

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

Cristian Alejandro Blanco Martínez

Universidad Tecnológica de Pereira  
Grupo de investigación Automática

August 20, 2025

# Energy Context & Relevance of Natural Gas

In Colombia, natural gas is widely used across the residential, commercial, industrial, and thermal power sectors. Its importance becomes evident during dry seasons, when hydroelectric generation is reduced and thermal plants, fueled by natural gas, step in to maintain electricity supply [Promigas, 2021].

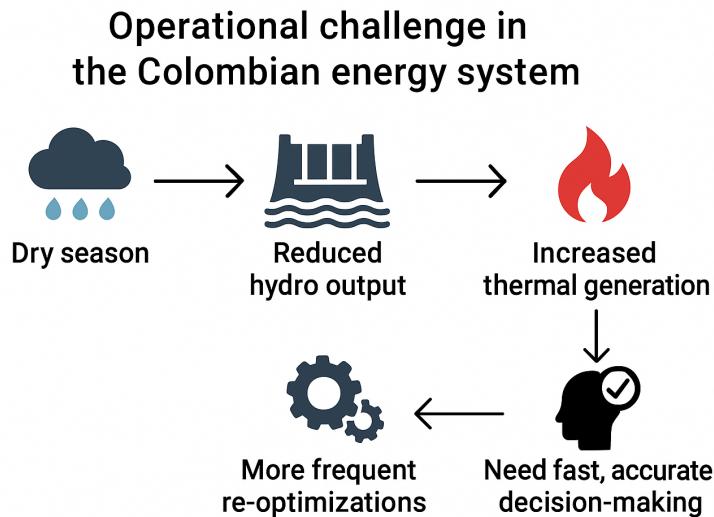


Figure 1: Variation in electricity generation mix between wet and dry seasons in Colombia.

# Motivation & Problem Statement

- **Large-scale natural gas networks** require solving complex optimization problems.
- Increasing network size and operational constraints lead to:
  - Long execution times
  - Difficulty in real-time or near real-time decision making
- Classical approaches are accurate but often **computationally expensive**.

**Main Problem:** There is a need for methods that preserve accuracy while significantly reducing computation time.

**Research Question:** How can an optimization tool be developed that seeks to improve the solution to the problem of natural gas transportation, taking into account computational cost, growing energy demand, and the uncertainty associated with this factor?

# Objectives

## General Objective:

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

## Specific Objectives:

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

# Natural Gas System as a Graph

A natural gas network can be represented as a directed graph, where nodes correspond to elements such as wells, users, or storage facilities, and edges represent pipelines and compressors.

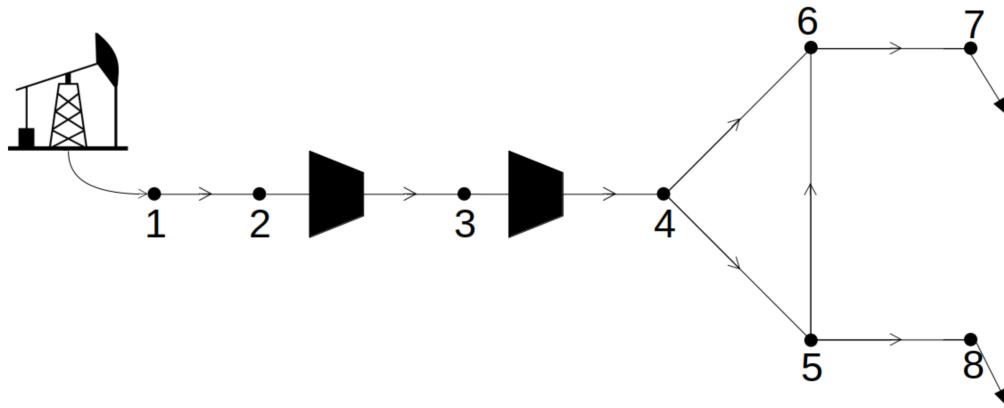


Figure 2: Example of an 8-node gas network represented.

