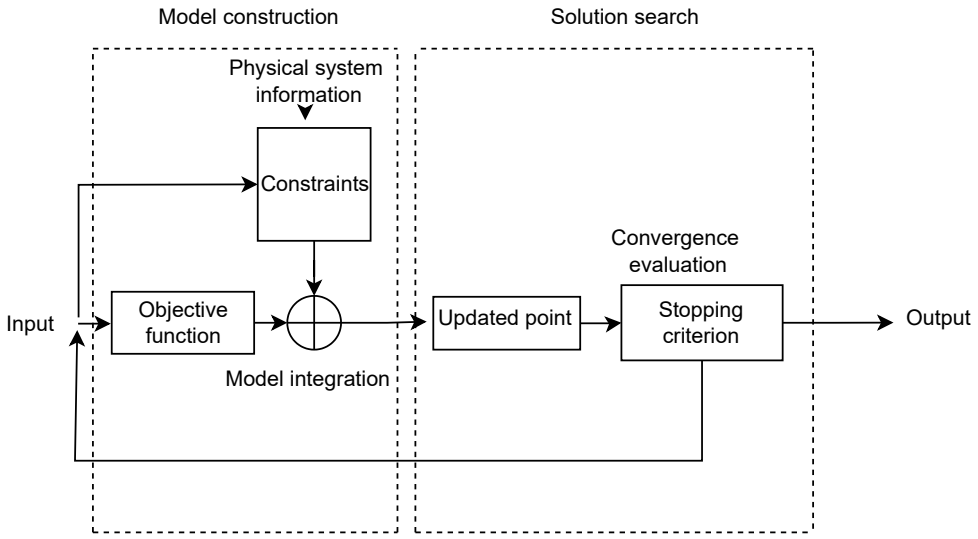

Figure 4-6: Classical optimization pipeline for NOPT

Convex Optimization Layers

The research [*Agrawal et al.*, 2019] introduces a significant application that restructures the traditional optimization problem into a novel framework:

$$\begin{aligned} &\min_{x} \ \varrho(\mathbf{x};\mathbf{s}) \\ &\text{s.t } h_N(\mathbf{x};\mathbf{s}) \leq 0 \\ &h_L(\mathbf{x};\mathbf{s}) = 0 \end{aligned} \tag{4-15}$$

As seen, it defines the optimization problem w.r.t. the parameters $\mathbf{s}$, an objective function ϱ, and the restriction functions h_N and h_L, facilitating comprehensive end-to-end learning with respect to $\mathbf{s}$.

As depicted in Figure **4-7**, this approach enables the integration of an optimization layer designed to be compatible with computational graphs. This allows for the application of AD in batch backpropagation contexts, ensuring compatibility with both TensorFlow and PyTorch. Furthermore, this method facilitates the use of various solvers included in its toolkit to address strictly convex problems.

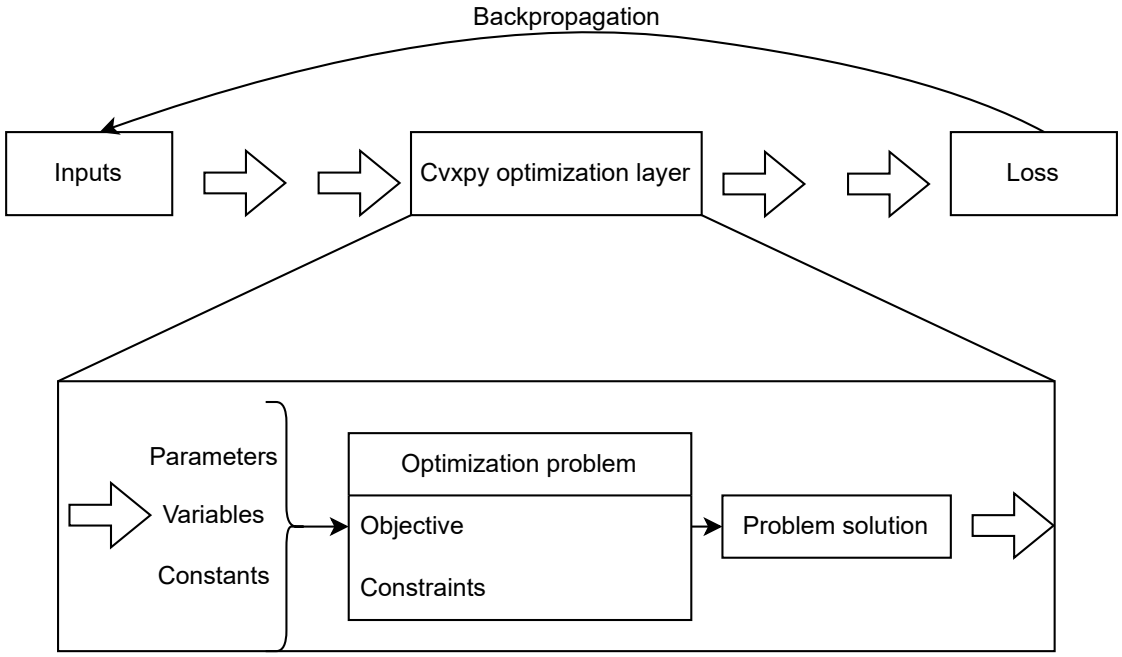

Figure 4-7: CVXPY main scheme. NN-based models are employed within an backpropagation scheme introducing the CVXPY layer.

4.1.5 Regularized Physics-Informed Neural Network (RPINN)

Let $\{y_r \in \mathcal{Y}, z_r \in \mathcal{Z}\}_{r=1}^{R}$ be an input-output set holding R samples. Our data-driven RPINN approach aims to couple the optimization problem in Eq. 4-14 as a penalty-based loss with bounded constraints from both network weights and learned features, as follows:

$$\begin{aligned}
&\min_{\tilde{\mathbf{X}},\tilde{\mathbf{Z}}} \frac{\lambda_{\mathcal{L}}}{R}\sum_{r=1}^{R}\mathcal{L}\left(y_r,\tilde{f}(z_r|\tilde{\mathbf{X}},\tilde{\mathbf{Z}})\right)+\sum_{i=1}^{C_L}\frac{\lambda_{Li}}{R}\sum_{r=1}^{R}\tilde{h}_{Li}(y_r,\tilde{f}(z_r|\tilde{\mathbf{X}},\tilde{\mathbf{Z}}))+\sum_{j=1}^{C_N}\frac{\lambda_{Nj}}{R}\sum_{r=1}^{R}\tilde{h}_{Nj}(y_r,\tilde{f}(z_r|\tilde{\mathbf{X}},\tilde{\mathbf{Z}}))\\
&\text{s.t.}\quad \lambda_{\mathcal{L}}+\sum_{i=1}^{C_L}\lambda_{Li}+\sum_{j=1}^{C_N}\lambda_{Nj}=1\\
&\qquad \boldsymbol{\zeta}_{\min}\leq\tilde{\mathbf{X}}\leq\boldsymbol{\zeta}_{\max}\\
&\qquad \boldsymbol{\psi}_{\min}\leq\tilde{f}(z_r|\tilde{\mathbf{X}},\tilde{\mathbf{Z}})\leq\boldsymbol{\psi}_{\max},\quad \forall r\in R;
\end{aligned} \tag{4-16}$$

where $\tilde{f}:\mathcal{Z}\to\mathcal{Y}$ is an ANN-based mapping function, $\mathcal{L}:\mathcal{Y}\times\mathcal{Y}\to\mathbb{R}$ is a given loss, $\tilde{\mathbf{X}}$ holds the network parameters, and $\tilde{\mathbf{Z}}$ gathers the learned features along layers. Also, $\tilde{h}_{Li}(\cdot,\cdot)$ and $\tilde{h}_{Ni}(\cdot,\cdot)$ are the i-th linear and j-th nonlinear penalty functions to follow the NOPT constraints set by the regularization terms $\lambda_{\mathcal{L}},\lambda_{Li},\lambda_{Nj}\in[0,1]$, where $i\in\{1,2,\ldots,C_L\}$ and $j\in\{1,2,\ldots,C_N\}$. Furthermore, $\boldsymbol{\zeta}_{\min}$ and $\boldsymbol{\zeta}_{\max}$ collect the network parameter limit values, and $\boldsymbol{\psi}_{\min}$ and $\boldsymbol{\psi}_{\max}$ capture the network output and feature bounds.

For a given input $z \in \mathcal{Z}$, our deep learning-based function with $\hat{L}$ layers yields:

$$\begin{aligned} \tilde{f}(z|\tilde{\mathbf{X}}, \tilde{\mathbf{Z}}) &= (f_{\hat{L}} \circ \cdots \circ f_1|\tilde{\mathbf{X}}, \tilde{\mathbf{Z}})(z), \\ \tilde{\mathbf{z}}_l &= f_l(\tilde{z}_{l-1}|\tilde{\mathbf{x}}_l, b_l) = \nu_l(\tilde{\mathbf{x}}_l^\top \tilde{\mathbf{z}}_{l-1} + b_l). \end{aligned} \tag{4-17}$$

In the l-th layer of Eq. 4-17, where $l \in \{1, 2, \ldots, \hat{L}\}$, the weights and bias are $\tilde{\mathbf{x}}_l, b_l \in \tilde{\mathbf{X}}$, the learned feature vector is $\tilde{\mathbf{z}}_l \in \tilde{\mathbf{Z}}$, and $\nu_l(\cdot)$ is a nonlinear activation function to deal with both network representation and customized bounds to fulfill the Eq. 4-16 limit constraints. Furthermore, the RPINN optimization problem can be solved via gradient descent with AD and back-propagation [*Murphy*, 2022].

It is worth noting that our baseline RPINN studies a supervised scenario for simplicity, but by addressing its regularized loss, we can easily achieve an unsupervised extension. Figure **4-8** depicts the RPINN main sketch.

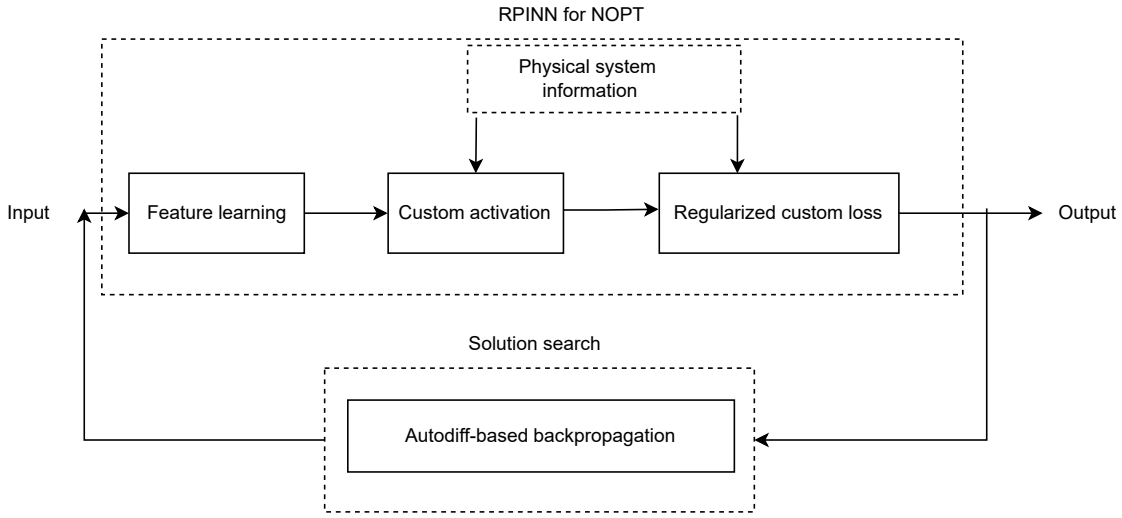

Figure 4-8: Regularized physics-informed neural network for data-driven nonlinear constrained optimization main sketch

4.2 Tested Scenarios for NOPT using RPINN

Adult Fly Population Model: In this study, we investigate the challenge of a NOPT problem in dynamic ecological environments. Specifically, our research focuses on the population of adult flies, as illustrated in Figure **4-9**.

In particular, the NOPT problem is written as:

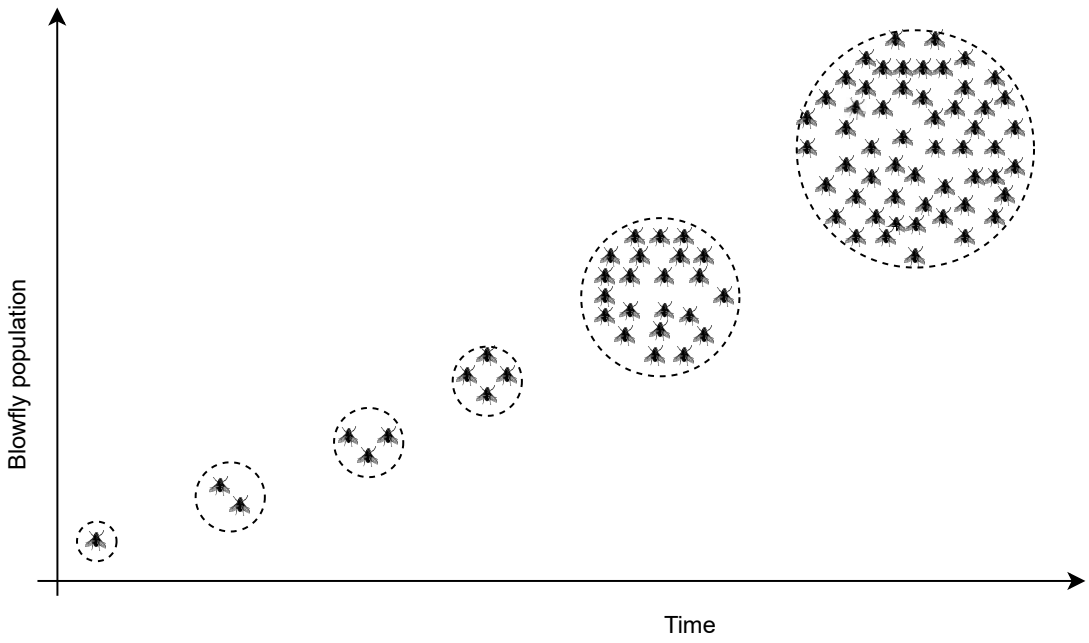

Figure 4-9: Adult Fly Population Model Diagram: The figure displays the random behavior in the fly population growth over time, graphically summarizing the dynamics of Equation 4-18.

$$
\begin{aligned}
&y_r = a y_{r-\tau-1} e^{\frac{-y_{r-\tau-1}}{y_0}} \tilde{a}_r + y_{r-1} e^{-\delta \tilde{\epsilon}_r} \\
&\text{In vector form:} \\
&\mathbf{x}_t = \begin{vmatrix} y_{r-1-\tau} \\ y_{r-1} \end{vmatrix}, \quad \mathbf{z}_t = \begin{vmatrix} a \exp\left(\frac{-y_{r-1-\tau}}{x_0}\right) \\ \exp(-\delta \tilde{\epsilon}_r) \end{vmatrix} \\
&y_t = \mathbf{x}_t^T \mathbf{z}_t
\end{aligned}
\tag{4-18}
$$

where $\tau \in \mathbb{N}$, $a, \delta \in \mathbb{R}$, $\tilde{a}_r \sim \mathcal{G}(\tilde{a}|1/\sigma_{\tilde{a}}^2, \sigma_{\tilde{a}}^2)$, and $\tilde{\epsilon}_r \sim \mathcal{G}(\tilde{\epsilon}|1/\sigma_{\tilde{\epsilon}}^2, \sigma_{\tilde{\epsilon}}^2)$. $\mathcal{G}(\cdot|\cdot,\cdot)$ stands for the Gamma distribution.

In the current context, we consider $y_t \in \mathbb{R}$, $\mathbf{z}_t$, $\mathbf{x}_t \in \mathbb{R}^P$, with P representing the number of variables involved. Additionally, the constants a, m_0 and ε are randomly taken from a normal distribution: $\mathcal{N}(\mu, \sigma^2)$:

$$
\min_{\mathbf{x}_t} \quad ||y_t - \mathbf{x}_t^T \mathbf{z}_t||_2^2 \tag{4-19}
$$

The optimization problem in Eq. 4-19 can be solved through our RPINN as follows:

$$\min_{\tilde{\mathbf{x}}} \frac{\lambda_{\mathcal{L}}}{R} \sum_{r=1}^{R} \mathcal{L}_H \left(y_{tr}, \tilde{f}(\mathbf{z}_{tr}|\tilde{\mathbf{X}}); \epsilon \right) \tag{4-20}$$

For concrete testing and to mitigate noisy samples, a Huber-based loss is used in Eq. 4-20:

$$\mathcal{L}_H \left(y_t, \tilde{f}(\mathbf{z_t}|\tilde{\mathbf{X}}); \epsilon \right) = \begin{cases} \frac{1}{2}\|y_t - \tilde{f}(\mathbf{z_t}|\tilde{\mathbf{X}})\|^2 & \|y_t - \tilde{f}(\mathbf{z_t}|\tilde{\mathbf{X}})\| \leq \epsilon \\ \epsilon \cdot (\|y_t - \tilde{f}(\mathbf{z_t}|\tilde{\mathbf{X}})\| - \frac{1}{2}\epsilon) & \|y_t - \tilde{f}(\mathbf{z_t}|\tilde{\mathbf{X}})\| > \epsilon, \end{cases} \tag{4-21}$$

where $\epsilon \in \mathbb{R}^+$.

Uniform Mixture Model: This task comprises a linear and bound-constrained optimization of a nonlinear cost [*González-Vanegas et al.*, 2019]:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{r=1}^{R} |y_r - \mathbf{x}^\top \mathbf{z}_r|_2^2 \\ \text{s.t.} \quad & \mathbf{0} \leq \mathbf{x} \leq \mathbf{1}, \\ & \mathbf{x}^\top \mathbf{1} = 1; \end{aligned} \tag{4-22}$$

where $y_r \in \mathbb{R}^+$ is the r-th target output, $\mathbf{x} \in \mathbb{R}^P$ denotes the mixing coefficients, and $\mathbf{z}_r \in \mathbb{R}^P$ holds random samples drawn from a uniform distribution as: $z_{rp} \sim \mathcal{U}(z|p-1, p)$,see Figure **4-10**. **0** and **1** are all-zeros and all-ones vectors of proper size.

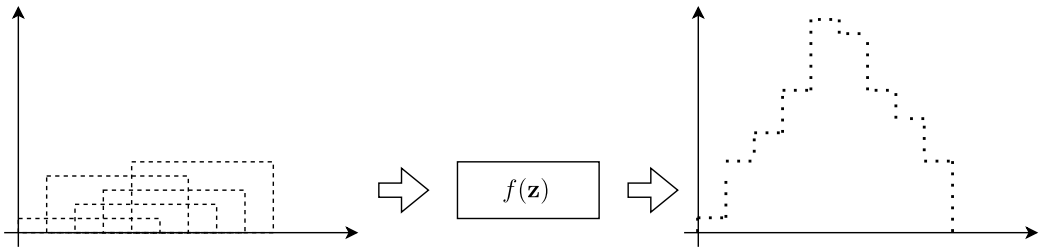

Figure 4-10: Uniform mixture model optimization. Left: weighted uniform probabilities. Right: visual representation of the mixing results

The optimization problem in Eq. 4-22 can be solved through our RPINN as follows:

$$\begin{aligned} \min_{\tilde{\mathbf{x}}} \quad & \frac{\lambda_{\mathcal{L}}}{R} \sum_{r=1}^{R} \mathcal{L}_H \left(y_r, \tilde{f}(\mathbf{z}_r|\tilde{\mathbf{X}}); \epsilon \right) + \frac{\lambda_L}{R} \tilde{\mathbf{x}}_{\hat{L}}^\top \mathbf{1} \\ \text{s.t.} \quad & \lambda_{\mathcal{L}} + \lambda_L = 1, \quad \forall \lambda_{\mathcal{L}}, \lambda_L \in [0, 1] \\ & \mathbf{0} \leq \tilde{\mathbf{x}}_{\hat{L}} \leq \mathbf{1}. \end{aligned} \tag{4-23}$$

For concrete testing and to mitigate noisy samples, a Huber-based loss is used in Eq. 4-23:

$$\mathcal{L}_H\left(y, \tilde{f}(\mathbf{z}|\tilde{\mathbf{X}}); \epsilon\right) = \begin{cases} \frac{1}{2}\|y - \tilde{f}(\mathbf{z}|\tilde{\mathbf{X}})\|^2 & \|y - \tilde{f}(\mathbf{z}|\tilde{\mathbf{X}})\| \leq \epsilon \\ \epsilon \cdot (\|y - \tilde{f}(\mathbf{z}|\tilde{\mathbf{X}})\| - \frac{1}{2}\epsilon) & \|y - \tilde{f}(\mathbf{z}|\tilde{\mathbf{X}})\| > \epsilon, \end{cases} \tag{4-24}$$

where $\epsilon \in \mathbb{R}^+$. Next, we fix a scaled exponential linear (SELU) activation for the network function composition, as follows:

$$\text{SELU}(x) = \begin{cases} \theta x & x > 0 \\ \theta\vartheta \cdot (e^x - 1) & x \leq 0, \end{cases} \tag{4-25}$$

where $\theta, \vartheta \in \mathbb{R}$. Then, $\tilde{\mathbf{x}}_{\hat{L}}$. To fulfill the former NOPT limit restriction, RPINN's weights at the output layer $\hat{L}$ hold a $l1$-based max-constraint.

4.3 Experimental Set-Up

This section will detail the experimental setups used to validate our proposal. We conducted two types of experiments to assess the effectiveness of our model. Initially, we compared outcomes in solving a constrained NOPT problem by employing bounded regularization. Subsequently, we assessed our model's performance on an unconstrained sequential NOPT problem.

4.3.1 Deep Learning Architectures

To address the uniform mixture model NOPT (supervised constrained optimization), our RPINN consists of two dense layers, as shown in Figure **4-11** and Table **4-1**.

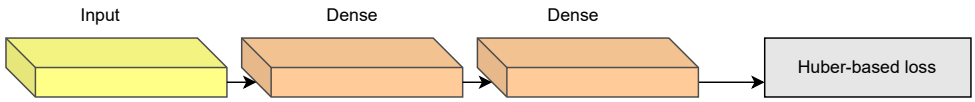

Figure 4-11: RPINN pipeline for the uniform mixture model-based NOPT

To solve the adult fly population model NOPT (see Eq. 4-19), it is essential to adjust the optimization variables in each cycle. For this purpose, a neural network was designed whose final layer multiplies our estimated variables, derived from the last dense layer, by the corresponding input (see Figure **4-12** and Table **4-2**).

Table 4-1: RPINN details for the uniform mixture model-based NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Param. #: number of trainable parameters. Total # of parameters: 30

Layer name	Type	Output shape	Param. #
Input	InputLayer	$(\tilde{R}, 5)$	0
Dense_1	Dense(SELU)	$(\tilde{R}, 5)$	25
Dense_2	Dense(SELU, $l1$-max-constraint)	$(\tilde{R}, 1)$	5

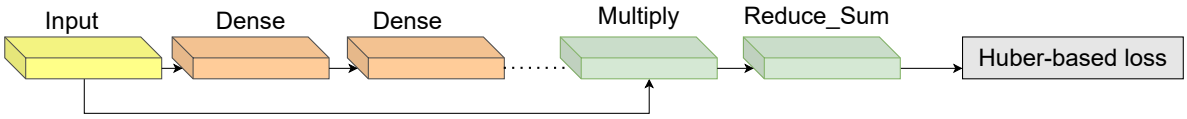

Figure 4-12: RPINN pipeline for the adult fly population model-based NOPT

Table 4-2: RPINN details for the adult fly population model-based NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Param. #: number of trainable parameters. Total # of parameters: 3330

Layer name	Type	Output shape	Param. #
input	InputLayer	$(\tilde{R}, 2)$	0
Dense_1	Dense(SELU,$l1l2$-constraint)	$(\tilde{R}, 32)$	96
Dense_2	Dense(SELU,$l1l2$-constraint)	$(\tilde{R}, 32)$	1056
Dense_3	Dense(SELU,$l1l2$-constraint)	$(\tilde{R}, 32)$	1056
Dense_4	Dense(SELU,$l1l2$-constraint)	$(\tilde{R}, 32)$	1056
W	Dense(SELU,$l1l2$-constraint)	$(\tilde{R}, 2)$	66
Multiply_1	TFOpLambda	$(\tilde{R}, 2)$	0
Reduce_Sum	TFOpLambda	$(\tilde{R},1)$	0

4.3.2 Training Details and Method Comparison

To evaluate the effectiveness of our methodology in addressing optimization problems, we utilized the mean absolute percentage error MAPE as the primary performance measure across all conducted experiments, defined as:

$$\text{MAPE}(\tilde{y}_r, \hat{y}_r) = \frac{100}{R} \sum_{r=1}^{R} \left| \frac{\tilde{y}_r - \hat{y}_r}{\tilde{y}_r} \right| [\%], \tag{4-26}$$

where $\tilde{y}_r, \hat{y}_r \in \mathbb{R}$ stands for r-th target and predicted value, $\text{MAPE}(\cdot, \cdot) \in [0, 100][\%]$, and $|\cdot|$ is the absolute value operator.

For the uniform mixture and adult fly population models, we generated 500 samples, each consisting of five and two variables respectively. We trained our RPINN architectures using a total of 400 samples, reserving 30% for the validation phase. The remaining 100 samples

were used to evaluate the model's performance. To analyze the efficacy of NOPT against noisy inputs, we added Gaussian white noise to the output of the uniform mixture model, maintaining the signal-to-noise ratio (SNR) at values of $\{-1, 3, 5\}$.

We implemented RPINN using Python 3.10.12 and the TensorFlow API 2.15.0 on Google Colaboratory. For training, we fixed 500 epochs, a batch size of 64 samples, an Adam optimizer, and a learning rate value of $1e-3$. Since IPOPT excels at solving NOPT, not to mention its open access, we fix it as a method comparison [Ipo, 2024]. Our codes and studied datasets are publicly available at `https://github.com/UN-GCPDS/python-gcpds.optimization` (accessed on 1 March 2024).

4.4 Results and Discussion

4.4.1 Adult Fly Population Model Results

As shown in Figure **4-13**, for chaotic data in the blowfly population model scenario, both our proposal and the solutions of conventional techniques exhibit similar results, which is due to the fact that the problem defined in Equation 4-19 is convex. Our RPINN, based on the Huber loss function, demonstrates greater robustness against data variability. The Huber function applies both the l1 and l2 norms to treat outliers and values close to our target (zero) differently, reducing sensitivity to extreme values and ensuring accuracy by penalizing smaller errors. In contrast, classical techniques solve the original problem based on the l2 norm, which is sensitive to outliers. Our proposal surpasses classical techniques in terms of the median MAPE, and for outliers, our solution is comparable to conventional methods.

4.4.2 Uniform Mixture Model Results

As shown in Figure **4-14** and Figure **4-15** (left), for noisy-free data on the uniform mixture model scenario, both our proposal and the SCS solution exhibit similar results. The similarity of the results stems from the fact that the problem defined in Eq. 4-22 is convex. Next, for noisy inputs, our RPINN, based on the Huber loss function, shows greater robustness against data variability and noise issues. In fact, the Huber function applies the $l1$-norm for errors exceeding a defined threshold, reducing sensitivity to extreme values, while for smaller errors it uses the $l2$-norm, ensuring accuracy by penalizing smaller errors. In contrast, the classical SCS technique uses an objective function based on the $l2$-norm, which is sensitive to outliers because it significantly penalizes large deviations. The weight distributions provide support for the latter hypothesis. Noise-free data leads to similar strength predictions for both RPINN and SCS. Conversely, for noisy inputs, our proposal regularizes the network

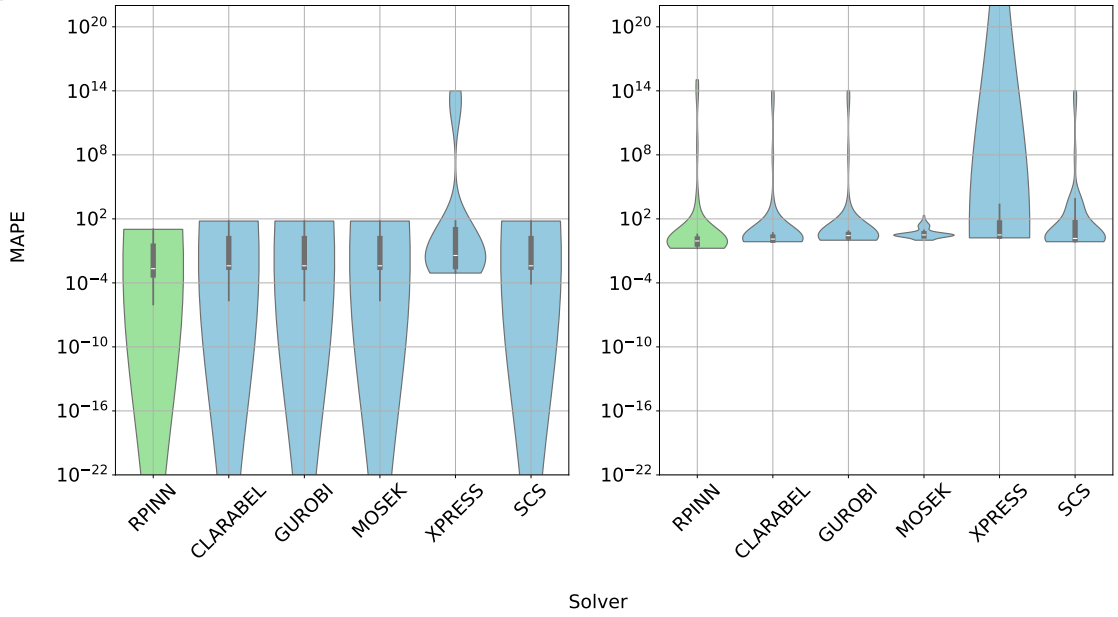

Figure 4-13: Blowfly population model MAPE results. Left: output error. Right: weights error

weights, yielding concentrated values to find the main output dynamics, and outperforms the SCS regarding the MAPE for all considered SNR values.

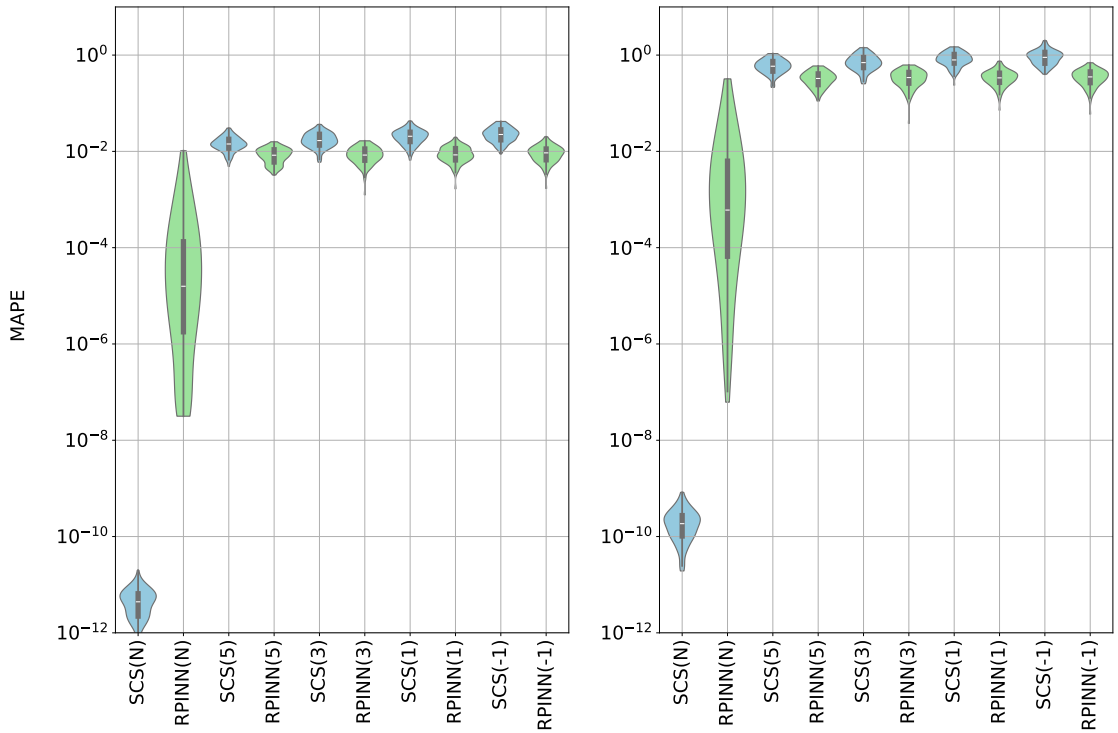

Figure 4-14: Uniform mixture model MAPE results. Left: output error. Right: weights error. (N): noisy-free. (-1),(3), and (5) stand for the SNR value

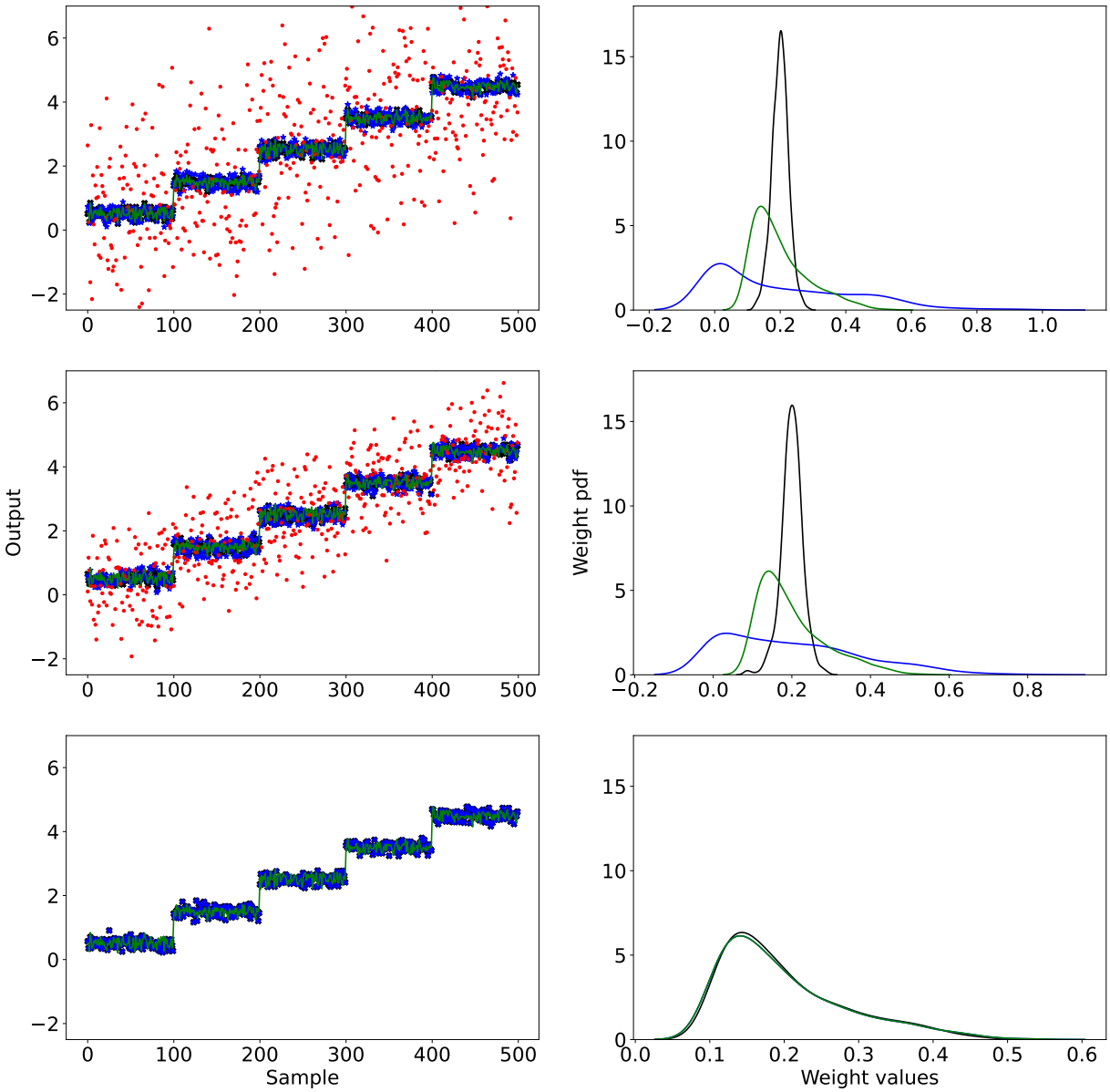

Figure 4-15: RPINN uniform mixture model-based NOPT results. First row: SNR= −1. Second row: SNR= 3. Third row: noise-free. Left: output prediction. Right: weight distribution. Green: target. Red: noisy target. Black: RPINN. Blue: SCS

4.4.3 Computational Cost Results

Figure **4-16** presents a comparison between the training and prediction times of RPINN and those of IPOPT. During the training phase, our model requires more time due to the execution of forward and backward steps in each iteration, an inherent characteristic of its ANN architecture. However, in the prediction phase, RPINN surpasses IPOPT, offering significantly shorter times for the adult fly population model. This is because, once the

weights are trained, our approach only needs to perform forward steps without additional weight updates, estimating all test set variables (100 samples) simultaneously, whereas IPOPT processes them individually. These features demonstrate RPINN's capability to generate fast and accurate predictions in NOPT solutions, reducing both processing times and interquartile ranges. However, in the case of uniform mixture models, both IPOPT and our proposal process the test data set as a single problem, resulting in a slight time advantage for the classical technique.

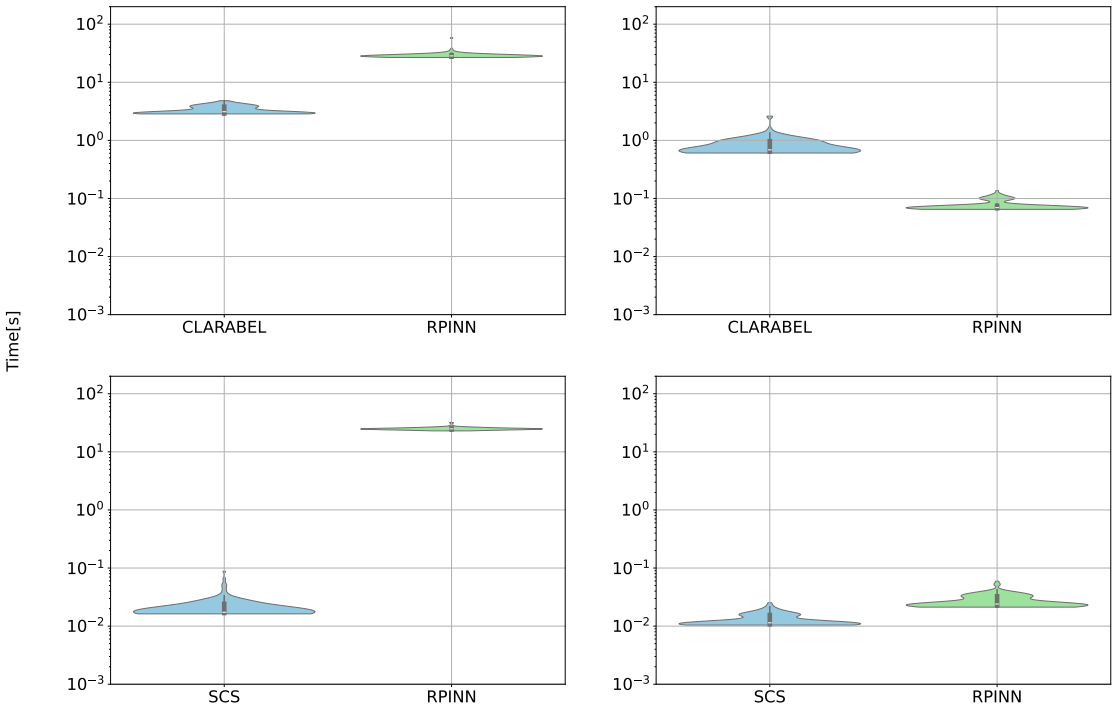

Figure 4-16: RPINN vs. Solver: The graph compares solution times for the test data between the classical technique (CLARABEL and SCS, in blue) and our strategy (RPINN, in green). On the left, training times are shown, while on the right, prediction times are displayed. The first and second rows correspond to the adult fly population model and the uniform mixture model, respectively.

4.5 Summary

This chapter introduced a neural network-based strategy, termed RPINN, to tackle NOPT challenges, examining two distinct contexts: a scenario composed of a uniform mixture distribution and an applied one focused on chaotic ecological dynamics. The models were evaluated under various degrees of interference using the CVXPY tool and validated through the implementation of test data sets. The effectiveness of these models was quantified using the MAPE metric, highlighting the consistent accuracy of neural networks despite variations in noise levels, demonstrating their robustness and superiority over conventional methods.

This efficiency is attributed to the properties of the Huber-based loss function, which shows greater robustness against data variability and the particular ability of NN to discern complex and non-linear patterns.

Although RPINN significantly outperforms traditional techniques in many aspects, there are domains where its performance is more modest, such as in extremely high noise levels or in the presence of complex non-linear constraints that can hinder RPINN's performance despite its robustness against data variability and noisy inputs. Nonetheless, NN maintains a competitive position compared to classical approaches, which, while facing fewer obstacles in these areas, are not without challenges. In summary, the purpose of this section is to outline a preliminary framework to address optimization challenges, such as the incorporation of constraints and the management of iterative problems, through the use of RPINN.

5 Unsupervised Regularized Physics - Informed Neural Network to Support Nonlinear Constrained Optimization

This chapter addresses two innovative strategies for solving specific nonlinear optimization problems in gas transportation, utilizing Physically Informed Neural Networks (PINN) in an approach we have termed Regularized PINN (RPINN). The first strategy develops dynamic and custom activation functions, designed to ensure compliance with physical limits and accurately model the dynamics associated with unmet demands and gas injections. The second strategy introduces regularization penalties that integrate the peculiarities of NOPT, such as node balance, the Weymouth equation, and the compression ratio, both into the network architecture and the cost function. This effectively guides the training process towards estimating optimal flows and pressures, aiming to minimize operational costs, reduce computation times, and ensure convergence to reliable and efficient results.

The chapter's structure is grounded on a solid mathematical foundation that supports our proposals, followed by a detailed description of the implemented methodology. It concludes with the presentation and analysis of the results obtained through these methodologies, comparing them with those achieved using the IPOPT software via GEKKO, to illustrate their effectiveness and applicability.

5.1 Theoretical Background

Gas network optimization, which involves the efficient management of pressure and flow, commonly encounters challenges stemming from the inherent complexity of NOPT. This can lead to outcomes that fall short of the optimum due to the non-convex nature of these problems. The integration of PINN emerges as a promising solution, recognized for its versatility in generalizing and tackling various challenges, while ensuring adherence to the system's fundamental components. The aim of this section is to elucidate the essential theoretical foundations necessary for understanding the assessments and contributions

discussed in this chapter.

5.1.1 Activation Functions

Activation functions are fundamental in the field of ANN, introducing the essential nonlinearities that enable the modeling of complex interactions in the data. In this section, we review various activation functions that have been widely implemented, highlighting their unique characteristics and application contexts.

We begin with the linear activation function, which stands out for its simplicity, as it does not alter the input. Mathematically, it is represented as:

$$f(x) = x \tag{5-1}$$

In this context, $x \in \mathbb{R}$ represents the input value for the function, essential for understanding the foundations upon which more complex activation variants are developed.

Regarding nonlinear activation functions, the sigmoid function [*KILIÇARSLAN et al.*, 2021] stands out, recognized for its ability to map input values to a range between 0 and 1, making it particularly useful for modeling probabilities. The sigmoid function is defined as:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{5-2}$$

The hyperbolic tangent function (tanh) extends the output range of the sigmoid function to values between -1 and 1 [*Rasamoelina et al.*, 2020]. This function is defined by the following expression:

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{5-3}$$

The Rectified Linear Unit (ReLU) function is widely recognized for its simplicity and effectiveness, particularly in convolutional networks and backpropagation processes [*Sakketou & Ampazis*, 2019]. This function allows positive values to remain unaltered while transforming negative values to zero.

$$\mathrm{ReLU}(x) = \max(0, x) \tag{5-4}$$

The Scaled Exponential Linear Unit (SELU) as a variant of ReLU, designed to promote self-normalization and mitigate some challenges present in the training of deep neural networks [*Sakketou & Ampazis*, 2019]. Its formulation is as follows:

$$\mathrm{SELU}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^{x} - \alpha & \text{if } x \leq 0 \end{cases} \tag{5-5}$$

Where λ and α are predefined parameters.