**Objective:** Achieve near-optimizer accuracy with significant runtime reduction.

# CensNet Layers: Integrating Node & Edge Features

**Motivation:** Traditional GCNs focus mainly on *node features* and ignore *edge features*, missing part of the graph's information.

## CensNet innovation:

- Alternates between **node layers** and **edge layers**.
- Each layer type updates its own features while using information from the other:
  - **Node layer:** Node embeddings are updated using both node adjacency and transformed edge features.
  - **Edge layer:** Edge embeddings are updated using edge adjacency and transformed node features.
- Uses an *incidence matrix*  $\mathbf{T}$  to switch between node and edge domains.

## Benefits:

- Captures both *structural* (adjacency) and *relational* (edge) information.
- Enhances long-range dependencies through alternating propagation.

# Mathematical Structure of CensNet Layers

**Node layer propagation:**

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

**Edge layer propagation:**

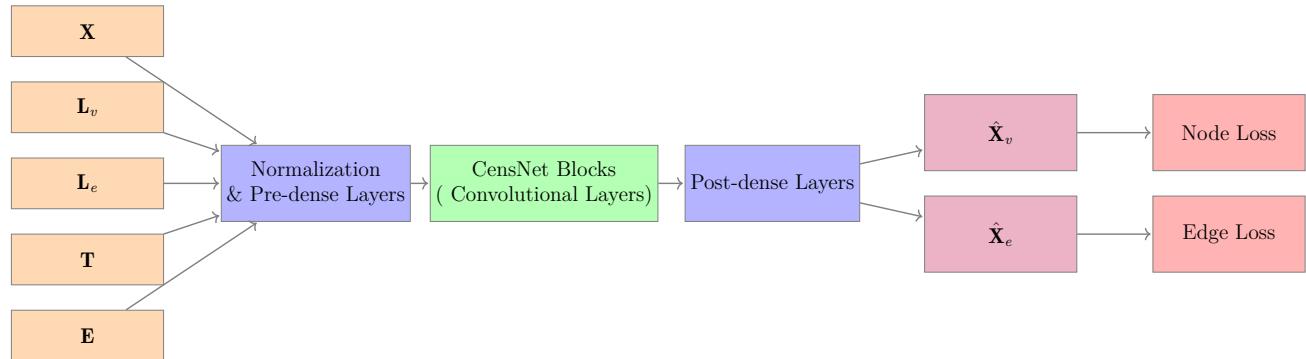
$$\mathbf{H}_e^{(l+1)} = \sigma \left( \mathbf{T}^\top \Phi(\mathbf{H}_v^{(l)} \mathbf{p}_v) \mathbf{T} \odot \tilde{\mathbf{A}}_e \mathbf{H}_e^{(l)} \mathbf{W}_e \right)$$

**Key components:**

- $\tilde{\mathbf{A}}_v, \tilde{\mathbf{A}}_e$ : normalized adjacency matrices (nodes/edges).
- $\mathbf{T}$ : incidence matrix mapping between node and edge domains.
- $\Phi(\cdot)$ : diagonal scaling from projected features.
- $\odot$ : element-wise filtering of adjacency by feature-derived weights.

**Core idea:** Information “switches” between node and edge spaces at each step, enriching representations.

# CensNet-based Model Architecture



**Inputs:** Encapsulate both physical attributes and topological information of the gas network:

- Node features ( $\mathbf{X}$ ): pressures, injections, withdrawals, demand forecasts.
- Node Laplacian ( $\mathbf{L}_v$ ) and Edge Laplacian ( $\mathbf{L}_e$ ): encode structural connectivity and enable topology-aware learning.
- Edge features ( $\mathbf{E}$ ): capacities, compressor ratios.
- Incidence matrix ( $\mathbf{T}$ ): explicit mapping between nodes and edges.

# CensNet-based Model Architecture

- **Processing:**

- Normalization and pre-dense layers transform inputs into a latent space.
- CensNet blocks perform message passing on both node and edge domains, enabling simultaneous learning of flow patterns and interactions.
- Post-dense layers refine features for prediction.