Finally, the Mish function has been introduced as a superior alternative to ReLU and its variants, designed to mitigate the vanishing gradient problem [*Misra*, 1908]. This function is described by the following equation:

$$\text{Mish}(x) = x \cdot \tanh(\ln(1 + e^x)) \tag{5-6}$$

5.1.2 Loss Function

In machine learning, loss functions are fundamental as they quantify the discrepancy between predictions and actual values, guiding models towards continuous improvement.

One of the most relevant evaluation metrics in the context of regression is the Mean Squared Error (MSE). This metric measures the model's accuracy by calculating the average of the squares of the differences between predictions and actual values, as described in [*Qiu et al.*, 2020]. Mathematically, MSE is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \tag{5-7}$$

In this context, $\hat{y}_i$ denotes the model prediction for the i-th observation, y_i represents the corresponding observed value, and n signifies the total number of observations.

Regarding the accuracy of predictive models, the Mean Absolute Error (MAE) is a widely employed metric. Unlike the Mean Squared Error (MSE), which averages squared errors, the MAE calculates the average of the absolute differences between predictions and observed values. This characteristic makes the MAE less sensitive to outliers compared to the MSE, as documented in [*Qi et al.*, 2020]. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{5-8}$$

Huber loss, which merges aspects of MSE and MAE, as an effective alternative in the presence of outliers, offering a robust error metric that smoothly minimizes errors. This property is particularly useful in contexts with highly anomalous data [*Wang et al.*, 2020a]. The formula for the Huber loss is:

$$L_\delta(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{if } |y - \hat{y}| \leq \delta, \\ \delta(|y - \hat{y}| - \frac{1}{2}\delta) & \text{otherwise.} \end{cases} \tag{5-9}$$

Where δ defines the threshold that distinguishes between the treatment of quadratic and absolute errors. This parameter is crucial for balancing precision in minor errors and tolerance to outliers.

5.1.3 Gas Distribution Networks

The primary objective of this study is to address a nonlinear optimization problem in gas transportation. To this end, it is crucial to familiarize oneself with the structures discussed in this document, where two types of network architectures are examined. The first is a synthetic gas network, comprising 8 nodes, based on the network described by [*García-Marín et al.*, 2019] (see Fig. **5-1**). This network includes 8 nodes, 2 compressors, a gas extraction source, and a storage unit.

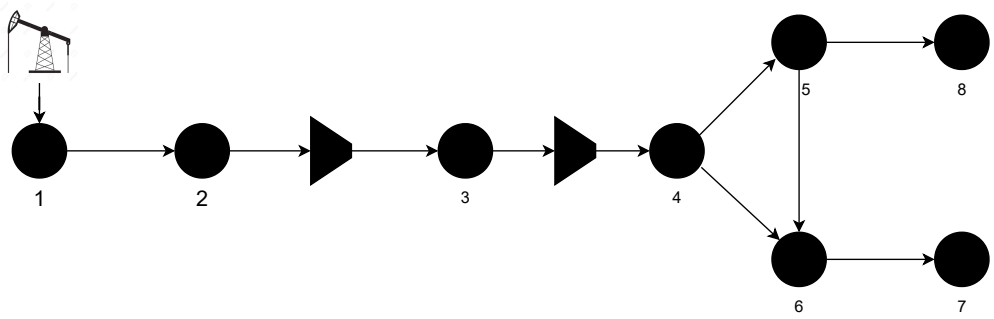

Figure 5-1: 8-node network: The diagram depicts nodes as dots. Arrows symbolize gas pipelines, denoting the flow direction or connectivity between nodes. Trapezoidal shapes represent compressors.

To broaden the comparison base and leverage the features of the 8-node radial network, the number of nodes was increased to 15, as shown in Fig. **5-2**.

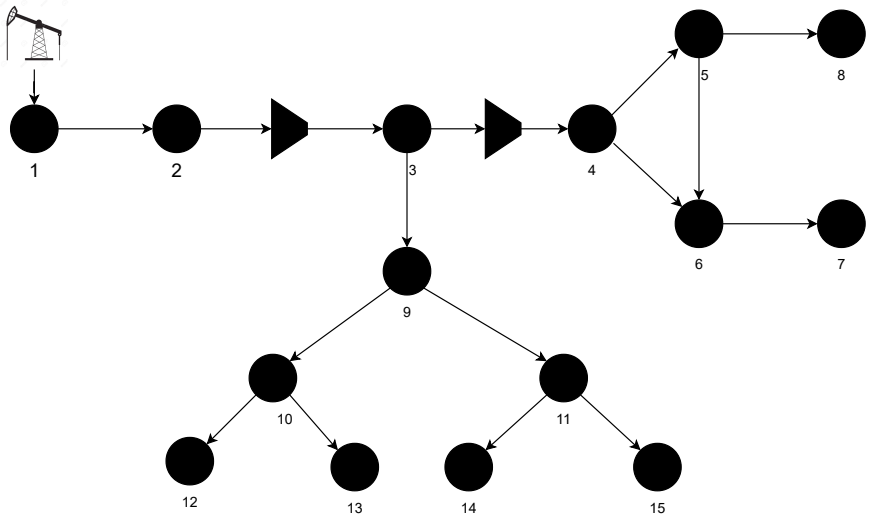

Figure 5-2: 15-node network: The diagram illustrates the expansion from the 8-node network, forming a synthetic network of 15 nodes.

Overall, Figures **5-1** and **5-2** depict a gas extraction network powered by an extraction source. These diagrams meticulously detail the essential components and their interconnections, providing a comprehensive view of the structure and operational flow of the network.

The second gasification network model represents a specific scenario within the Colombian framework, based on the national gas infrastructure [*García-Marín et al.*, 2022]. This network consists of 63 nodes, including 13 extraction points, 14 compression stations, and a storage facility. The distribution and structure of this network are illustrated in Figure **5-3**.

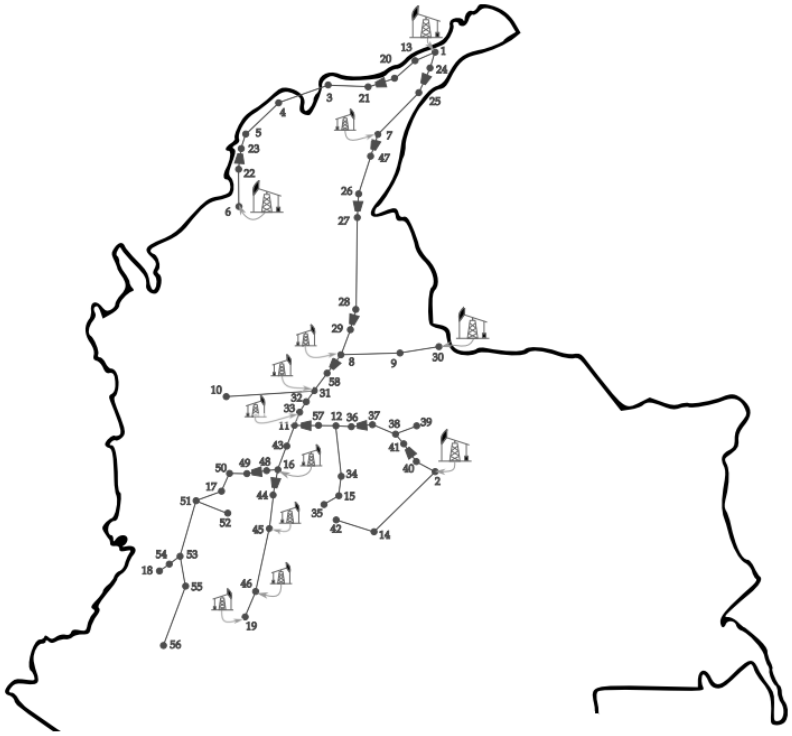

Figure 5-3: Colombian Network: The diagram displays the geographical distribution of Colombia's network along with its elements.

5.1.4 Mathematical Model

In Colombia, the issue of gas transport is currently addressed through an optimization model based on linear programming [*CREG*, 2022a]. The mathematical formulation is described below.

$$\min_{f} \sum_{i=1}^{\dot{W}} \sum_{j=1}^{N} f_{i,j} \times (C_i^w + C_{i,j}^t) \tag{5-10}$$

Here, W represents the number of sources in the system, $C_i^w \in R^{\dot{W}x1}$ and $C_{i,j}^t \in R^{\dot{W}xN}$ denote the extraction cost from the i source and the transportation cost between the i and j nodes, respectively. The variable $f_{i,j} \in R^{\dot{W}xN}$ is the gas flow between source i and node j. N denotes the number of nodes. For this model, the constraints are:

$$\sum_{j=1}^{N} f_{i,j} \leq f_i^w \tag{5-11}$$

This equation means that the gas allocated to each node cannot exceed the capacity of source i, denoted by $f_i^{\dot{w}} \in R^{\dot{W}x1}$, and must match the demand of each node, represented by $L_j \; in R^{Nx1}$, see the following equation:

$$\sum_{i=1}^{\dot{W}} f_{i,j} = L_j \tag{5-12}$$

5.2 Tested Scenarios for NOPT using RPINN

We study a gas-powered system as a function of flow and pressure. For this purpose, a synthetic network of eight nodes is used, as detailed in [*García-Marín et al.*, 2019] and illustrated in Figure **5-1**.

In particular, the NOPT problem is written as:

$$\begin{aligned}
\min_{\mathbf{x},\boldsymbol{\pi}} \quad & \mathbf{x}^\top \mathbf{a} \\
\text{s.t.} \quad & \mathbf{B}\mathbf{x} = \mathbf{z} \\
& x_q = \operatorname{sgn}\left(\pi^2_{w(q)} - \pi^2_{w'(q)}\right) \sqrt{k_q |\pi^2_{w(q)} - \pi^2_{w'(q)}|}, \\
& \forall q \in Q; \quad w(q), w'(q) \in W \\
& \beta_{\min}(n,n') \leq \frac{\pi_n}{\pi_{n'}} \leq \beta_{\max}(n,n'), \quad \forall n, n' \in V \\
& \boldsymbol{\gamma}_{\min} \leq \boldsymbol{\pi} \leq \boldsymbol{\gamma}_{\max} \\
& \boldsymbol{\delta}_{\min} \leq \mathbf{x} \leq \boldsymbol{\delta}_{\max},
\end{aligned} \tag{5-13}$$

where $\mathbf{a} \in \mathbb{R}^P$ represents the gas transport costs for the P flows in $\mathbf{x} \in \mathbb{R}^P$. The incidence matrix $\mathbf{B} \in \mathbb{R}^{W \times P}$ encodes the gas network structure, with W nodes and $\mathbf{z} \in \mathbb{R}^W$ the input gas demand. The first equality constraint is what encodes the linear-based flow and gas demand equilibrium along the network nodes. Next, the node pressure is stored in $\boldsymbol{\pi} \in \mathbb{R}^W$. In turn, the q-th flow $x_q \in \mathbf{x}$ is selected according to the network structure

from $\mathbf{B}$ to fulfill the Weymouth equality with $k_q \in \mathbb{R}$ and $Q \leq P$ [*García-Marín et al.*, 2022]. Then, the function $w(q)$ extracts the related pressure $\pi_{w(q)} \in \boldsymbol{\pi}$ regarding such a Weymouth-based physic constraint. Furthermore, $\pi_n, \pi_{n'} \in \boldsymbol{\pi}$ chose the inlet and outlet pressures to fulfill the system compression ratio, with V components $(n, n' \in \{1, 2, \ldots, V\}$, $V \leq W)$ and compression factor limits $\beta_{\min}(n, n'), \beta_{\max}(n, n') \in \mathbb{R}^+$. Also, $\boldsymbol{\gamma}_{\min}, \boldsymbol{\gamma}_{\max} \in \mathbb{R}^W$ and $\boldsymbol{\delta}_{\min}, \boldsymbol{\delta}_{\max} \in \mathbb{R}^P$ are the minimum and maximum pressure and flow limits, respectively.

Now, let $\{\mathbf{z}_r \in \mathbb{R}^W\}_{r=1}^R$, be an unsupervised input set concerning the required gas demand for R observations. Our RPINN solution of Eq. 5-13 is as follows:

$$\begin{aligned}
\min_{\{\tilde{\mathbf{z}}_r, \tilde{\boldsymbol{\pi}}_r\}_{r=1}^R} & \frac{\lambda_{\mathcal{L}}}{R} \sum_{r=1}^{R} \tilde{\mathbf{z}}_r^\top \mathbf{a} + \frac{\lambda_{L1}}{R} \sum_{r=1}^{R} \tilde{h}_{L1}(\mathbf{B}\tilde{\mathbf{z}}_r, \mathbf{z}_r; \epsilon_{L1}) + \\
& \frac{\lambda_{N1}}{R} \sum_{r=1}^{R} \tilde{h}_{N1}(\tilde{\mathbf{z}}_r, \tilde{\boldsymbol{\pi}}_r; \mathbf{B}, \epsilon_{N1}) + \frac{\lambda_{N2}}{R} \sum_{r=1}^{R} \tilde{h}_{N2}(\tilde{\boldsymbol{\pi}}_r; \mathbf{B}, \boldsymbol{\beta}, \epsilon_{N2}) \\
\text{s.t.} \quad & \lambda_{\mathcal{L}} + \lambda_{L1} + \lambda_{N1} + \lambda_{N2} = 1 \\
& \tilde{f} = \tilde{f}^\dagger \cup \tilde{f}^\ddagger \\
& \tilde{\mathbf{z}}_r = \tilde{f}^\dagger(\mathbf{z}_r | \tilde{\mathbf{X}}^\dagger, \tilde{\mathbf{Z}}^\dagger), \tilde{\boldsymbol{\pi}}_r = \tilde{f}^\ddagger(\mathbf{z}_r | \tilde{\mathbf{X}}^\ddagger, \tilde{\mathbf{Z}}^\ddagger) \\
& \boldsymbol{\gamma}_{\min} \leq \tilde{\boldsymbol{\pi}}_r \leq \boldsymbol{\gamma}_{\max} \\
& \boldsymbol{\delta}_{\min} \leq \tilde{\mathbf{z}}_r \leq \boldsymbol{\delta}_{\max}, \quad \forall r \in R.
\end{aligned} \tag{5-14}$$

Given the r-th gas demand vector $\mathbf{z}_r \in \mathbb{R}^W$, $\tilde{\mathbf{z}}_r \in \mathbb{R}^P$ predicts the flow vector based on $f^\dagger$, and $\tilde{\boldsymbol{\pi}}_r \in \mathbb{R}^W$ the corresponding pressure vector using $f^\ddagger$. Moreover:

$$\begin{aligned}
\tilde{h}_{L1}(\mathbf{B}\tilde{\mathbf{z}}_r, \mathbf{z}_r; \epsilon_{L1}) =& \mathcal{L}_H(\mathbf{B}\tilde{\mathbf{z}}_r, \mathbf{z}_r; \epsilon_{L1}) \\
\tilde{h}_{N1}(\tilde{\mathbf{z}}_r, \tilde{\boldsymbol{\pi}}_r; \mathbf{B}, \epsilon_{N1}) =& \frac{1}{Q} \sum_{q=1}^{Q} \mathcal{L}_H(\tilde{z}_{rq}, \varphi_q(\tilde{\boldsymbol{\pi}}_r; \mathbf{B}); \epsilon_{N1}) \\
\tilde{h}_{N2}(\tilde{\boldsymbol{\pi}}_r; \mathbf{B}, \boldsymbol{\beta}) =& \frac{1}{V^2} \sum_{n,n' \in V} \mathcal{L}_{\tilde{H}}(\tilde{\pi}_{rn}, \tilde{\pi}_{rn'}; \mathbf{B}, \boldsymbol{\beta}, \epsilon_{N2}),
\end{aligned} \tag{5-15}$$

where notation $\mathcal{L}_H(\cdot, \cdot; \epsilon.)$ stands for a Huber-based penalty (see Eq. 4-24), $\varphi(\tilde{\boldsymbol{\pi}}_r; \mathbf{B}) \in \mathbb{R}^Q$ holds elements:

$$\varphi_q(\tilde{\boldsymbol{\pi}}_r; \mathbf{B}) = \operatorname{sgn}\left(\tilde{\pi}^2_{rw(q)} - \tilde{\pi}^2_{rw'(q)}\right) \sqrt{k_q |\tilde{\pi}^2_{rw(q)} - \tilde{\pi}^2_{rw'(q)}|}, \ \forall q \in Q, \tag{5-16}$$

and:

$$\mathcal{L}_{\tilde{H}}(\tilde{\pi}_{rn}, \tilde{\pi}_{rn'}; \mathbf{B}, \boldsymbol{\beta}, \epsilon_{N2}) = \begin{cases} 0 & \beta_{\min}(n,n') \leq \frac{\tilde{\pi}_{rn}}{\tilde{\pi}_{rn'}} \leq \beta_{\max}(n,n') \\ \left(\frac{\frac{\tilde{\pi}_{rn}}{\tilde{\pi}_{rn'}} - 0.5\beta_{\min}(n,n')}{\beta_{\max}(n,n')} \right)^2 & \text{otherwise.} \end{cases} \tag{5-17}$$

It is worth mentioning that the custom penalty in Eq. 5-15 aims to deal with noisy inputs while preserving the NOPT limits and constraints. In particular, $\mathcal{L}_{\tilde{H}}(\cdot, \cdot; \mathbf{B}, \boldsymbol{\beta}, \epsilon_{N2})$ penalizes pressures that are far from the middle of the compression factor range, according to $\beta_{\min}(n,n'), \beta_{\max}(n,n') \in \boldsymbol{\beta}$.

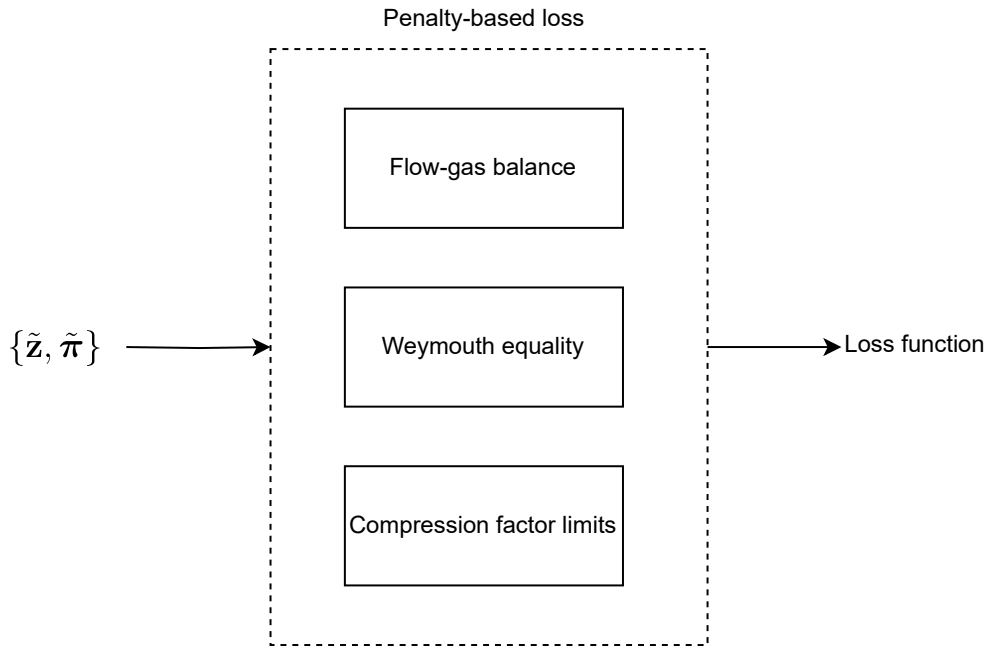

Figure 5-4: Loss Functions: The diagram summarizes the loss functions used to train the PINN, which are aimed at ensuring compliance with the restrictions of the balance of the nodes, Weymouth and the limitations of the compression ratio.

To ensure compliance with model constraints, especially limit constraints, it is necessary to keep the solution within acceptable ranges. For this purpose, a scaled sigmoid function $\tilde{\sigma}(\cdot) \in [u_{\min}, u_{\max}]$ is used to address the predicted flow and pressure limits in Equation 5-14, as follows:

$$\tilde{\sigma}(x) = \alpha \frac{1}{1 + e^{-x}} + \iota, \tag{5-18}$$

where $\alpha, \iota \in \mathbb{R}$. Additionally, the constraints modeling the dynamics of unmet gas supply and demand are considered. These constraints require flexible adaptability to variations in input parameters, meaning they must be dynamic functions.

5.2.1 Activation Function for Modeling Gas Supply

Modeling the behavior of gas in the distribution network is crucial, as it allows optimizing the capacity of the sources to meet all gas demands as much as possible. Therefore, this section introduces the concept of a dynamic activation function for a sample r:

$$\tilde{z}_{lp(s)} = \begin{cases} 0 & \text{if } \mathbf{z}^T\mathbf{1} \leq 0, \\ \mathbf{z}^T\mathbf{1} & \text{if } 0 < \mathbf{z}^T\mathbf{1} < \delta_{p(s)},\ \forall p(s) \in P, \\ \delta_{p(s)} & \text{if } \mathbf{z}^T\mathbf{1} \geq \delta_{p(s)}, \end{cases} \tag{5-19}$$

This activation function ensures that the flow from the extraction source $\tilde{z}_{lp(s)}$ is always greater than zero and less than $\delta_{p(s)} \in \boldsymbol{\delta}$, which essentially corresponds to the maximum source capacity. This ensures that these sources do not exceed their capacity.