- **Outputs:**

- Node-level flows ( $\hat{\mathbf{X}}_v$ ) — predicted gas supply/demand at each node.
- Edge-level transported flows ( $\hat{\mathbf{X}}_e$ ) — predicted flow rates in each pipeline.

- **Training:** Node and edge predictions are penalized separately using task-specific loss terms, ensuring both accuracy and physical consistency.

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

where  $Y_r$  and  $\hat{Y}_r$  are target and predicted values for task  $r$ , and  $\lambda \|\Theta\|_p$  is a regularization term to prevent overfitting.

# Optimization Problem: Objective and Constraints

The gas network consists of wells  $\mathcal{W}$  (supply nodes), users  $\mathcal{U}$  (demand nodes), pipelines  $\mathcal{P}$ , and compressors  $\mathcal{C}$ .

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

where  $C_i$  is the cost per unit flow at element  $i$ , and  $f_i$  is the corresponding gas flow.

$$\underline{f}_w \leq f_w \leq \overline{f}_w \quad \forall w \in \mathcal{W} \quad \text{Well capacity} \tag{1}$$

$$-\overline{f}_p \leq f_p \leq \overline{f}_p \quad \forall p \in \mathcal{P} \quad \text{Pipeline limits} \tag{2}$$

$$0 \leq f_u \leq \overline{f}_u \quad \forall u \in \mathcal{U} \quad \text{Demand satisfaction} \tag{3}$$

$$\sum_{m: (m, n) \in \mathcal{A}} f_m = \sum_{m': (n, m') \in \mathcal{A}} f_{m'} \quad \text{Nodal gas balance} \tag{4}$$

# Dataset

Two different gas network configurations were used to evaluate the proposed CensNet-based model. The first is a small synthetic case for controlled experimentation, while the second corresponds to the real Colombian natural gas transportation system.

Case	Nodes	Pipelines	Compressors	Users
Small Network	8	6	2	2
Colombian Network	63	48	14	27

**Table 1:** Datasets used in the study.

# Experimental Setup

Different network operation scenarios were generated by adding 5%–25% noise to user demand values and solving each scenario with the linear constrained optimization model using APOPT (via GEKKO). The outputs served as ground truth for CensNet model training.

- **Data split:** 60% training, 20% validation, 20% testing.
- **Samples:** 2000 (8-node network), 2400 (63-node network).
- **Training:** 1500 epochs, Adam optimizer, Leaky ReLU activation ( $\alpha = 0.2$ ).
- **Learning rate:** Initial  $1 \times 10^{-2}$ , exponential decay (rate 0.9, steps 1000).
- **Hyperparameters:**
  - $N_{\text{channels}}$ : 16–64
  - $N_{\text{layers}}$ : 1–5 (CensNet convolutional)
  - $N_{\text{dense}}$ : 2–32 (post-dense layers)
- **Losses tested:**
  - Nodal loss only
  - Nodal + Edge loss

# Main Results: Prediction Accuracy

**Networks:** 8-node (synthetic) and 63-node (Colombian gas system)

**Models:** CensNet (N) – nodal loss only, CensNet (N+E) – nodal + edge loss

- **Nodal flow predictions:** High  $R^2$  in both settings
  - 8-node:  $R^2 = 0.988$  (N), 0.988 (N+E)
  - 63-node:  $R^2 = 0.996$  (N), 0.996 (N+E)
- **Edge flow predictions:**
  - (N) performs poorly on edges:  $R^2$  low
  - (N+E) improves:  $R^2$  up to 0.999 (8-node) and 0.996 (63-node)