5.2.2 Activation Function for Modeling Unsupplied Gas

To analyze the behavior of unsupplied gas flows, this section introduces a dynamic activation function, designed to flexibly adapt to system variations. This function is implemented through a piecewise activation structure, formulated as follows for a sample r:

$$\tilde{z}_{lw(d)} = \begin{cases} 0 & \text{if } \mathbf{z}^T\mathbf{1} \leq \tilde{z}_{p(s)} \\ \tilde{z}_{l_{-1}w(d)} & \text{if } \mathbf{z}^T\mathbf{1} > \tilde{z}_{p(s)} \text{ and } 0 < \tilde{z}_{l_{-1}w(d)} \leq z_{w(d)}\ \forall w(d) \in W \\ z_{w(d)} & \text{if } \mathbf{z}^T\mathbf{1} > \tilde{z}_{p(s)} \text{ and } \tilde{z}_{l_{-1}w(d)} > z_{w(d)} \end{cases} \tag{5-20}$$

The objective of this structure is to manage two critical scenarios within the system. First, it ensures that when the total system demand ($\mathbf{z}^T\mathbf{1}$) is within the capacity limits of the sources, the unsupplied gas flows $\tilde{z}_{lw(d)}$ remain at zero, reflecting the system's ability to fully meet the demand under such conditions. Second, when the demand exceeds these capacities, a supply deficit is generated, resulting in a non-zero unsupplied demand flow. In these cases, it is crucial for the RPINN to accurately evaluate this variable. To this end, a partial output $\tilde{z}_{l_{-1}w(d)}$ is used to estimate $\tilde{z}_{lw(d)}$, saturating the values between 0 and the specific demand of each node $z_{w(d)}$ to maintain node balance.

It is essential to highlight that the activation functions, both for supply and non-supply of gas, are determined by the demand values at the nodes $\mathbf{z}$. This means that the activation fluctuates throughout the network in each iteration of each season during training, which is also considered when making predictions and distinguishes these functions from those with fixed intervals. We have termed this approach dynamic activation. Figure **5-5** illustrates our contribution in the area of activation functions.

Fig. **5-6** provides a general synthesis of our contribution, summarizing the entire scope of the idea presented in this section.

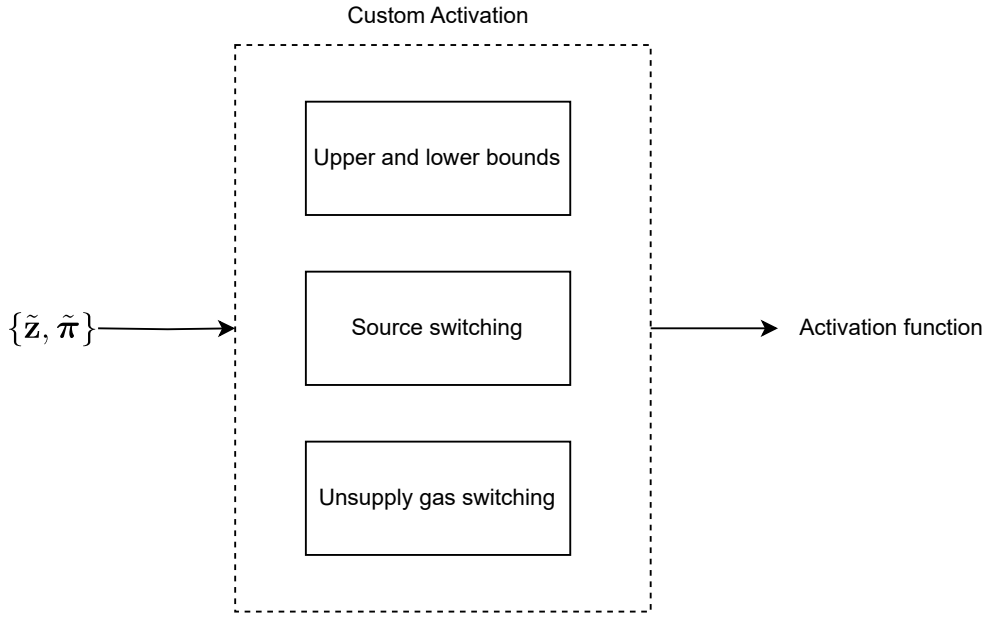

Figure 5-5: Activation functions: The diagram illustrates the suggested functions that convert an input x into $\bar{x}$ to comply with the supply, unsupplied gas constraints, and the physical limitations of the variables.

5.3 Experimental Set-Up

The scenarios in Section 5.2 will be used to test our RPINN in unsupervised settings. They will be utilized to look at sample variability, noisy input measurements, and nonlinear constraints.

5.3.1 Deep Learning Architectures

To address the uniform mixture model NOPT (unsupervised constrained optimization), our RPINN consists of two dense layers, as shown in Figure **4-11** and Table **4-1**.

Next, as seen in Figure **5-7** and Table **5-1**, a wide ANN architecture is proposed for our RPINN-based gas-powered system scenario. We can specifically focus on essential variables—flows and pressures—in our sketch, adapting it to the unique characteristics of the gas network. To achieve this, our model incorporates blocks of dense layers designed to map input data, as well as batch normalization layers that help stabilize and normalize the features and gradient along the back-propagation. Additionally, it includes custom layers named: custom dense, bounded dense, source switching, and unsupply gas switching. We

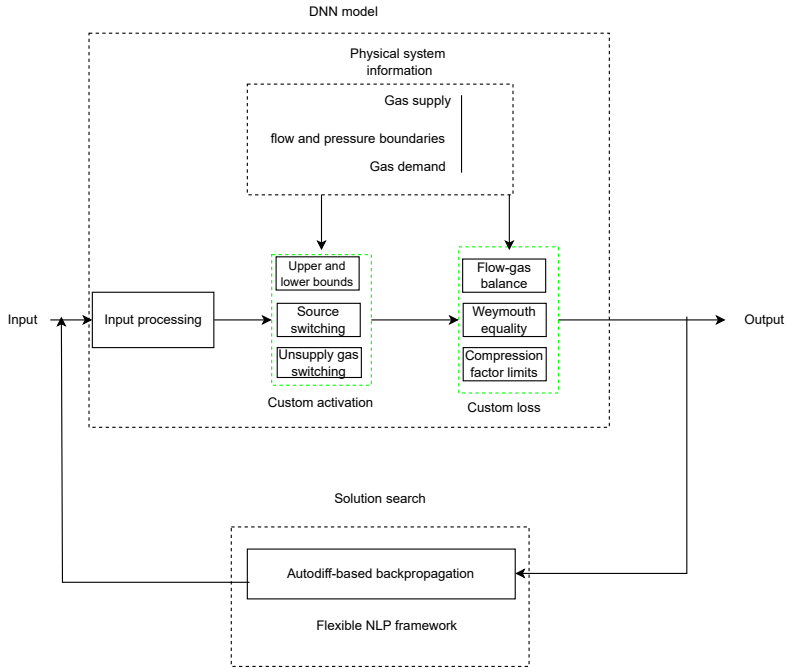

Figure 5-6: Enhanced PINN for NLP-based Gas System Optimization: The figure displays the overall scheme of the proposed approach to tackle the problem discussed in this chapter, highlighting in green the two strategies designed to deal with the constraints of the NLP problem in gas transportation.

design these to encode the source behavior of the system, manage unmet demand, and delineate system boundaries.

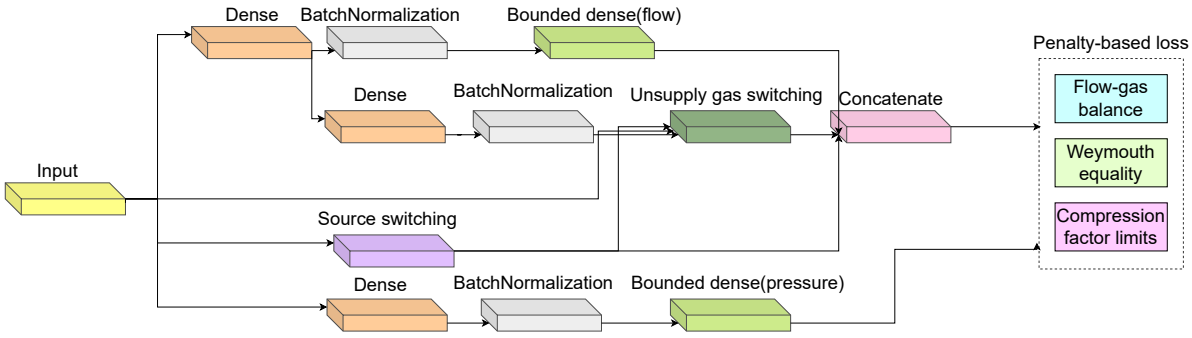

Figure 5-7: RPINN pipeline for the gas-powered system-based NOPT

As seen, a shallow and straightforward approach is a simple NOPT task for the uniform mixture model, which we must elucidate in relation to fixed architectures. Additionally, in order to mitigate overfitting and accommodate numerous constraints and a linear loss in the gas-powered system, we implemented a shallow and wide network. Nevertheless, our RPINN approach is adaptable in terms of network architecture, enabling the implementation of more complex schemes as needed.

Table 5-1: RPINN architecture details for the gas-powered system NOPT. $\tilde{R}$: batch-size for AD-based back-propagation. Source switching, unsupply gas switching, custom dense, and bounded dense stand for specific switching, limited, and scaled layers, as explained in Section 5.2. Param. #: number of trainable parameters. Total # of parameters: 12707 (49.67K6B)

Layer name	Type	Output shape	Param. #	Memory size
Input	InputLayer	($\tilde{R}$, 8)	0	0Kb
Dense_1	Dense(SELU)	($\tilde{R}$, 236)	2124	8.3Kb
Dense_2	Dense(SELU)	($\tilde{R}$, 8)	1896	7.41Kb
Source switching	CustomDense	($\tilde{R}$, 1)	1	4B
BatchNormalization_1	BatchNormalization	($\tilde{R}$, 236)	944	3.69Kb
BatchNormalization_2	BatchNormalization	($\tilde{R}$, 8)	32	0.12Kb
Partial flows	BoundedDense	($\tilde{R}$, 50)	2274	8.88Kb
Unsupply gas switching	CustomDense	($\tilde{R}$, 8)	0	0Kb
Flow prediction	Concatenate	($\tilde{R}$, 59)	0	0Kb
Dense_3	Dense(SELU)	($\tilde{R}$, 236)	2124	8.3Kb
BatchNormalization_3	BatchNormalization	($\tilde{R}$, 236)	944	3.69Kb
Pressure prediction	BoundedDense	($\tilde{R}$, 8)	1896	7.41Kb
Node balance	CustomDense	($\tilde{R}$, 8)	472	1.84Kb
Weymouth	CustomDense	($\tilde{R}$, 14)	0	0Kb

5.3.2 Training Details and Method Comparison

To evaluate the effectiveness of our methodology in addressing optimization problems, we utilized two performance measures across all conducted experiments: MAPE (see Eq. 4-26), and the mean absolute error (MAE), defined as:

$$\text{MAE}(\tilde{y}_r, \hat{y}_r) = \frac{1}{R}\sum_{r=1}^{R} |\tilde{y}_r - \hat{y}_r| \,, \tag{5-21}$$

where $\tilde{y}_r, \hat{y}_r \in \mathbb{R}$ stands for r-th target and predicted value, $\text{MAE}(\cdot,\cdot) \in \mathbb{R}^+$, and $|\cdot|$ is the absolute value operator.

Now, for the gas-powered system, we defined three distinct scenarios to evaluate the network's capacity under varying demand conditions. This process yielded a total of 20.000 samples, of which 30% were designated for testing. We produced 320 samples using GEKKO v1.0.6 to compare the model's performance with IPOPT v3.12 [*Owerko et al.*, 2022].

We implemented RPINN using Python 3.10.12 and the TensorFlow API 2.15.0 on Google Colaboratory. For training, we fixed 600 epochs, a batch size of 256 and an Adamax optimizer. Also, an initial learning rate of 1×10^{-2} with decreasing scheduling was employed. The regularization hyperparameters, namely $\lambda_{\cdot}$ in Eq. 4-16, were experimentally fixed within the range $[0, 1]$. Since IPOPT excels at solving NOPT, not to mention its open access, we

fix it as a method comparison [Ipo, 2024]. Our codes and studied datasets are publicly available at `https://github.com/UN-GCPDS/python-gcpds.optimization` (accessed on 1 March 2024).specific conditions.

5.3.3 Synthetic Gas Networks

Scalability evaluation was conducted on networks of 8 and 15 nodes. Subsequently, the nodes of the 15-node network were grouped to compact it, allowing for the evaluation and comparison of metrics between the classical IPOPT technique and our initiative in networks of 8 to 15 nodes.

For the analysis of the Colombian scenario, the network structure shown in Fig. **5-3** was employed. Using its radial configuration, the nodes were reorganized to form an 18-node network, detailed in Fig. **5-8**.

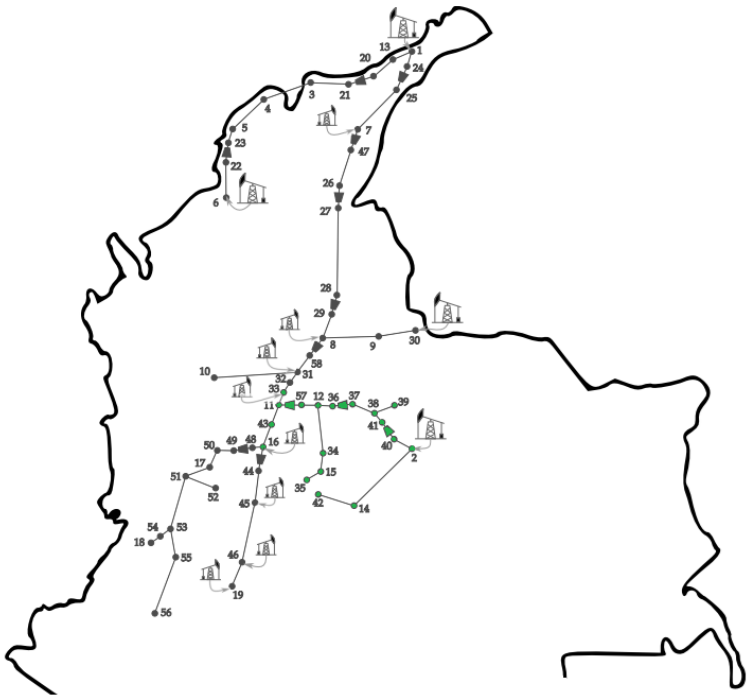

Figure 5-8: Colombian Network: The diagram shows in green the elements used to validate our proposal in the network of Colombia with their respective geographical distribution

The method was evaluated on gas networks containing up to three compressors, under the assumption of no gas storage and a single supply node. Additionally, it was considered that the gas demand at all nodes was exclusively residential and followed a uniform distribution.

5.4 Results and Discussion

5.4.1 Loss Function Selection

In Fig. **5-9**, the behavior of the MSE, MAE, and Huber cost functions is illustrated while addressing two fundamental constraints: node balance and the Weymouth equation, and their impact on the transportation costs within our NOPT model. The results indicate that the Huber cost function offers superior performance, achieving a significant reduction in operational costs.

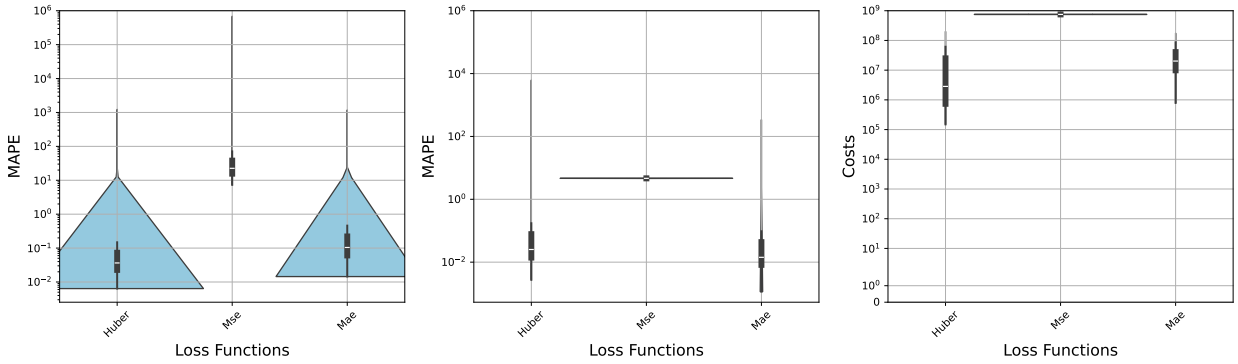

Figure 5-9: The chart presents the results of performance evaluation in three critical aspects. From left to right, the first image shows the performance of the three loss functions when evaluating the balance of nodes. The second image shows the result of the Weymouth equation, and the third, the costs of the solutions.

5.4.2 Dynamic activation functions

Figure **5-10** illustrates the behavior of our customized activation functions, designed to capture the dynamics of unmet supply and demand. The figure shows that when the system exceeds supply capacity, depletion occurs and the shortage flows are not zero. This phenomenon does not appear when assuming fixed limits for unmet demand, represented in the graph by the curve labeled 'Unsupplied Gas', which violates the node balance constraint. Finally, comparing the sum of the dynamic supply function and the unmet flows with the sum without these functions demonstrates that this application better fits the real behavior of the flows.

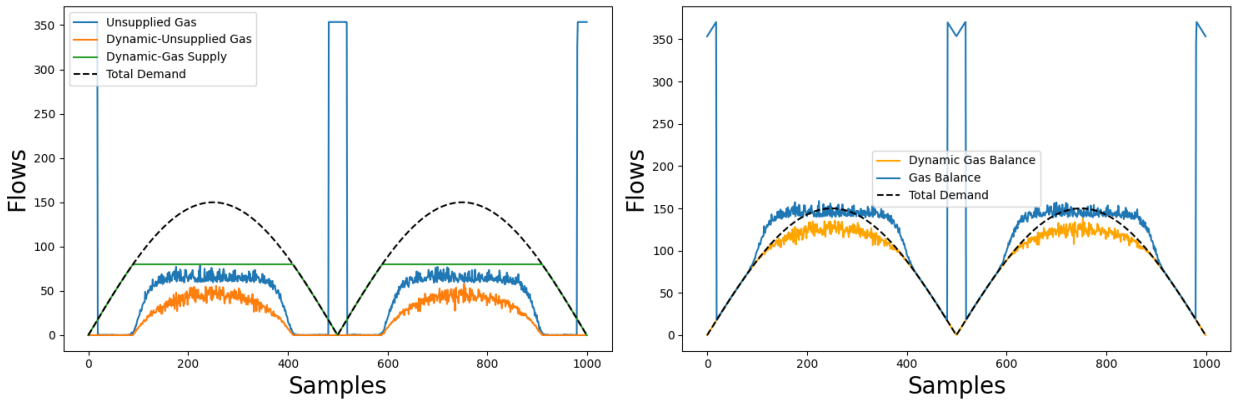

Figure 5-10: The figure illustrates the behavior of dynamic activation functions in a simulated scenario of oscillating demand. In the figure on the left, the behavior of the injection flow is observed when the source capacity is exceeded, and how unsupplied demand ceases to be zero in this context. On the other hand, to the right, the sum of supply flows with and without the proposed dynamic function is presented.

5.4.3 Unsupervised Constrained Optimization Results

Figure **5-11** depicts our RPINN regularized penalty illustration for the gas-powered system NOPT. We adopted a standard variant of the Huber loss for the node balance and Weymouth constraints. As shown, the threshold ϵ. transitions between the $l1$ and $l2$ norms. Regarding the constraint on the compression ratio limit, it was essential to alter the structure due to its inequality behavior. This enhancement stabilizes the transition between the $l2$ and $l1$ norms at zero, based on the distance to the required range's central value. Furthermore, it is crucial to correctly integrate these cost functions into our RPINN. Then, the right plot in Figure **5-11** shows the Weymouth (blue), compression ratio (orange), and compression factor constraint (green) penalty evolution. The resulting loss shows a decreasing trend, indicating that the Huber-based approach can handle the physical limitations of the gas-powered NOPT.

In turn, We have designed three evaluation scenarios to compare the IPOPT framework with our approach and validate the performance of regularization functions in data generation. In the first scenario, the data remain below the source's maximum capacity. In the second scenario, 50% of the samples exceed this capacity, while in the third, about 100% of the data surpass it. Figure **5-12** shows that, although IPOPT has a lower MAPE, its precision (variance) varies considerably over the iterations. This indicates that conventional methods for NOPT are not robust against data variability and nonlinear constraints. In contrast, our RPINN achieves acceptable MAPE with low variability across all experiments, thanks to its regularized strategy based on ANN. In the first two cases, traditional solutions to the Weymouth equation outperform our proposal. However, in the third case, our proposal is superior due to its greater stability and lower sensitivity to outliers, thanks to the Huber-based penalty. It is important to note that, although MAPE is a valuable metric for evaluating model performance, its sensitivity to scale can distort the perception of error when actual

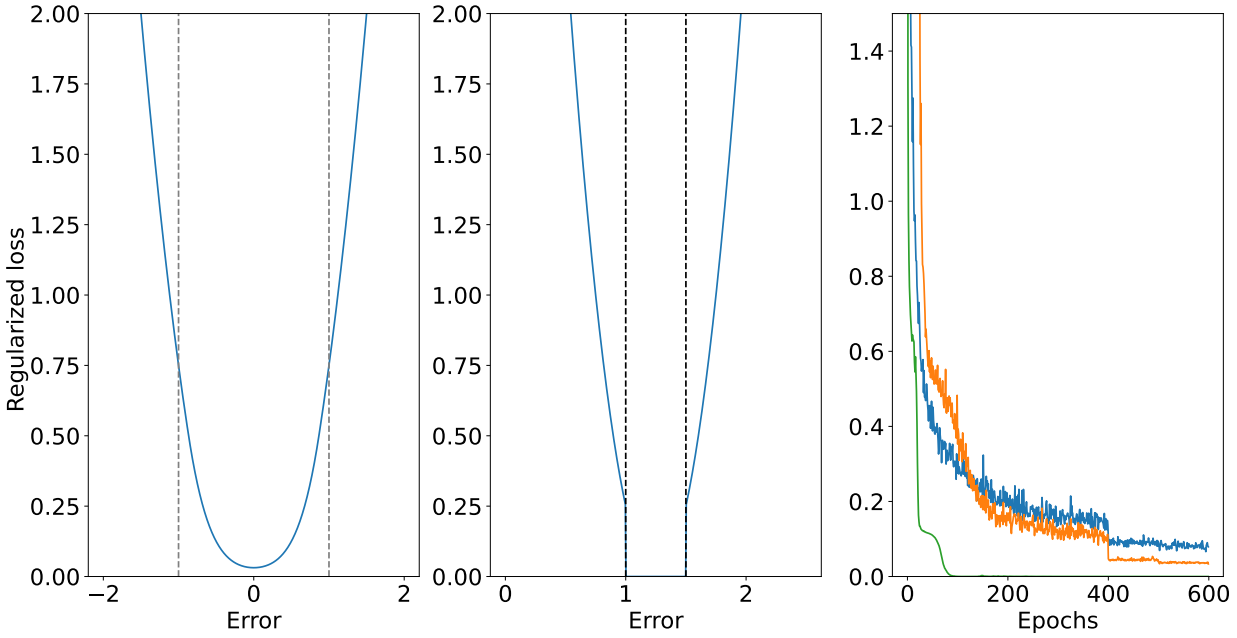

Figure 5-11: Gas-powered system regularized loss illustration. Left: node balance and Weymouth penalties based on conventional Huber-loss. Middle: Compression factor limit constraint using our Huber-based enhancement. (see Eq. 5-15). Right: Gas-powered system custom penalty evolution (Blue: Weymouth equality constraint; Orange: compression ratio limit constraint; Green: Compression factor constraint).

values are extremely small. To address this limitation, we also used the Mean Absolute Error (MAE) in our analysis, as illustrated in Figure **5-13**. The figure shows that our RPINN approach maintains low variability in MAE compared to IPOPT. This reinforces the robustness of our regularized strategy based on ANNs, particularly in scenarios where data values vary significantly. Both approaches share similar costs and adhere to compression ratio constraints.