**Gas balance consistency (APOPT – CensNet differences):**

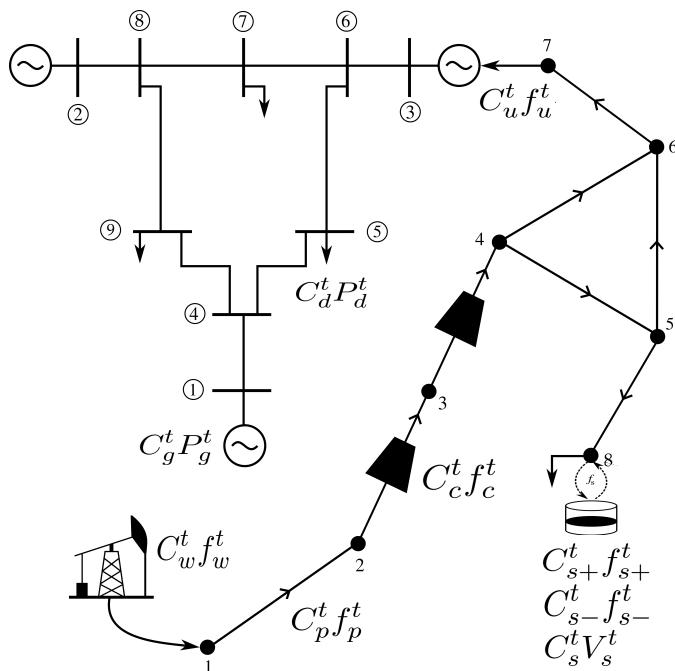
- **8-node:** Edge error reduced from 62.92 (N) → 39.24 (N+E)
- **63-node:** Edge error reduced from 62.47 (N) → -0.07 (N+E)
- Nodal differences < 0.1 units in all cases

**Prediction speed:** T-tests confirm **CensNet is significantly faster** than APOPT

- 8-node:  $T = 14.94$ ,  $p = 1.32 \times 10^{-34}$
- 63-node:  $T = 47.29$ ,  $p = 4.92 \times 10^{-110}$

# Interconnected system - Definition

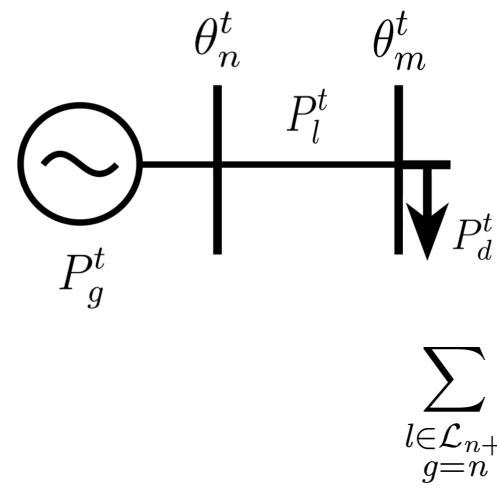
A power-gas interconnected system is a hybrid infrastructure that integrates natural gas and power networks, enhancing overall system efficiency [Duan et al., 2022].



$$\begin{aligned} \min_{\mathcal{P}, \mathcal{F}} \quad & \sum_{g \in \mathcal{G}} C_g^t P_g^t + \sum_{d \in \mathcal{D}} C_d^t P_d^t + \\ & \sum_{w \in \mathcal{W}} C_w^t f_w^t + \sum_{p \in \mathcal{P}} C_p^t f_p^t + \\ & \sum_{c \in \mathcal{C}} C_c^t f_c^t + \sum_{u \in \mathcal{U}} C_u^t f_u^t + \quad (5) \\ & \sum_{s \in \mathcal{S}} C_{s+}^t f_{s+}^t + \sum_{s \in \mathcal{S}} C_{s-}^t f_{s-}^t + \\ & \sum_{s \in \mathcal{S}} C_s^t V_s^t \end{aligned}$$

# Power system constraints

The constraints of the electrical system represent the technical limits of the different elements that compose it, as well as the physical laws that govern it. For certain applications, the DC model is sufficient [Yang et al., 2019].



$$\underline{P}_g^t \leq P_g^t \leq \overline{P}_g^t \quad \forall g \in \mathcal{G}, \quad (6a)$$

$$-\overline{P}_l^t \leq P_l^t \leq \overline{P}_l^t \quad \forall l \in \mathcal{L}, \quad (6b)$$

$$P_l^t = B_{nm}(\theta_n - \theta_m) \quad \forall l = (n, m) \in \mathcal{L}, \quad (6c)$$

$$0 \leq P_d^t \leq \overline{P}_d^t \quad \forall d \in \mathcal{D}, \quad (6d)$$

$$-\overline{\theta}_n^t \leq \theta_m^t \leq \overline{\theta}_n^t \quad \forall n \in \mathcal{N}_P, \quad (6e)$$

$$\sum_{\substack{l \in \mathcal{L}_{n+} \\ g=n}} P_l^t + P_g^t = \sum_{\substack{l \in \mathcal{L}_{n-} \\ d=n}} P_l^t + P_d^t \quad \forall n \in \mathcal{N}_P \quad (6f)$$

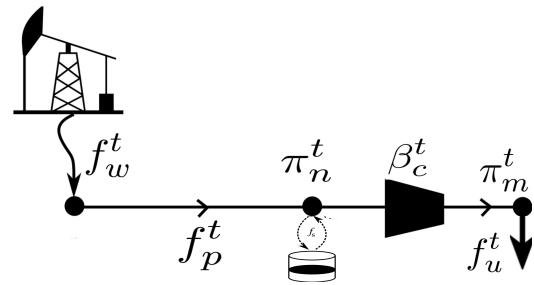
# Interconnection constraints

The second set of restrictions interconnects the natural gas and electric power systems through gas-fired power plants that generate electricity.  $f_n^t$  stands for the natural gas fuel consumption to generate a power  $P_n^t$  at generator bus  $n \in \mathcal{N}_I$ , the heat-rate  $\text{HR}_n$  defines the generator efficiency, and the set  $\mathcal{N}_I = \mathcal{G} \cap \mathcal{U}$  holds all the units in the interconnected system belonging to both the power generator and gas demand sets.

$$f_n^t = P_n^t \cdot \text{HR}_n, \quad \forall n \in \mathcal{N}_I, \quad (7)$$

# Gas system constraints

This set of constraints ensures that wells, pipelines, nodal pressures, compressors, unsupplied demand and storage facilities operate within proper operating limits. [García-Marín et al., 2022].



$$\underline{f}_w^t \leq f_w^t \leq \overline{f}_w^t \quad \forall w \in \mathcal{W} \quad (8a)$$

$$-\overline{f}_p^t \leq f_p^t \leq \overline{f}_p^t \quad \forall p \in \mathcal{P} \quad (8b)$$

$$\underline{\pi}_n^t \leq \pi_n^t \leq \overline{\pi}_n^t \quad \forall n \in \mathcal{N}_f \quad (8c)$$

$$\pi_m^t \leq \beta_c^t \pi_n^t \quad \forall c = (n, m) \in \mathcal{C} \quad (8d)$$

$$0 \leq f_u^t \leq \overline{f}_u^t \quad \forall u \in \mathcal{U} \quad (8e)$$

$$0 \leq f_{s+}^t \leq V_{0s} - \underline{V}_s \quad \forall s \in \mathcal{S} \quad (8f)$$

$$0 \leq f_{s-}^t \leq \overline{V}_s - V_{0s} \quad \forall s \in \mathcal{S} \quad (8g)$$

$$V_s^t = V_s^{t-1} + f_{s-}^{t-1} - f_{s+}^{t-1} \quad \forall s \in \mathcal{S} \quad (8h)$$

# Gas system constraints

The gas balance constraints guarantee that the amount of gas entering a node equals the amount leaving it. On the other hand, the Weymouth equation encapsulates the complex physics of gas flow through pipelines between flow rates and pressures.

$$\sum_{m:(m,n) \in \mathcal{A}} f_m^t = \sum_{m':(n,m') \in \mathcal{A}} f_{m'}^t \quad \forall n \in \mathcal{N}_f \quad (9a)$$

$$\text{sgn}(f_p^t)(f_p^t)^2 = K_{nm}((\pi_n^t)^2 - (\pi_m^t)^2) \quad \forall p = (n, m) \in \mathcal{P} \quad (9b)$$

Modeling this equation as a nonconvex and discontinuous equality constraint poses significant challenges for optimization algorithms. Specifically, the nonconvexity leads to multiple local optima, making it difficult to find the global optimum [Jiang et al., 2021].