Finally, to see if our RPINN model can handle the limits in Eq. 5-14 well, we looked at the results of the flow and pressure prediction layers and how they behaved, as shown in Figure **5-14**. The parameters analyzed, including injection and pipe flows as well as pipeline pressures, remain within acceptable limits. This behavior is attributed to custom activation in Eq. 5-18, which ensures a smooth and steady transition between the established ranges.

A crucial aspect to consider is the detection of outliers in both the nodal balance constraint and the Weymouth equation. In this study, we established a criterion that accepts deviations below 1×10^{-1}, analyzing the data that exceeded this threshold. Figure **5-15** shows that, in our approach, less than 1% of the data surpass this threshold, except in the specific network of Colombia, where it reaches up to 30%.

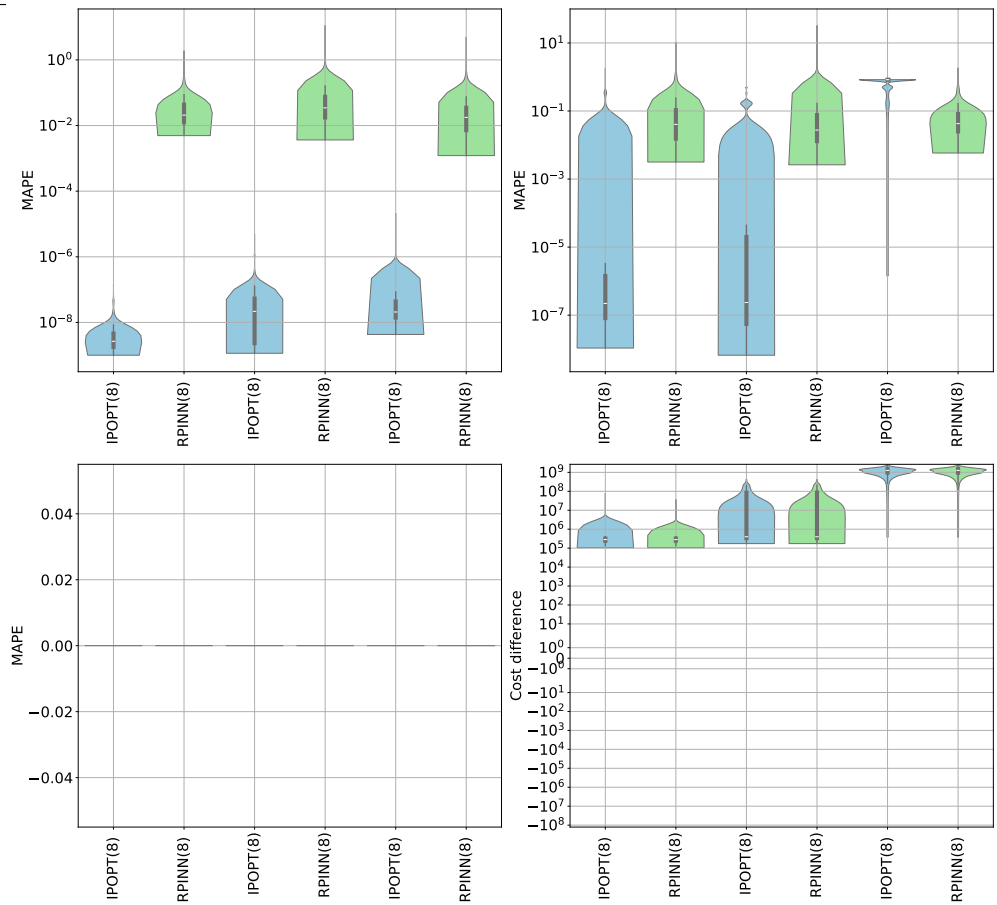

Figure 5-12: Gas-powered system objective cost and constraint compliance MAPE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function) between RPINN and IPOPT.

5.4.4 Convergence

The convergence rate is a crucial factor in ensuring the reliability of solutions to optimization problems. Figure **5-16** shows the failure rate in convergence using IPOPT across various network configurations and different data variations. It can be observed that as variability increases, so does the rate of non-convergence. Similarly, where IPOPT did not provide a response, our RPINN-based approach maintains the trend observed in previous results, see Figure **5-13**, as shown in Figure **5-17**. It is important to highlight that these results are satisfactory, as the conventional technique failed to produce a solution.

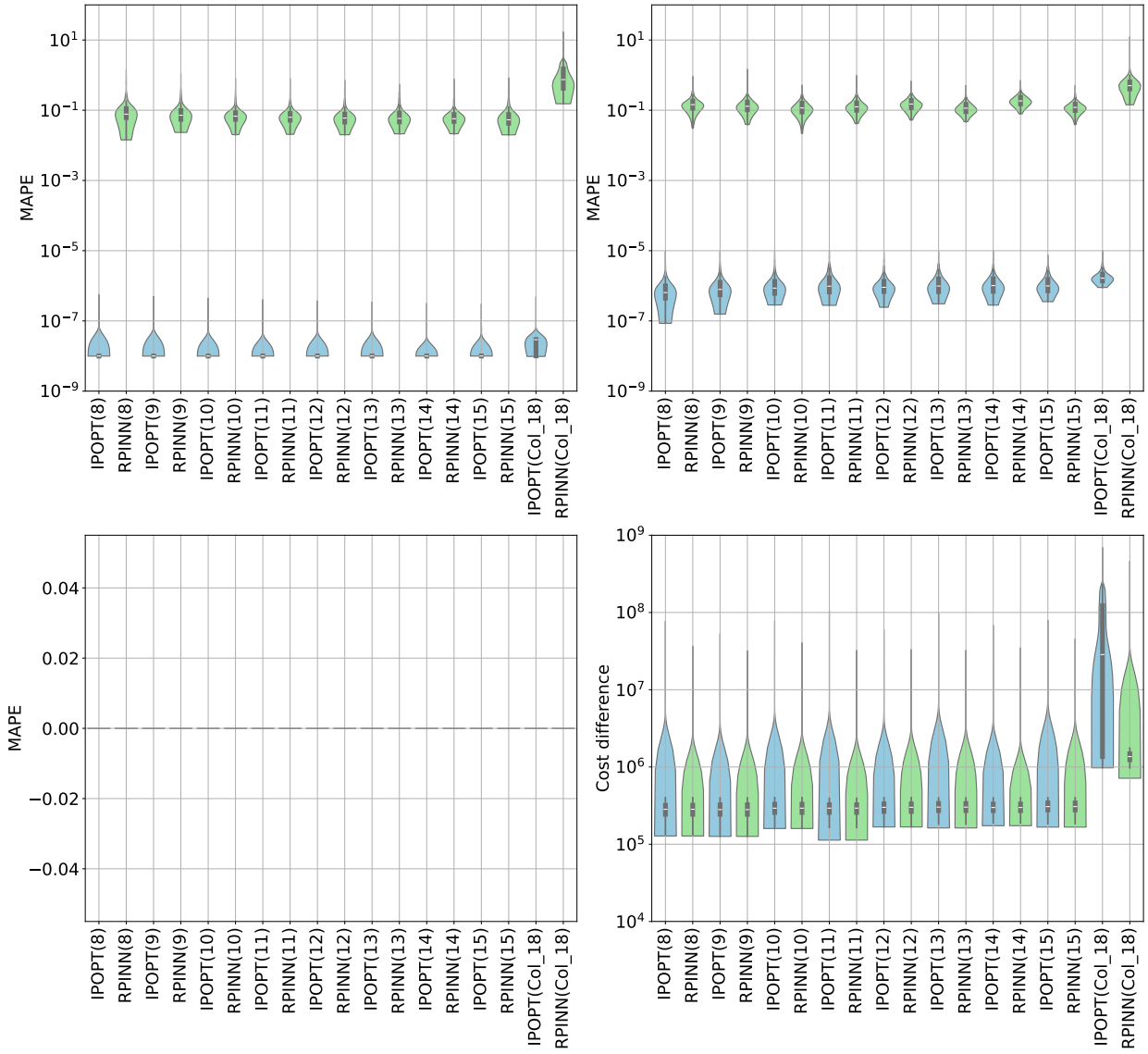

Figure 5-13: Gas-powered system objective cost and constraint compliance MAE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function) between RPINN and IPOPT.

5.4.5 Computational Cost Results

Figure **5-18** shows the training and prediction times needed by the RPINN compared to IPOPT. Our model needs more time to process during the training phase because it has to do both forward and backward passes in each iteration within an ANN-based framework. However, in the prediction phases, our RPINN outperforms IPOPT, resulting in significantly shorter prediction times. This is due to the fact that our approach only requires forward passes after weight training. These tools demonstrate the RPINN's capability to generate fast and accurate predictions for NOPT solutions, not only by reducing processing times but

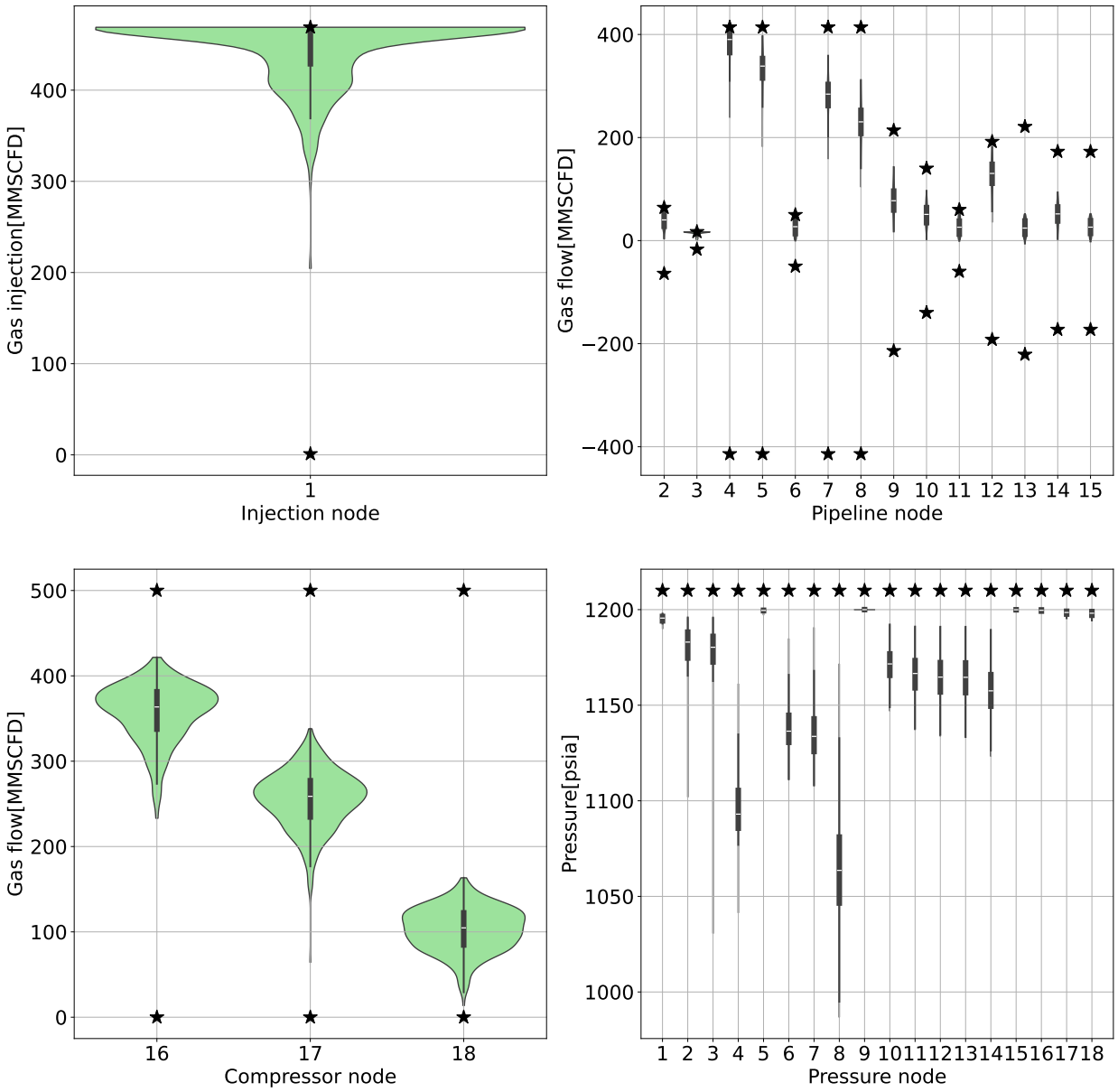

Figure 5-14: Gas-powered system bound constraint MAPE results. The star symbol on this graph denotes the defined limits for each of the sources, compressors, pipelines, and pressures as well as their behavior. The number on the x-axis indicates the node to which the information belongs. MMSCFD: Million standard cubic feet per day. psia: pounds per square inch absolute.

also by narrowing interquartile ranges. Figure **5-19** shows the times across different networks, where it can be seen that our proposal maintains stable computation times regardless of the network size. In contrast, IPOPT's times tend to grow exponentially as the number of nodes increases, highlighting the scalability of our approach.

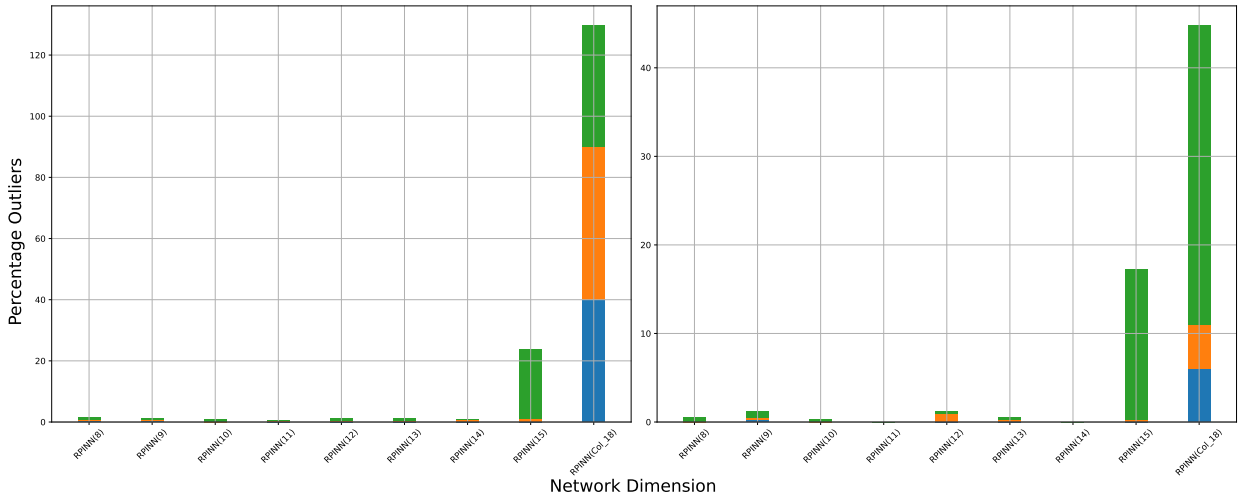

Figure 5-15: The graph compares the samples that exceed 1×10^0 in our strategy. The data that remain below the network's capacity are represented in blue. In orange, the data where 50% of the samples exceed the capacity are shown, and in green, those that completely surpass it. On the left, the results of the node balance are presented, while on the right, the corresponding ones to the Weymouth equation.

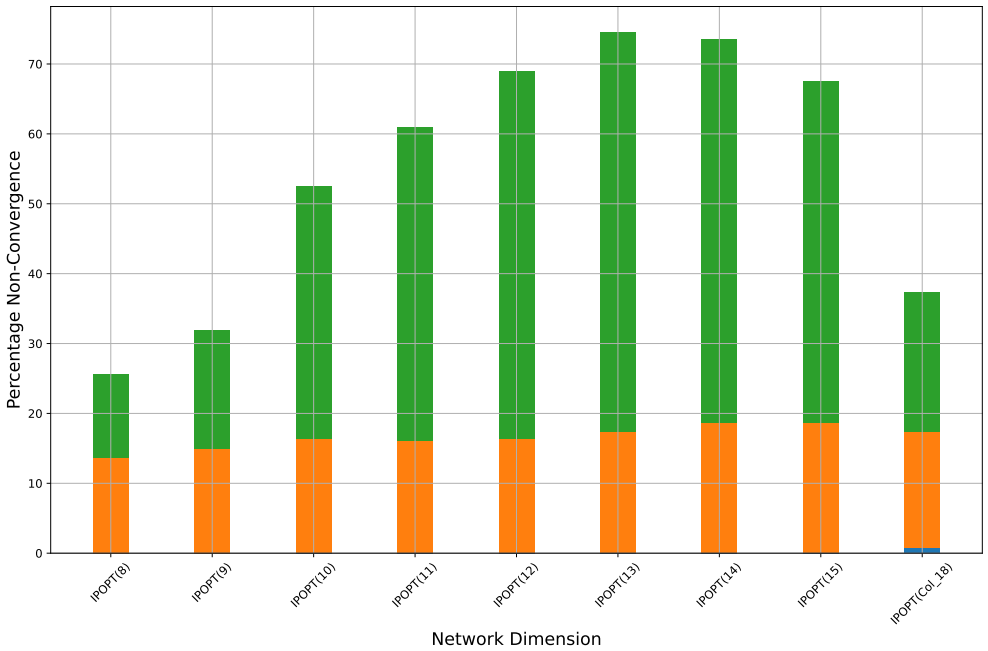

Figure 5-16: Illustration representing the occurrence of convergence failures in each gas network, with numerical values indicating the nodes of each network.The data that remain below the network's capacity are represented in blue. In orange, the data where 50% of the samples exceed the capacity are shown, and in green, those that completely surpass it.

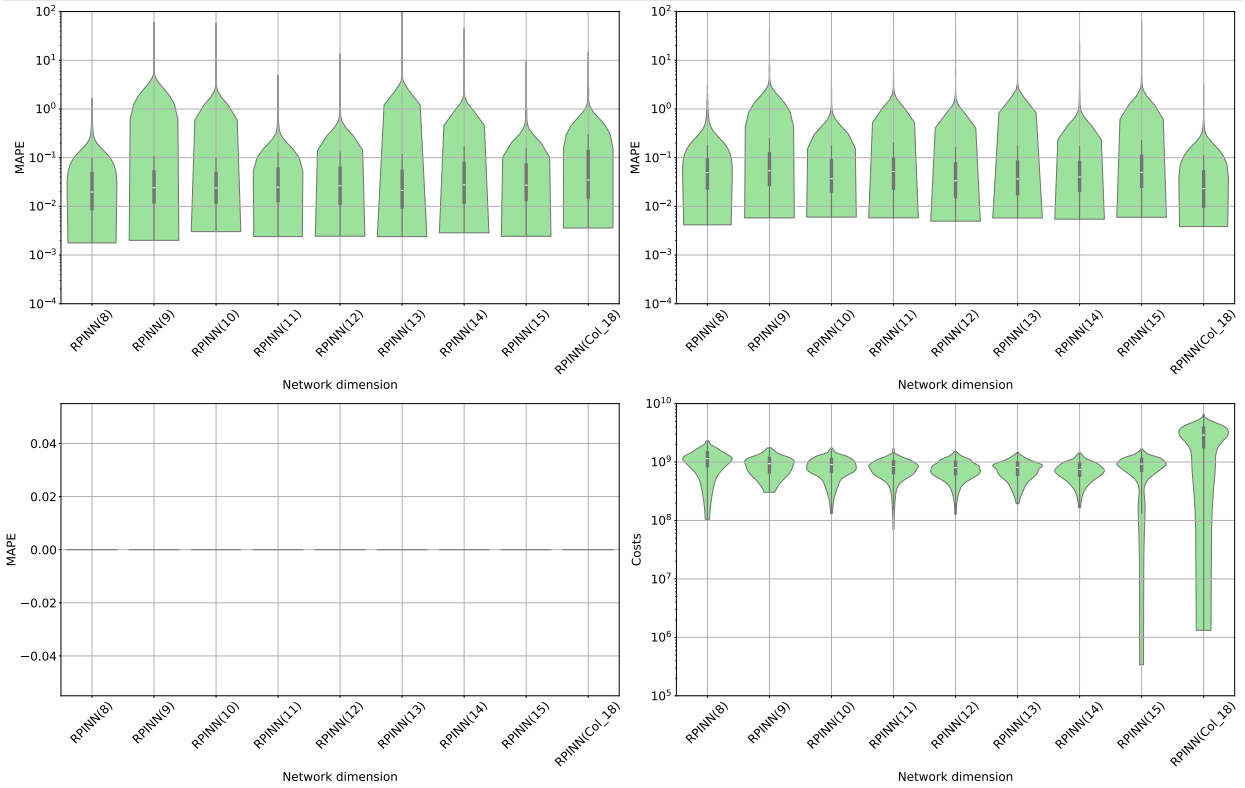

Figure 5-17: Gas-powered system objective cost and constraint compliance MAE results. Upper left: node balance. Upper right: Weymouth constraint. Bottom left: compression ratio constraint. Bottom right: cost difference (objective function).

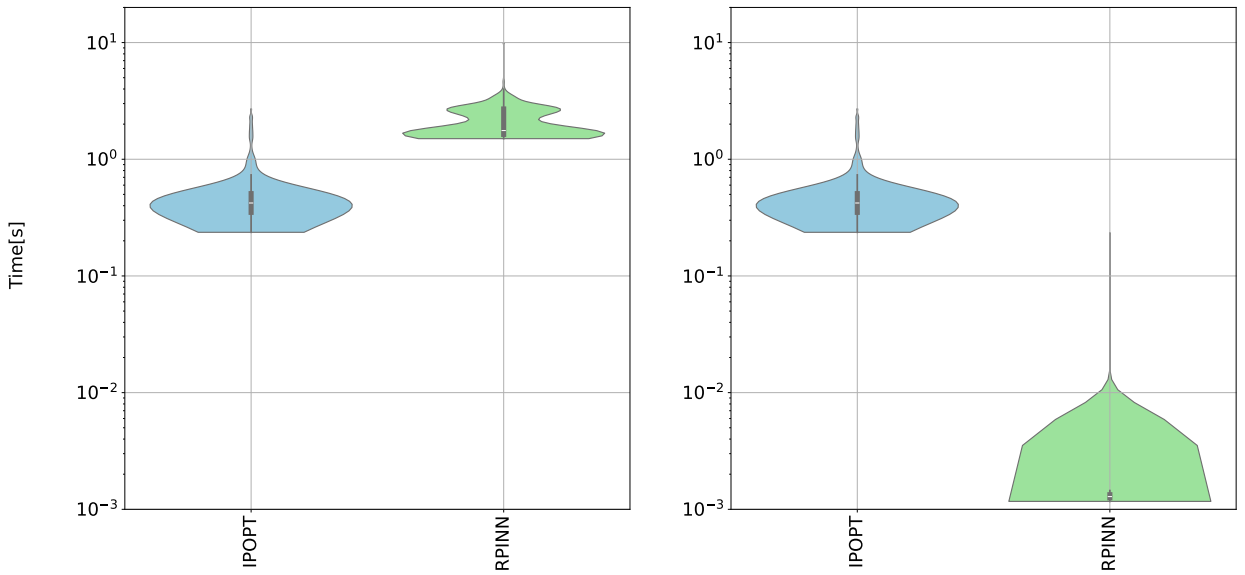

Figure 5-18: RPINN vs. IPOPT computational cost results. The graph compares solution times for the test data between the classical technique (IPOPT, in blue) and our strategy (RPINN, in green). On the left, the training times are shown, while on the right, the prediction times are displayed.

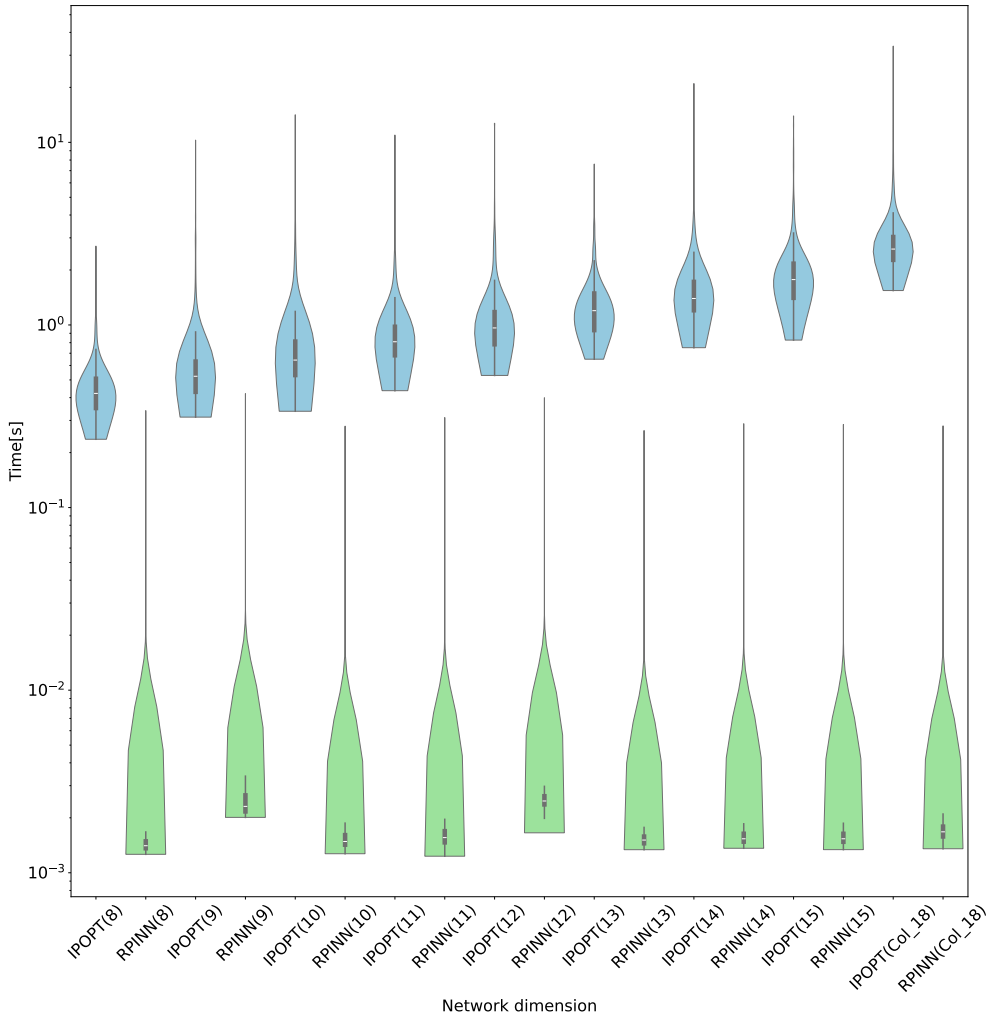

Figure 5-19: RPINN vs. IPOPT computational cost results. The graph compares solution times for the test data between the classical technique (IPOPT, in blue) and our strategy (RPINN, in green).

5.5 Summary

In this section, we present a novel regularized PINN framework, termed RPINN, as an NOPT optimization tool for unsupervised data-driven scenarios. As a result, we address three key NOPT issues. First, we tackle data variability and noisy input measurements by appropriately adapting custom activation and regularization penalties within an ANN scheme. Second, we effectively integrate nonlinear constraints into the network architecture, adhering to the principles of model physics. Specifically, we utilize the network weights and/or learned features within a functional composition framework to determine the NOPT variables. Third, our ANN-based strategy employs automatic differentiation (AD) training, which favors system scalability and computational time through batch-based back-propagation. Experimental results from unsupervised data-driven NOPT tasks confirm that our proposal is robust and competitive against state-of-the-art optimization approaches. The primary advantage of our proposal lies in its stability against noisy input measurements, making it a particularly valuable solution in contexts with fluctuating information. Furthermore, because RPINN is based on ANN, it offers flexibility in terms of network architecture.

6 Final Remarks

6.1 Conclusions

We introduce a novel Regularized Physics-Informed Neural Network (RPINN) framework, named RPINN, presenting a significant advancement in addressing the challenges associated with nonlinear constrained optimization. By integrating custom activation functions and regularization penalties within an ANN architecture, RPINN effectively handles data variability and noisy inputs. Besides, the incorporation of physics principles into the network architecture allows for the computation of optimization variables based on network weights and learned features, leading to competitive performance compared to state-of-the-art solvers. Furthermore, the use of automatic differentiation for training enhances scalability and reduces computation time, making RPINN a robust solution for various NOPT tasks. Experimental results included two scenarios regarding supervised and unsupervised datasets.

The uniform mixture model experiments (supervised constrained NOPT) show that the RPINN is good at dealing with data variability and noisy samples. For noise-free data, both RPINN and the IPOPT solver achieved similar results due to the convex nature of the problem. Still, in scenarios with noisy inputs, RPINN significantly outperformed IPOPT. The RPINN framework, leveraging the Huber loss function, showed greater robustness against noise by effectively regularizing the network weights. This resulted in more accurate and stable output predictions compared to IPOPT, which relied on an objective function based on the $l2$-norm and was more sensitive to outliers. The RPINN weight distributions were concentrated, which showed that the model could find the main output dynamics even when noise was present, as shown by the lower mean absolute percentage error across all signal-to-noise ratio values.

Then, the results of the gas-powered system (unsupervised constrained optimization) highlight the capability of the RPINN framework to effectively manage complex, nonlinear constraints under varying conditions of gas demand. Compared to the IPOPT framework, the RPINN showed consistent performance with low changes in the mean absolute percentage error. This was especially true when the gas demand was higher than the source's maximum capacity. While IPOPT showed lower MAPE in terms of node balance and Weymouth constraints, its

precision fluctuated significantly with data variability. In contrast, RPINN maintained stable performance, ensuring compliance with physical constraints such as the Weymouth equation and compression ratio limits. The custom penalty functions within RPINN facilitated this stability, proving particularly valuable when traditional methods struggled with outliers and extreme values. Overall, RPINN offered a robust, scalable solution with reduced prediction times.

6.2 Future Work

As future work, authors plan to include Bayesian hyperparameter optimization for RPINN fine tuning [*Eleftheriadis et al.*, 2023]. We will also look at normalized and information theoretic learning-based loss as ways to deal with noisy inputs and complicated constraints [*Ma et al.*, 2020b, *Jeon & Van Roy*, 2022]. Finally, Bayesian PINN and graph neural networks will be coupled with our RPINN for representation learning enhancement [*Yang et al.*, 2021, *Thangamuthu et al.*, 2022].

6.3 Academic Products

6.3.1 Academic Discussion

- Perez-Rosero, D.A.; Álvarez-Meza, A.M.; Castellanos-Dominguez, C.G. A Regularized Physics-Informed Neural Network to Support Data-Driven Nonlinear Constrained Optimization. Computers 2024, 13, 176. https://doi.org/10.3390/computers13070176. Q2-A2 Minciencias

6.3.2 Software and Repositories

- Nonlinear Optimization : Baseline library for nonlinear optimization experiments.

References

[CRE, 2023] , 2023; Creg; https://creg.gov.co/; Último acceso en 2024.

[Cor, 2023] , 2023; Actualidad del sector energético colombiano; https://investigaciones.corficolombiana.com/analisis-sectorial-y-sostenibilidad/perspectiva-sectorial-energia/actualidad-del-sector-energetico-colombiano/informe_1290865; Último acceso en 2024.

[Ene, 2023] , 2023; Cómo se genera la electricidad; https://www.enel.com.co/es/empresas/enel-generacion/como-se-genera-la-electricidad.html; Último acceso en 2024.

[REA, 2023] , 2023; Renewable energy association; https://www.r-e-a.net; Último acceso en 2024.

[Uni, 2023] , 2023; Colombia puede liderar la transición energética que exige el planeta; https://uniandes.edu.co/es/noticias/ingenieria/colombia-puede-liderar-la-transicion-energetica-que-exige-el-planeta; Último acceso en 2024.

[FIC, 2024] , 2024; Xpress optimization; https://www.fico.com/en/products/fico-xpress-optimization; Último acceso en 2024.

[Gur, 2024] , 2024; Gurobi optimization; https://www.gurobi.com/; Último acceso en 2024.

[Ipo, 2024] , 2024; Ipopt deprecated features; https://coin-or.github.io/Ipopt/deprecated.html; Último acceso en 2024.

[Mos, 2024] , 2024; Mosek; https://www.mosek.com/; Último acceso en 2024.

[*Ab Wahab et al.*, 2020] **Ab Wahab, M. N.**; **Nefti-Meziani, S.** & **Atyabi, A.**: , 2020; A comparative review on mobile robot path planning: Classical or meta-heuristic methods?; *Annual Reviews in Control*; **50**: 233–252.

[*Abubakar & Kumam*, 2019] **Abubakar, A. B.** & **Kumam, P.**: , 2019; A descent dai-liao conjugate gradient method for nonlinear equations; *Numerical Algorithms*; **81**: 197–210.

[*Agrawal & Boyd*, 2020] **Agrawal, A.** & **Boyd, S.**: , 2020; Disciplined quasiconvex programming; *Optimization Letters*; to appear.

[*Agrawal et al.*, 2019] **Agrawal, A.**; **Amos, B.**; **Barratt, S.**; **Boyd, S.**; **Diamond, S.** & **Kolter, J. Z.**: , 2019; Differentiable convex optimization layers; *Advances in neural information processing systems*; **32**.

[Andrei *et al.*, 2020] **Andrei, N.** *et al.*: , 2020; *Nonlinear conjugate gradient methods for unconstrained optimization*; Springer.

[*ANH*, 2022] **ANH**: , 2022; Informe de reservas y recursos contingentes de hidrocarburos 2022; `https://www.anh.gov.co/documents/21617/Informe_de_Reservas__y_Recursos_Contingentes_de_Hidrocarburos_2022_pfMyhzQ.pdf`; Último acceso en 2024.

[*Applegate et al.*, 2021] **Applegate, D.**; **Diaz, M.**; **Hinder, O.**; **Lu, H.**; **Lubin, M.**; **O' Donoghue, B.** & **Schudy, W.**: , 2021; Practical large-scale linear programming using primal-dual hybrid gradient; en *Advances in Neural Information Processing Systems*, tomo 34 (Editado por **Ranzato, M.**; **Beygelzimer, A.**; **Dauphin, Y.**; **Liang, P.** & **Vaughan, J. W.**); Curran Associates, Inc.; págs. 20243–20257; URL `https://proceedings.neurips.cc/paper_files/paper/2021/file/a8fbbd3b11424ce032ba813493d95ad7-Paper.pdf`.

[*Arya*, 2022] **Arya, A. K.**: , 2022; A critical review on optimization parameters and techniques for gas pipeline operation profitability; *Journal of Petroleum Exploration and Production Technology*; **12** (11): 3033–3057.

[*Arya et al.*, 2022] **Arya, A. K.**; **Jain, R.**; **Yadav, S.**; **Bisht, S.** & **Gautam, S.**: , 2022; Recent trends in gas pipeline optimization; *Materials Today: Proceedings*; **57**: 1455–1461.

[*Asgharieh Ahari & Kocuk*, 2023] **Asgharieh Ahari, S.** & **Kocuk, B.**: , 2023; A mixed-integer exponential cone programming formulation for feature subset selection in logistic regression; *EURO Journal on Computational Optimization*; **11**: 100069; doi: https://doi.org/10.1016/j.ejco.2023.100069; URL `https://www.sciencedirect.com/science/article/pii/S2192440623000138`.

[*Attari & Torkayesh*, 2018] **Attari, M. Y. N.** & **Torkayesh, A. E.**: , 2018; Developing benders decomposition algorithm for a green supply chain network of mine industry: Case of iranian mine industry; *Operations Research Perspectives*; **5**: 371–382.

[*Awwal et al.*, 2019] **Awwal, A. M.**; **Kumam, P.** & **Abubakar, A. B.**: , 2019; A modified conjugate gradient method for monotone nonlinear equations with convex constraints; *Applied Numerical Mathematics*; **145**: 507–520.

[*Azad et al.*, 2020] **Azad, A. S.**; **Rahaman, M. S. A.**; **Watada, J.**; **Vasant, P.** & **Vintaned, J. A. G.**: , 2020; Optimization of the hydropower energy generation using meta-heuristic approaches: A review; *Energy Reports*; **6**: 2230–2248.

[*Bahrami & Mohammadi*, 2019] **Bahrami, S.** & **Mohammadi, A.**: , 2019; *Smart microgrids*; Springer.

[*Baker*, 2020] **Baker, K.**: , 2020; A learning-boosted quasi-newton method for ac optimal power flow; *arXiv preprint arXiv:2007.06074*.

[*Bayat & Bagheri*, 2019] **Bayat, A.** & **Bagheri, A.**: , 2019; Optimal active and reactive power allocation in distribution networks using a novel heuristic approach; *Applied Energy*; **233**: 71–85.

[*Baydin et al.*, 2018] **Baydin, A. G.**; **Pearlmutter, B. A.**; **Radul, A. A.** & **Siskind, J. M.**: , 2018; Automatic differentiation in machine learning: a survey; *Journal of Marchine Learning Research*; **18**: 1–43.

[*Beal et al.*, 2018] **Beal, L.**; **Hill, D.**; **Martin, R.** & **Hedengren, J.**: , 2018; Gekko optimization suite; *Processes*; **6** (8): 106; doi:10.3390/pr6080106.

[*Beyza et al.*, 2020] **Beyza, J.**; **Bravo, V. M.**; **Garcia-Paricio, E.**; **Yusta, J. M.** & **Artal-Sevil, J. S.**: , 2020; Vulnerability and resilience assessment of power systems: From deterioration to recovery via a topological model based on graph theory; en *2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, tomo 4; págs. 1–6; doi:10.1109/ROPEC50909.2020.9258709.

[*Bolte & Pauwels*, 2021] **Bolte, J.** & **Pauwels, E.**: , 2021; Conservative set valued fields, automatic differentiation, stochastic gradient methods and deep learning; *Mathematical Programming*; **188**: 19–51.

[*Böttcher et al.*, 2023] **Böttcher, L.**; **Wolf, H.**; **Jung, B.**; **Lutat, P.**; **Trageser, M.**; **Pohl, O.**; **Tao, X.**; **Ulbig, A.** & **Grohe, M.**: , 2023; Solving ac power flow with graph neural networks under realistic constraints; en *2023 IEEE Belgrade PowerTech*; IEEE; págs. 1–7.

[*Buitrago-Villada et al.*, 2021] **Buitrago-Villada, M. d. P.**; **García-Marín, S.**; **Zuluaga-Orozco, J. E.** & **Murillo-Sánchez, C. E.**: , 2021; On the importance of using an ac or dc network model in the multi-period secure stochastic optimal power flow for settling a multidimensional day-ahead market; *IEEE Latin America Transactions*; **19** (12): 2003–2010; doi:10.1109/TLA.2021.9480141.

[*Casacio et al.*, 2019] **Casacio, L.**; **Lyra, C.** & **Oliveira, A. R.**: , 2019; Interior point methods for power flow optimization with security constraints; *International Transactions in Operational Research*; **26** (1): 364–378.

[*CEOWORLD*, 2022] **CEOWORLD**: , 2022; The power of optimization; `https://ceoworld.biz/2022/12/05/the-power-of-optimization/`; Último acceso en 2024.

[*Chang et al.*, 2022] **Chang, H.**; **Chen, Q.**; **Lin, R.**; **Shi, Y.**; **Xie, L.** & **Su, H.**: , 2022; Controlling pressure of gas pipeline network based on mixed proximal policy optimization; en *2022 China Automation Congress (CAC)*; IEEE; págs. 4642–4647.

[*Chen et al.*, 2020a] **Chen, C.**; **Liu, S.**; **Lin, Z.**; **Yang, L.**; **Wu, X.**; **Qiu, W.**; **Gao, Q.**; **Zhu, T.**; **Xing, F.** & **Zhang, J.**: , 2020a; Optimal coordinative operation strategy of the electric–thermal–gas integrated energy system considering csp plant; *IET Energy Systems Integration*; **2** (3): 187–195.

[*Chen et al.*, 2021] **Chen, J.**; **Wang, L.**; **Wang, C.**; **Yao, B.**; **Tian, Y.** & **Wu, Y.-S.**: , 2021; Automatic fracture optimization for shale gas reservoirs based on gradient descent method and reservoir simulation; *Advances in Geo-Energy Research*; **5** (2): 191–201.

[*Chen et al.*, 2022] **Chen, M.**; **Lu, H.** & **Zheng, C.**: , 2022; An integrated energy system optimization model coupled with power-to-gas and carbon capture; en *2022 International Conference on Renewable Energies and Smart Technologies (REST)*, tomo 1; IEEE; págs. 1–5.

[*Chen et al.*, 2020b] **Chen, Q.**; **Zuo, L.**; **Wu, C.**; **Bu, Y.**; **Lu, Y.**; **Huang, Y.** & **Chen, F.**: , 2020b; Short-term supply reliability assessment of a gas pipeline system under demand variations; *Reliability Engineering & System Safety*; **202**: 107004.

[*Chen et al.*, 2023] **Chen, T.-C.**; **Alvarez, J. R. N.**; **Dwijendra, N. K. A.**; **Kadhim, Z. J.**; **Alayi, R.**; **Kumar, R.**; **PraveenKumar, S.** & **Velkin, V. I.**: , 2023; Modeling and optimization of combined heating, power, and gas production system based on renewable energies; *Sustainability*; **15** (10): 7888.

[*Chevalier & Chatzivasileiadis*, 2023] **Chevalier, S.** & **Chatzivasileiadis, S.**: , 2023; Global performance guarantees for neural network models of ac power flow.

[*Chowdhury & Kamalasadan*, 2021] **Chowdhury, M. M.-U.-T.** & **Kamalasadan, S.**: , 2021; A new second-order cone programming model for voltage control of power distribution system with inverter-based distributed generation; *IEEE Transactions on Industry Applications*; **57** (6): 6559–6567.

[*Cotter et al.*, 2019] **Cotter, A.**; **Jiang, H.**; **Gupta, M.**; **Wang, S.**; **Narayan, T.**; **You, S.** & **Sridharan, K.**: , 2019; Optimization with non-differentiable constraints with applications to fairness, recall, churn, and other goals; *Journal of Machine Learning Research*; **20** (172): 1–59.

[*CREG*, 2022a] **CREG**: , 2022a; Anexo 7; `https://gestornormativo.creg.gov.co/Publicac.nsf/52188526a7290f8505256eee0072eba7/efe89a093559fe0b052589b9007d308b/$FILE/Anexo7.pdf`; Último acceso en 2024.

[*CREG*, 2022b] **CREG**: , 2022b; Regulación de la creg sobre energías renovables comprende seis grandes temas; `https://creg.gov.co/publicaciones/8625/`

regulacion-de-la-creg-sobre-energias-renovables-comprende-seis-grandes-temas/; Último acceso en 2024.

[*CREG*, 2022c] **CREG**: , 2022c; Reglamento Único de la comisión de regulación de energía y gas (creg); https://gestornormativo.creg.gov.co/gestor/entorno/docs/ru_creg_gn.htm; Último acceso en 2024.

[*Delgado et al.*, 2022] **Delgado, J. A.**; **Baptista, E. C.**; **Balbo, A. R.**; **Soler, E. M.**; **Silva, D. N.**; **Martins, A. C.** & **Nepomuceno, L.**: , 2022; A primal–dual penalty-interior-point method for solving the reactive optimal power flow problem with discrete control variables; *International Journal of Electrical Power & Energy Systems*; **138**: 107917.

[*Diamond & Boyd*, 2016] **Diamond, S.** & **Boyd, S.**: , 2016; CVXPY: A Python-embedded modeling language for convex optimization; *Journal of Machine Learning Research*; **17** (83): 1–5.

[*Diehl*, 2019] **Diehl, F.**: , 2019; Warm-starting ac optimal power flow with graph neural networks; en *33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*; págs. 1–6.

[*EIA*, 2022] **EIA**: , 2022; Natural gas explained; https://www.eia.gov/energyexplained/natural-gas/; Último acceso en 2024.

[*Eleftheriadis et al.*, 2023] **Eleftheriadis, P.**; **Leva, S.** & **Ogliari, E.**: , 2023; Bayesian hyperparameter optimization of stacked bidirectional long short-term memory neural network for the state of charge estimation; *Sustainable Energy, Grids and Networks*; **36**: 101160.

[*Energy Master*, 2023] **Energy Master**: , 2023; Consumo de energía; https://energymaster.co/consumo-de-energia-2/; Último acceso en 2024.

[*Energy5*, 2024] **Energy5**: , 2024; Optimización de las redes de gasoductos de gas natural para mejorar la gestión de la cadena de suministro; https://energy5.com/es/optimizacion-de-las-redes-de-gasoductos-de-gas-natural-para-mejorar-la-gestion-de- Último acceso en 2024.

[*Falconer & Mones*, 2022] **Falconer, T.** & **Mones, L.**: , 2022; Leveraging power grid topology in machine learning assisted optimal power flow; *IEEE Transactions on Power Systems*; **38** (3): 2234–2246.

[*Falconer & Mones*, 2023] **Falconer, T.** & **Mones, L.**: , 2023; Leveraging power grid topology in machine learning assisted optimal power flow; *IEEE Transactions on Power Systems*; **38** (3): 2234–2246; doi:10.1109/TPWRS.2022.3187218.

[*Fallahi & Maghouli*, 2020] **Fallahi, F.** & **Maghouli, P.**: , 2020; Integrated unit commitment and natural gas network operational planning under renewable generation uncertainty; *International Journal of Electrical Power & Energy Systems*; **117**: 105647.

[*Falsone et al.*, 2020] **Falsone, A.**; **Notarnicola, I.**; **Notarstefano, G.** & **Prandini, M.**: , 2020; Tracking-admm for distributed constraint-coupled optimization; *Automatica*; **117**: 108962.

[*Fan et al.*, 2023] **Fan, Y.**; **Chen, Y.** & **Cui, M.**: , 2023; Electric-hydrogen integrated energy system optimization considering ladder-type carbon trading mechanism and user-side flexible load; en *2023 IEEE 6th International Electrical and Energy Conference (CIEEC)*; IEEE; págs. 1780–1784.

[*Farrokhifar et al.*, 2020] **Farrokhifar, M.**; **Nie, Y.** & **Pozo, D.**: , 2020; Energy systems planning: A survey on models for integrated power and natural gas networks coordination; *Applied Energy*; **262**: 114567.

[*Füllner & Rebennack*, 2022] **Füllner, C.** & **Rebennack, S.**: , 2022; Non-convex nested benders decomposition; *Mathematical Programming*; **196** (1-2): 987–1024.

[*Gao et al.*, 2022] **Gao, G.**; **Wang, Y.**; **Vink, J. C.**; **Wells, T. J.** & **Saaf, F. J.**: , 2022; Distributed quasi-newton derivative-free optimization method for optimization problems with multiple local optima; *Computational Geosciences*; **26** (4): 847–863.

[*Gao & Li*, 2021] **Gao, H.** & **Li, Z.**: , 2021; A benders decomposition based algorithm for steady-state dispatch problem in an integrated electricity-gas system; *IEEE Transactions on Power Systems*; **36** (4): 3817–3820.

[*Gao et al.*, 2019] **Gao, S.**; **Yu, Y.**; **Wang, Y.**; **Wang, J.**; **Cheng, J.** & **Zhou, M.**: , 2019; Chaotic local search-based differential evolution algorithms for optimization; *IEEE Transactions on Systems, Man, and Cybernetics: Systems*; **51** (6): 3954–3967.

[*García-Marín et al.*, 2019] **García-Marín, S.**; **González-Vanegas, W.** & **Murillo-Sánchez, C.**: , 2019; MPNG: MATPOWER-Natural Gas; `https://github.com/MATPOWER/mpng`; [Online; accessed (fecha de acceso)].

[*García-Marín et al.*, 2022] **García-Marín, S.**; **González-Vanegas, W.** & **Murillo-Sánchez, C.**: , 2022; Mpng: A matpower-based tool for optimal power and natural gas flow analyses; *IEEE Transactions on Power Systems*: 1–9; doi:10.1109/TPWRS.2022.3195684.

[*Gholizadeh et al.*, 2020] **Gholizadeh, S.**; **Danesh, M.** & **Gheyratmand, C.**: , 2020; A new newton metaheuristic algorithm for discrete performance-based design optimization of steel moment frames; *Computers & Structures*; **234**: 106250.

[*González et al.*, 2021] **González, A. B. P.**; **Silva, B. D. J.** & **Macia, Y. M.**: , 2021; Transición energética en américa latina y el caribe: diálogos inter y transdisciplinarios en tiempos de pandemia por covid-19; *Revista LIDER*: 33–61.

[*Gonzalez & Peña-Vinces*, 2023] **Gonzalez, C. C.** & **Peña-Vinces, J.**: , 2023; A framework for a green accounting system-exploratory study in a developing country context, colombia; *Environment, Development and Sustainability*; **25** (9): 9517–9541.

[*González-Vanegas et al.*, 2019] **González-Vanegas, W.**; **Álvarez Meza, A.**; **Hernández-Muriel, J.** & **Orozco-Gutiérrez,** : , 2019; Akl-abc: An automatic approximate bayesian computation approach based on kernel learning; *Entropy*; **21** (10); doi:10.3390/e21100932; URL `https://www.mdpi.com/1099-4300/21/10/932`.

[*Goulart & Chen*, 2024] **Goulart, P.** & **Chen, Y.**: , 2024; Clarabel documentation; `https://oxfordcontrol.github.io/ClarabelDocs/stable/`; Último acceso en 2024.

[*Greyson*, 2021] **Greyson, K. A.**: , 2021; Vehicles power consumption measurements: Case study of dart in tanzania; en *2021 IEEE AFRICON*; págs. 1–7; doi:10.1109/AFRICON51333.2021.9570955.

[*Habib & Yildirim*, 2022] **Habib, A.** & **Yildirim, U.**: , 2022; Developing a physics-informed and physics-penalized neural network model for preliminary design of multi-stage friction pendulum bearings; *Engineering Applications of Artificial Intelligence*; **113**: 104953.

[*Haghighat et al.*, 2021] **Haghighat, E.**; **Raissi, M.**; **Moure, A.**; **Gomez, H.** & **Juanes, R.**: , 2021; A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics; *Computer Methods in Applied Mechanics and Engineering*; **379**: 113741; doi:https://doi.org/10.1016/j.cma.2021.113741; URL `https://www.sciencedirect.com/science/article/pii/S0045782521000773`.

[*Haji & Abdulazeez*, 2021] **Haji, S. H.** & **Abdulazeez, A. M.**: , 2021; Comparison of optimization techniques based on gradient descent algorithm: A review; *PalArch's Journal of Archaeology of Egypt/Egyptology*; **18** (4): 2715–2743.

[*Han et al.*, 2024] **Han, M.**; **Du, Z.**; **Yuen, K. F.**; **Zhu, H.**; **Li, Y.** & **Yuan, Q.**: , 2024; Walrus optimizer: A novel nature-inspired metaheuristic algorithm; *Expert Systems with Applications*; **239**: 122413.

[*Harrington*, 2018] **Harrington, K. R.**: , 2018; Panorama actual sobre eficiencia energética en américa latina; *Revista VIRTUALPRO*; **200**.

[*Hasanzadeh et al.*, 2021] **Hasanzadeh, A.**; **Chitsaz, A.**; **Mojaver, P.** & **Ghasemi, A.**: , 2021; Stand-alone gas turbine and hybrid mcfc and sofc-gas turbine systems: Comparative life cycle cost, environmental, and energy assessments; *Energy Reports*; **7**: 4659–4680.

[*Heipcke & Colombani*, 2020] **Heipcke, S.** & **Colombani, Y.**: , 2020; Xpress mosel: Modeling and programming features for optimization projects; en *Operations Research Proceedings 2019: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), Dresden, Germany, September 4-6, 2019*; Springer; págs. 677–683.

[*Hu & Jiang*, 2024] **Hu, J.** & **Jiang, Y.**: , 2024; Accelerating benders decomposition approach for shared parking spaces allocation considering parking unpunctuality and no-shows; *Expert Systems with Applications*; **240**: 122346.

[*Hu & Zhang*, 2023] **Hu, Z.** & **Zhang, H.**: , 2023; Optimal power flow based on physical-model-integrated neural network with worth-learning data generation.

[*Huang & Wang*, 2023] **Huang, B.** & **Wang, J.**: , 2023; Applications of physics-informed neural networks in power systems - a review; *IEEE Transactions on Power Systems*; **38** (1): 572–588; doi:10.1109/TPWRS.2022.3162473.

[*Huang et al.*, 2023] **Huang, W.**; **Zhang, X.**; **Cheng, H.**; **Xie, J.** *et al.*: , 2023; Metamodel-based optimization method for traffic network signal design under stochastic demand; *Journal of Advanced Transportation*; **2023**.

[*Ibrahim & Hossain*, 2021] **Ibrahim, I. A.** & **Hossain, M. J.**: , 2021; Low voltage distribution networks modeling and unbalanced (optimal) power flow: A comprehensive review; *IEEE Access*; **9**: 143026–143084.

[*Impacto TIC*, 2023] **Impacto TIC**: , 2023; Energías renovables en colombia: Situación, retos y proyectos; `https://impactotic.co/innovacion/sostenibilidad/energias-renovables-en-colombia-situacion-retos-y-proyectos/`; Último acceso en 2024.

[*International Energy Agency (IEA)*, 2022] **International Energy Agency (IEA)**: , 2022; World energy investment 2022; `https://www.iea.org/reports/world-energy-investment-2022`; license: CC BY 4.0.

[*International Energy Agency (IEA)*, 2023] **International Energy Agency (IEA)**: , 2023; Global energy and climate model; `https://www.iea.org/reports/global-energy-and-climate-model`; license: CC BY 4.0.

[*Jakovetić et al.*, 2020] **Jakovetić, D.**; **Bajović, D.**; **Xavier, J.** & **Moura, J. M.**: , 2020; Primal–dual methods for large-scale and distributed convex optimization and data analytics; *Proceedings of the IEEE*; **108** (11): 1923–1938.

[*Jamii et al.*, 2022] **Jamii, J.**; **Trabelsi, M.**; **Mansouri, M.**; **Mimouni, M. F.** & **Shatanawi, W.**: , 2022; Non-linear programming-based energy management for a wind farm coupled with pumped hydro storage system; *Sustainability*; **14** (18); doi:10.3390/su141811287; URL `https://www.mdpi.com/2071-1050/14/18/11287`.

[*Jeon & Van Roy*, 2022] **Jeon, H. J.** & **Van Roy, B.**: , 2022; An information-theoretic framework for deep learning; *Advances in Neural Information Processing Systems*; **35**: 3279–3291.

[*Kardoš et al.*, 2022] **Kardoš, J.**; **Kourounis, D.**; **Schenk, O.** & **Zimmerman, R.**: , 2022; Beltistos: A robust interior point method for large-scale optimal power flow problems; *Electric Power Systems Research*; **212**: 108613.

[*Karimi et al.*, 2019] **Karimi, M.**; **Shahriari, A.**; **Aghamohammadi, M.**; **Marzooghi, H.** & **Terzija, V.**: , 2019; Application of newton-based load flow methods for determining steady-state condition of well and ill-conditioned power systems: A review; *International Journal of Electrical Power & Energy Systems*; **113**: 298–309.

[*Khan et al.*, 2020] **Khan, I. U.**; **Javaid, N.**; **Gamage, K. A.**; **Taylor, C. J.**; **Baig, S.** & **Ma, X.**: , 2020; Heuristic algorithm based optimal power flow model incorporating stochastic renewable energy sources; *IEEE Access*; **8**: 148622–148643.

[*KILIÇARSLAN et al.*, 2021] **KILIÇARSLAN, S.**; **Kemal, A.** & **Çelik, M.**: , 2021; An overview of the activation functions used in deep learning algorithms; *Journal of New Results in Science*; **10** (3): 75–88.

[Kim *et al.*, 2023] **Kim, H.** *et al.*: , 2023; Self-supervised equality embedded deep lagrange dual for approximate constrained optimization; *arXiv preprint arXiv:2306.06674*.

[*Klatzer et al.*, 2022] **Klatzer, T.**; **Bachhiesl, U.** & **Wogrin, S.**: , 2022; State-of-the-art expansion planning of integrated power, natural gas, and hydrogen systems; *International Journal of Hydrogen Energy*; **47** (47): 20585–20603.

[*Kohjitani et al.*, 2022] **Kohjitani, H.**; **Koda, S.**; **Himeno, Y.**; **Makiyama, T.**; **Yamamoto, Y.**; **Yoshinaga, D.**; **Wuriyanghai, Y.**; **Kashiwa, A.**; **Toyoda, F.**; **Zhang, Y.** *et al.*: , 2022; Gradient-based parameter optimization method to determine membrane ionic current composition in human induced pluripotent stem cell-derived cardiomyocytes; *Scientific Reports*; **12** (1): 19110.

[*Koksal & Aydin*, 2023] **Koksal, E. S.** & **Aydin, E.**: , 2023; Physics informed piecewise linear neural networks for process optimization; *Computers Chemical Engineering*; **174**: 108244; doi:https://doi.org/10.1016/j.compchemeng.2023.108244; URL `https://www.sciencedirect.com/science/article/pii/S009813542300114X`.

[*Kuhns & Shaw*, 2018] **Kuhns, R. J.** & **Shaw, G. H.**: , 2018; *Coal and Natural Gas*; Springer International Publishing, Cham; ISBN 978-3-319-22783-2; págs. 65–69; doi:10.1007/978-3-319-22783-2_8; URL `https://doi.org/10.1007/978-3-319-22783-2_8`.

[*Kumar & Rahaman*, 2020] **Kumar, J.** & **Rahaman, O.**: , 2020; Lower bound limit analysis using power cone programming for solving stability problems in rock mechanics for generalized hoek–brown criterion; *Rock Mechanics and Rock Engineering*; **53** (7): 3237–3252.

[*Laue*, 2022] **Laue, S.**: , 2022; On the equivalence of automatic and symbolic differentiation.

[*Lee et al.*, 2023] **Lee, M.**; **Kim, K. T.** & **Park, J.**: , 2023; A numerically efficient output-only system-identification framework for stochastically forced self-sustained oscillators; *arXiv preprint arXiv:2305.02801*.

[*Leon et al.*, 2020] **Leon, L. M.**; **Bretas, A. S.** & **Rivera, S.**: , 2020; Quadratically constrained quadratic programming formulation of contingency constrained optimal power flow with photovoltaic generation; *Energies*; **13** (13): 3310.

[*Li et al.*, 2022] **Li, S.**; **Ding, T.**; **Jia, W.**; **Huang, C.**; **Catalão, J. P. S.** & **Li, F.**: , 2022; A machine learning-based vulnerability analysis for cascading failures of integrated

power-gas systems; *IEEE Transactions on Power Systems*; **37** (3): 2259–2270; doi: 10.1109/TPWRS.2021.3119237.

[*Li et al.*, 2017] **Li, X.**; **Tomasgard, A.** & **Barton, P. I.**: , 2017; Natural gas production network infrastructure development under uncertainty; *Optimization and Engineering*; **18**: 35–62.

[*Liang & Zhao*, 2023] **Liang, H.** & **Zhao, C.**: , 2023; Deepopf-u: A unified deep neural network to solve ac optimal power flow in multiple networks.

[*Lin et al.*, 2022a] **Lin, Y.**; **Zhang, X.**; **Wang, J.**; **Shi, D.** & **Bian, D.**: , 2022a; Voltage stability constrained optimal power flow for unbalanced distribution system based on semidefinite programming; *Journal of Modern Power Systems and Clean Energy*; **10** (6): 1614–1624; doi:10.35833/MPCE.2021.000220.

[*Lin et al.*, 2022b] **Lin, Y.**; **Zhang, X.**; **Wang, J.**; **Shi, D.** & **Bian, D.**: , 2022b; Voltage stability constrained optimal power flow for unbalanced distribution system based on semidefinite programming; *Journal of Modern Power Systems and Clean Energy*; **10** (6): 1614–1624.

[*Liu et al.*, 2021] **Liu, B.**; **Yang, Q.**; **Zhang, H.** & **Wu, H.**: , 2021; An interior-point solver for ac optimal power flow considering variable impedance-based facts devices; *IEEE Access*; **9**: 154460–154470.

[*Lopez-Garcia & Domínguez-Navarro*, 2023] **Lopez-Garcia, T. B.** & **Domínguez-Navarro, J. A.**: , 2023; Power flow analysis via typed graph neural networks; *Engineering Applications of Artificial Intelligence*; **117**: 105567.

[*Ma et al.*, 2020a] **Ma, M.**; **Fan, L.**; **Miao, Z.**; **Zeng, B.** & **Ghassempour, H.**: , 2020a; A sparse convex ac opf solver and convex iteration implementation based on 3-node cycles; *Electric Power Systems Research*; **180**: 106169; doi:https://doi.org/10.1016/j.epsr.2019.106169; URL `https://www.sciencedirect.com/science/article/pii/S0378779619304882`.

[*Ma et al.*, 2020b] **Ma, X.**; **Huang, H.**; **Wang, Y.**; **Romano, S.**; **Erfani, S.** & **Bailey, J.**: , 2020b; Normalized loss functions for deep learning with noisy labels; en *International conference on machine learning*; PMLR; págs. 6543–6553.

[*Mahapatra & Rajan*, 2020] **Mahapatra, D.** & **Rajan, V.**: , 2020; Multi-task learning with user preferences: Gradient descent with controlled ascent in pareto optimization; en *International Conference on Machine Learning*; PMLR; págs. 6597–6607.

[*Mai & Mortari*, 2022] **Mai, T.** & **Mortari, D.**: , 2022; Theory of functional connections applied to quadratic and nonlinear programming under equality constraints; *Journal of Computational and Applied Mathematics*; **406**: 113912.

[*Mannel & Rund*, 2021] **Mannel, F.** & **Rund, A.**: , 2021; A hybrid semismooth quasi-newton method for nonsmooth optimal control with pdes; *Optimization and Engineering*; **22** (4): 2087–2125.

[*McKinsey & Company*, 2023] **McKinsey & Company**: , 2023; Global energy perspective 2023; `https://www.mckinsey.com/industries/oil-and-gas/our-insights/global-energy-perspective-2023`; accedido el [fecha de acceso].

[*Meng et al.*, 2020a] **Meng, L.**; **Zhang, C.**; **Ren, Y.**; **Zhang, B.** & **Lv, C.**: , 2020a; Mixed-integer linear programming and constraint programming formulations for solving distributed flexible job shop scheduling problem; *Computers & industrial engineering*; **142**: 106347.

[*Meng et al.*, 2020b] **Meng, Z.**; **Zhao, Z.**; **Mei, D.** & **Zhou, Y.**: , 2020b; Numerical differentiation for two-dimensional functions by a fourier extension method; *Inverse Problems in Science and Engineering*; **28** (1): 126–143.

[*Mhanna & Mancarella*, 2021] **Mhanna, S.** & **Mancarella, P.**: , 2021; An exact sequential linear programming algorithm for the optimal power flow problem; *IEEE Transactions on Power Systems*; **37** (1): 666–679.

[*Ministerio de Minas y Energía*, 2020] **Ministerio de Minas y Energía**: , 2020; Resolución 40311 de 2020; `https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_minminas_40311_2020.htm`; Último acceso en 2024.

[*Ministerio de Minas y Energía*, 2021] **Ministerio de Minas y Energía**: , 2021; Resolución 40279 de 2021; `https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_minminas_40279_2021.htm`; Último acceso en 2024.

[*Misra*, 1908] **Misra, D.**: , 1908; Mish: A self regularized non-monotonic activation function. arxiv 2019; *arXiv preprint arXiv:1908.08681*.

[*Misyris et al.*, 2020] **Misyris, G. S.**; **Venzke, A.** & **Chatzivasileiadis, S.**: , 2020; Physics-informed neural networks for power systems; en *2020 IEEE power & energy society general meeting (PESGM)*; IEEE; págs. 1–5.

[*Misyris et al.*, 2021] **Misyris, G. S.**; **Stiasny, J.** & **Chatzivasileiadis, S.**: , 2021; Capturing power system dynamics by physics-informed neural networks and optimization; en *2021 60th IEEE Conference on Decision and Control (CDC)*; IEEE; págs. 4418–4423.

[*Moehle et al.*, 2017] **Moehle, N.**; **Mattingley, J.** & **Boyd, S.**: , 2017; Embedded convex optimization with cvxpy; *url= https://github. com/moehle/cvxpy codegen*.

[*Mokhtari & Ribeiro*, 2020] **Mokhtari, A.** & **Ribeiro, A.**: , 2020; Stochastic quasi-newton methods; *Proceedings of the IEEE*; **108** (11): 1906–1922.

[*Montoya et al.*, 2020a] **Montoya, O.**; **Gil-González, W.**; **Hernández, J. C.**; **Giral-Ramírez, D. A.** & **Medina-Quesada, A.**: , 2020a; A mixed-integer nonlinear programming model for optimal reconfiguration of dc distribution feeders; *Energies*; **13** (17): 4440.

[*Montoya et al.*, 2020b] **Montoya, O. D.**; **Gil-González, W.** & **Grisales-Noreña, L.**: , 2020b; Relaxed convex model for optimal location and sizing of dgs in dc grids using sequential quadratic programming and random hyperplane approaches; *International Journal of Electrical Power & Energy Systems*; **115**: 105442.

[*Mora-Camino & Nunes Cosenza*, 2018] **Mora-Camino, F.** & **Nunes Cosenza, C. A.**: , 2018; *Fuzzy Dual Numbers*; Springer International Publishing, Cham; ISBN 978-3-319-65418-8; págs. 11–16; doi:10.1007/978-3-319-65418-8_3; URL `https://doi.org/10.1007/978-3-319-65418-8_3`.

[*Mugel et al.*, 2022] **Mugel, S.**; **Kuchkovsky, C.**; **Sanchez, E.**; **Fernandez-Lorenzo, S.**; **Luis-Hita, J.**; **Lizaso, E.** & **Orus, R.**: , 2022; Dynamic portfolio optimization with real datasets using quantum processors and quantum-inspired tensor networks; *Physical Review Research*; **4** (1): 013006.

[*Muñoz et al.*, 2022] **Muñoz, P.**; **Franceschini, E. A.**; **Levitan, D.**; **Rodriguez, C. R.**; **Humana, T.** & **Perelmuter, G. C.**: , 2022; Comparative analysis of cost, emissions and fuel consumption of diesel, natural gas, electric and hydrogen urban buses; *Energy Conversion and Management*; **257**: 115412.

[*Murphy*, 2022] **Murphy, K. P.**: , 2022; *Probabilistic machine learning: an introduction*; MIT press.

[*Naderi et al.*, 2021] **Naderi, E.**; **Pourakbari-Kasmaei, M.**; **Cerna, F. V.** & **Lehtonen, M.**: , 2021; A novel hybrid self-adaptive heuristic algorithm to handle single-and multi-objective optimal power flow problems; *International Journal of Electrical Power & Energy Systems*; **125**: 106492.

[*Nellikkath & Chatzivasileiadis*, 2022a] **Nellikkath, R.** & **Chatzivasileiadis, S.**: , 2022a; Physics-informed neural networks for ac optimal power flow; *Electric Power Systems Research*; **212**: 108412; doi:https://doi.org/10.1016/j.epsr.2022.108412; URL `https://www.sciencedirect.com/science/article/pii/S0378779622005636`.

[*Nellikkath & Chatzivasileiadis*, 2022b] **Nellikkath, R.** & **Chatzivasileiadis, S.**: , 2022b; Minimizing worst-case violations of neural networks; *arXiv preprint arXiv:2212.10930*.

[*Nellikkath & Chatzivasileiadis*, 2022c] **Nellikkath, R.** & **Chatzivasileiadis, S.**: , 2022c; Physics-informed neural networks for ac optimal power flow; *Electric Power Systems Research*; **212**: 108412.

[*Notarnicola & Notarstefano*, 2019] **Notarnicola, I.** & **Notarstefano, G.**: , 2019; Constraint-coupled distributed optimization: A relaxation and duality approach; *IEEE Transactions on Control of Network Systems*; **7** (1): 483–492.

[*O'Donoghue*, 2021] **O'Donoghue, B.**: , 2021; Operator splitting for a homogeneous embedding of the linear complementarity problem; *SIAM Journal on Optimization*; **31**: 1999–2023.

[*O'Donoghue et al.*, 2016] **O'Donoghue, B.**; **Chu, E.**; **Parikh, N.** & **Boyd, S.**: , 2016; Conic optimization via operator splitting and homogeneous self-dual embedding; *Journal of Optimization Theory and Applications*; **169** (3): 1042–1068; URL `http://stanford.edu/~boyd/papers/scs.html`.

[*Oliva et al.*, 2019] **Oliva, D.**; **Abd Elaziz, M.**; **Elsheikh, A. H.** & **Ewees, A. A.**: , 2019; A review on meta-heuristics methods for estimating parameters of solar cells; *Journal of Power Sources*; **435**: 126683.

[*Owerko et al.*, 2020] **Owerko, D.**; **Gama, F.** & **Ribeiro, A.**: , 2020; Optimal power flow using graph neural networks; en *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; IEEE; págs. 5930–5934.

[*Owerko et al.*, 2022] **Owerko, D.**; **Gama, F.** & **Ribeiro, A.**: , 2022; Unsupervised optimal power flow using graph neural networks; *arXiv preprint arXiv:2210.09277*.

[*Pan et al.*, 2021] **Pan, X.**; **Zhao, T.**; **Chen, M.** & **Zhang, S.**: , 2021; Deepopf: A deep neural network approach for security-constrained dc optimal power flow; *IEEE Transactions on Power Systems*; **36** (3): 1725–1735; doi:10.1109/TPWRS.2020.3026379.

[*Pan et al.*, 2022] **Pan, X.**; **Chen, M.**; **Zhao, T.** & **Low, S. H.**: , 2022; Deepopf: A feasibility-optimized deep neural network approach for ac optimal power flow problems; *IEEE Systems Journal*; **17** (1): 673–683.

[*Pan et al.*, 2023] **Pan, X.**; **Chen, M.**; **Zhao, T.** & **Low, S. H.**: , 2023; Deepopf: A feasibility-optimized deep neural network approach for ac optimal power flow problems; *IEEE Systems Journal*; **17** (1): 673–683; doi:10.1109/JSYST.2022.3201041.

[*Park et al.*, 2023] **Park, S.**; **Chen, W.**; **Mak, T. W. K.** & **Hentenryck, P. V.**: , 2023; Compact optimization learning for ac optimal power flow.

[*Pas et al.*, 2022] **Pas, P.**; **Schuurmans, M.** & **Patrinos, P.**: , 2022; Alpaqa: A matrix-free solver for nonlinear mpc and large-scale nonconvex optimization; en *2022 European Control Conference (ECC)*; págs. 417–422; doi:10.23919/ECC55457.2022.9838172.

[*Paszke et al.*, 2017] **Paszke, A.**; **Gross, S.**; **Chintala, S.**; **Chanan, G.**; **Yang, E.**; **DeVito, Z.**; **Lin, Z.**; **Desmaison, A.**; **Antiga, L.** & **Lerer, A.**: , 2017; Automatic differentiation in pytorch.

[*Pérez-Cedeño et al.*, 2022] **Pérez-Cedeño, R. O.**; **Stanescu, C. L. V.** *et al.*: , 2022; Relación entre el consumo energético y las emisiones de co2 con el índice de desarrollo humano en países americanos: un análisis de eficiencia usando dea; *Publicaciones en Ciencias y Tecnología*; **16** (1): 3–15.

[*Pillutla et al.*, 2023] **Pillutla, K.**; **Roulet, V.**; **Kakade, S. M.** & **Harchaoui, Z.**: , 2023; Modified gauss-newton algorithms under noise; en *2023 IEEE Statistical Signal Processing Workshop (SSP)*; págs. 51–55; doi:10.1109/SSP53291.2023.10207977.

[*Pinheiro et al.*, 2022] **Pinheiro, R. B.**; **Balbo, A. R.**; **Cabana, T. G.** & **Nepomuceno, L.**: , 2022; Solving nonsmooth and discontinuous optimal power flow problems via interior-point p-penalty approach; *Computers & Operations Research*; **138**: 105607.

[*Pritchard et al.*, 1999] **Pritchard, J. K.**; **Seielstad, M. T.**; **Perez-Lezaun, A.** & **Feldman, M. W.**: , 1999; Population growth of human y chromosomes: a study of y chromosome microsatellites.; *Molecular biology and evolution*; **16** (12): 1791–1798.

[*Qi et al.*, 2020] **Qi, J.**; **Du, J.**; **Siniscalchi, S. M.**; **Ma, X.** & **Lee, C.-H.**: , 2020; On mean absolute error for deep neural network based vector-to-vector regression; *IEEE Signal Processing Letters*; **27**: 1485–1489.

[*Qiu et al.*, 2020] **Qiu, B.**; **Huang, Z.**; **Liu, X.**; **Meng, X.**; **You, Y.**; **Liu, G.**; **Yang, K.**; **Maier, A.**; **Ren, Q.** & **Lu, Y.**: , 2020; Noise reduction in optical coherence tomography images using a deep neural network with perceptually-sensitive loss function; *Biomed. Opt. Express*; **11** (2): 817–830; doi:10.1364/BOE.379551; URL `https://opg.optica.org/boe/abstract.cfm?URI=boe-11-2-817`.

[*Rasamoelina et al.*, 2020] **Rasamoelina, A. D.**; **Adjailia, F.** & **Sinčák, P.**: , 2020; A review of activation function for artificial neural network; en *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*; IEEE; págs. 281–286.

[*Raynaud et al.*, 2022] **Raynaud, G.**; **Houde, S.** & **Gosselin, F. P.**: , 2022; Modalpinn: An extension of physics-informed neural networks with enforced truncated fourier decomposition for periodic flow reconstruction using a limited number of imperfect sensors; *Journal of Computational Physics*; **464**: 111271.

[*Ríos-Mercado & Borraz-Sánchez*, 2015] **Ríos-Mercado, R. Z.** & **Borraz-Sánchez, C.**: , 2015; Optimization problems in natural gas transportation systems: A state-of-the-art review; *Applied Energy*; **147**: 536–555.

[*Ritchie & Rosado*, 2020] **Ritchie, H.** & **Rosado, P.**: , 2020; Energy mix; *Our World in Data*; https://ourworldindata.org/energy-mix.

[*Robuschi et al.*, 2021] **Robuschi, N.**; **Zeile, C.**; **Sager, S.** & **Braghin, F.**: , 2021; Multiphase mixed-integer nonlinear optimal control of hybrid electric vehicles; *Automatica*; **123**: 109325.

[*Rumelhart et al.*, 1986] **Rumelhart, D. E.**; **Hinton, G. E.** & **Williams, R. J.**: , 1986; Learning representations by back-propagating errors; *nature*; **323** (6088): 533–536.

[*Sadat & Sahraei-Ardakani*, 2021] **Sadat, S. A.** & **Sahraei-Ardakani, M.**: , 2021; Customized sequential quadratic programming for solving large-scale ac optimal power flow; en *2021 North American Power Symposium (NAPS)*; IEEE; págs. 1–6.

[*Sadat & Sahraei-Ardakani*, 2022] **Sadat, S. A.** & **Sahraei-Ardakani, M.**: , 2022; Tuning successive linear programming to solve ac optimal power flow problem for large networks; *International Journal of Electrical Power & Energy Systems*; **137**: 107807.

[*Sakketou & Ampazis*, 2019] **Sakketou, F.** & **Ampazis, N.**: , 2019; On the invariance of the selu activation function on algorithm and hyperparameter selection in neural network recommenders; en *Artificial Intelligence Applications and Innovations: 15th IFIP WG 12.5 International Conference, AIAI 2019, Hersonissos, Crete, Greece, May 24–26, 2019, Proceedings 15*; Springer; págs. 673–685.

[*Saldarriaga-Cortés et al.*, 2019] **Saldarriaga-Cortés, C.**; **Salazar, H.**; **Moreno, R.** & **Jiménez-Estévez, G.**: , 2019; Stochastic planning of electricity and gas networks: An asynchronous column generation approach; *Applied energy*; **233**: 1065–1077.

[*Sangaiah et al.*, 2020] **Sangaiah, A. K.**; **Tirkolaee, E. B.**; **Goli, A.** & **Dehnavi-Arani, S.**: , 2020; Robust optimization and mixed-integer linear programming model for lng supply chain planning problem; *Soft computing*; **24**: 7885–7905.

[*Schiassi et al.*, 2021] **Schiassi, E.**; **De Florio, M.**; **D'Ambrosio, A.**; **Mortari, D.** & **Furfaro, R.**: , 2021; Physics-informed neural networks and functional interpolation for data-driven parameters discovery of epidemiological compartmental models; *Mathematics*; **9** (17): 2069.

[*Schreider et al.*, 2015] **Schreider, S.**; **Plummer, J.**; **McInnes, D.** & **Miller, B.**: , 2015; Sensitivity analysis of gas supply optimization models; *Annals of Operations Research*; **226**: 565–588.

[*Schuster et al.*, 2020] **Schuster, R.**; **Hanson, J. O.**; **Strimas-Mackey, M.** & **Bennett, J. R.**: , 2020; Exact integer linear programming solvers outperform simulated annealing for solving conservation planning problems; *PeerJ*; **8**: e9258.

[*Shi & Hong*, 2020] **Shi, Q.** & **Hong, M.**: , 2020; Penalty dual decomposition method for nonsmooth nonconvex optimization—part i: Algorithms and convergence analysis; *IEEE Transactions on Signal Processing*; **68**: 4108–4122.

[*Sosa & Navarro*, 2020] **Sosa, P. V.** & **Navarro, D. M.**: , 2020; Crecimiento, complejidad económica y emisiones de co2: un análisis para colombia; *Revista CIFE: Lecturas de Economía Social*; **22** (37): 21–41.

[*Stiasny et al.*, 2021] **Stiasny, J.**; **Chevalier, S.** & **Chatzivasileiadis, S.**: , 2021; Learning without data: Physics-informed neural networks for fast time-domain simulation; en *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*; IEEE; págs. 438–443.

[*Street*, 1988] **Street, J. H.**: , 1988; The contribution of simon s. kuznets to institutionalist development theory; *Journal of Economic Issues*; **22** (2): 499–509.

[*Strelow et al.*, 2023] **Strelow, E. L.**; **Gerisch, A.**; **Lang, J.** & **Pfetsch, M. E.**: , 2023; Physics informed neural networks: A case study for gas transport problems; *Journal of Computational Physics*; **481**: 112041.

[*Sun et al.*, 2022] **Sun, X.**; **Zhang, Y.**; **Zhang, Y.**; **Xie, J.** & **Sun, B.**: , 2022; Operation optimization of integrated energy system considering power-to-gas technology and carbon trading; *International Transactions on Electrical Energy Systems*; **2022**: 1–12.

[*Sun et al.*, 2020] **Sun, Y.**; **Zhang, B.**; **Ge, L.**; **Sidorov, D.**; **Wang, J.** & **Xu, Z.**: , 2020; Day-ahead optimization schedule for gas-electric integrated energy system based on second-order cone programming; *CSEE Journal of Power and Energy Systems*; **6** (1): 142–151.

[*Thangamuthu et al.*, 2022] **Thangamuthu, A.**; **Kumar, G.**; **Bishnoi, S.**; **Bhattoo, R.**; **Krishnan, N.** & **Ranu, S.**: , 2022; Unravelling the performance of physics-informed graph neural networks for dynamical systems; *Advances in Neural Information Processing Systems*; **35**: 3691–3702.

[*Unigas*, 2023] **Unigas**: , 2023; Panorama energético en colombia: retos del segundo semestre de 2023; `https://www.unigas.com.co/blog/panorama-energetico-en-colombia-retos-del-segundo-semestre-de-2023/`; Último acceso en 2024.

[*UPME*, 2019] **UPME**: , 2019; Informe final del plan energético nacional: Indicadores de eficiencia energética; URL `https://www1.upme.gov.co/DemandayEficiencia/Documents/Informe_PEVI_final.pdf`; accessed: 2024-07-15.

[*UPME*, 2021a] **UPME**: , 2021a; Resolución 0283 de 2021; `https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_upme_0283_2021.htm`; Último acceso en 2024.

[*UPME*, 2021b] **UPME**: , 2021b; Resolución 0330 de 2021; `https://normas.cra.gov.co/gestor/docs/resolucion_upme_0330_2021.htm`; Último acceso en 2024.

[*UPME*, 2022a] **UPME**: , 2022a; Proyectos de eficiencia energética; `https://www1.upme.gov.co/DemandayEficiencia/Paginas/Proyectos-de-eficiencia-energetica.aspx`; Último acceso en 2024.

[*UPME*, 2022b] **UPME**: , 2022b; Resolución 0289 de 2022; `https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_upme_0289_2022.htm`; Último acceso en 2024.

[*UPME*, 2022c] **UPME**: , 2022c; Resolución 0468 de 2022; `https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_upme_0468_2022.htm`; Último acceso en 2024.

[*Van de Graaf & Colgan*, 2016] **Van de Graaf, T.** & **Colgan, J.**: , 2016; Global energy governance: a review and research agenda; *Palgrave Communications*; **2** (1): 1–12.

[*Vo et al.*, 2023] **Vo, T. Q. T.**; **Baiou, M.**; **Nguyen, V. H.** & **Weng, P.**: , 2023; Improving subtour elimination constraint generation in branch-and-cut algorithms for the tsp with machine learning; en *Learning and Intelligent Optimization* (Editado

por **Sellmann, M.** & **Tierney, K.**); Springer International Publishing, Cham; ISBN 978-3-031-44505-7; págs. 537–551.

[*Vásquez Cordano & Zellou*, 2020] **Vásquez Cordano, A. L.** & **Zellou, A. M.**: , 2020; Super cycles in natural gas prices and their impact on latin american energy and environmental policies; *Resources Policy*; **65**: 101513; doi:https://doi.org/10.1016/j.resourpol.2019.101513; URL `https://www.sciencedirect.com/science/article/pii/S0301420718302034`.

[*Wang et al.*, 2023] **Wang, G.**; **Zhao, W.**; **Qiu, R.**; **Liao, Q.**; **Lin, Z.**; **Wang, C.** & **Zhang, H.**: , 2023; Operational optimization of large-scale thermal constrained natural gas pipeline networks: A novel iterative decomposition approach; *Energy*; **282**: 128856.

[*Wang et al.*, 2007] **Wang, H.**; **Murillo-Sanchez, C. E.**; **Zimmerman, R. D.** & **Thomas, R. J.**: , 2007; On computational issues of market-based optimal power flow; *IEEE Transactions on Power Systems*; **22** (3): 1185–1193; doi:10.1109/TPWRS.2007.901301.

[*Wang & Zhou*, 2023] **Wang, L.** & **Zhou, B.**: , 2023; Optimal planning of electric vehicle fast-charging stations considering uncertain charging demands via dantzig–wolfe decomposition; *Sustainability*; **15** (8): 6588.

[*Wang et al.*, 2020a] **Wang, Q.**; **Ma, Y.**; **Zhao, K.** & **Tian, Y.**: , 2020a; A comprehensive survey of loss functions in machine learning; *Annals of Data Science*: 1–26.

[*Wang et al.*, 2020b] **Wang, Y.**; **Gao, S.**; **Zhou, M.** & **Yu, Y.**: , 2020b; A multi-layered gravitational search algorithm for function optimization and real-world problems; *IEEE/CAA Journal of Automatica Sinica*; **8** (1): 94–109.

[*Wolpert & Macready*, 1997] **Wolpert, D. H.** & **Macready, W. G.**: , 1997; No free lunch theorems for optimization; *IEEE transactions on evolutionary computation*; **1** (1): 67–82.

[*Wu et al.*, 2018] **Wu, X.**; **Li, C.**; **He, Y.** & **Jia, W.**: , 2018; Operation optimization of natural gas transmission pipelines based on stochastic optimization algorithms: a review; *Mathematical Problems in Engineering*; **2018**.

[*XM*, 2023] **XM**: , 2023; Informes mensuales de análisis del mercado; `https://www.xm.com.co/nuestra-empresa/informes/informes-de-la-operacion-y-el-mercado/informes-mensuales-de-analisis-del-mercado`; Último acceso en 2024.

[*Yang et al.*, 2021] **Yang, L.**; **Meng, X.** & **Karniadakis, G. E.**: , 2021; B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data; *Journal of Computational Physics*; **425**: 109913.

[*Yang et al.*, 2019] **Yang, W.**; **Wang, J.**; **Niu, T.** & **Du, P.**: , 2019; A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting; *Applied energy*; **235**: 1205–1225.

[*Yu et al.*, 2023] **Yu, H.**; **Ma, W.**; **Li, R.**; **Qiu, Z.**; **Li, H.**; **Zhou, J.** & **Wu, K.**: , 2023; Optimization of electric-gas integrated energy system based on reversible solid oxide cell considering gas composition; en *2023 International Conference on Smart Electrical Grid and Renewable Energy (SEGRE)*; IEEE; págs. 214–220.

[*Yu et al.*, 2021] **Yu, W.**; **Huang, W.**; **Wen, Y.**; **Li, Y.**; **Liu, H.**; **Wen, K.**; **Gong, J.** & **Lu, Y.**: , 2021; An integrated gas supply reliability evaluation method of the large-scale and complex natural gas pipeline network based on demand-side analysis; *Reliability Engineering & System Safety*; **212**: 107651.

[*Zhang et al.*, 2023] **Zhang, Z.**; **Sun, M.**; **Deng, R.**; **Kang, C.** & **Chow, M.-Y.**: , 2023; Physics-constrained robustness evaluation of intelligent security assessment for power systems; *IEEE Transactions on Power Systems*; **38** (1): 872–884; doi: 10.1109/TPWRS.2022.3169139.

[*Zhao et al.*, 2020] **Zhao, Z.**; **Liu, S.**; **Zhou, M.** & **Abusorrah, A.**: , 2020; Dual-objective mixed integer linear program and memetic algorithm for an industrial group scheduling problem; *IEEE/CAA Journal of Automatica Sinica*; **8** (6): 1199–1209.

[*Zhou et al.*, 2023] **Zhou, M.**; **Chen, M.** & **Low, S. H.**: , 2023; Deepopf-ft: One deep neural network for multiple ac-opf problems with flexible topology; *IEEE Transactions on Power Systems*; **38** (1): 964–967; doi:10.1109/TPWRS.2022.3217407.

[*Zimmerman & Murillo-Sánchez*, 2020] **Zimmerman, R. D.** & **Murillo-Sánchez, C. E.**: , 2020; *MATPOWER User's Manual*; Zenodo; doi:10.5281/zenodo.4074122.