# Proposed complementarity-based approach

Complementarity constraints can be used to represent non-uniform or discontinuous operators, such as absolute value, sign, and min/max [Baumrucker et al., 2008].

$$y = \operatorname{sgn}(f_T^t) \quad (10)$$

The sign equation can be represented by an optimization problem whose constraints avoid the use of binary variables.

$$\mathcal{O}_W : \min_{y_p^t} -y_p^t f_p^t \quad (11a)$$

$$\text{s.t. } y_p^t (f_p^t)^2 = K_{nm} ((\pi_n^t)^2 - (\pi_m^t)^2) \quad (11b)$$

$$-1 \leq y_p^t \leq 1 \quad (11c)$$

$$f_p^t = f_{p+}^t - f_{p-}^t \quad (11d)$$

$$0 \leq f_{p+}^t \perp (y_p^t + 1) \geq 0 \quad (11e)$$

$$0 \leq f_{p-}^t \perp (1 - y_p^t) \geq 0 \quad (11f)$$

# Proposed complementarity-based approach

The program  $\mathcal{O}_W$  misses the KKT conditions, because the conventional constraint qualifications for nonlinear programming are typically not satisfied in the case of MPCC [Bouza and Still, 2007]. Hence, instead of working with the original problem, the relaxed problem  $\mathcal{O}_\epsilon$  is considered:

$$\mathcal{O}_\epsilon : \min_{y_p^t} -y_p^t f_p^t \quad (12a)$$

$$\text{s.t. } y_p^t (f_p^t)^2 = K_{nm} ((\pi_n^t)^2 - (\pi_m^t)^2) \quad (12b)$$

$$f_p^t = f_{p+}^t - f_{p-}^t \quad (12c)$$

$$-1 \leq y_p^t \leq 1 \quad (12d)$$

$$f_{p+}^t (y_p^t + 1) \leq \epsilon \quad (12e)$$

$$f_{p-}^t (1 - y_p^t) \leq \epsilon \quad (12f)$$

The regularized optimization model  $\mathcal{O}_\epsilon$  offers several key properties that justify its use in tackling challenging MPCC

# Databases

Each of the mentioned databases was selected based on its unique characteristics and the significant contributions it could make to the study.

	<b>Topology</b>	<b>Connection points</b>	<b>Closed loops</b>	<b>Problem</b>
<b>Case 1</b>	9-bus 8-system	1	1	Small system with one loop
<b>Case 2</b>	118-bus 48-system	9	7	Contains several interconnected loops
<b>Case 3</b>	96-bus 63-system	10	0	Fully radial but considers bidirectional flows.

**Table 2:** Databases used in the study

# Experimental setup

In order to establish a baseline for comparison, two alternative methodologies were employed:

- Taylor series approximation [Fodstad et al., 2015].
- Second Order Cone Programming (SOC) [Schwele et al., 2019].

The objective function value for each of the mentioned databases was computed by solving the optimization problem using:

- The IPOPT [Wächter and Biegler, 2005] solver.
- The GEKKO [Beal et al., 2018] package.

# Experimental setup

Since understanding and quantifying the inherent errors introduced by any constraint approximation approach supports its real-world pertinence, the considered validation trades off the reached cost function and constraint error values.

$$WE_p^t = |f_p^t - (K_{nm}|(\pi_n^t)^2 - (\pi_m^t)^2|)|^{1/2}, \quad \forall p = (n, m) \in \mathcal{P} \quad (13)$$

Hence,  $WE_p^t$  metric explains the inherent sensitivity of tested approaches and validates the significance of their differences.

## Results - Case 1

To assess the performance of Weymouth approximation approaches on the 8/9 system, a Monte Carlo experiment estimates the cost function and Weymouth error distributions by solving the optimization problem for one single day one hundred times with uniformly sampled natural gas demands.

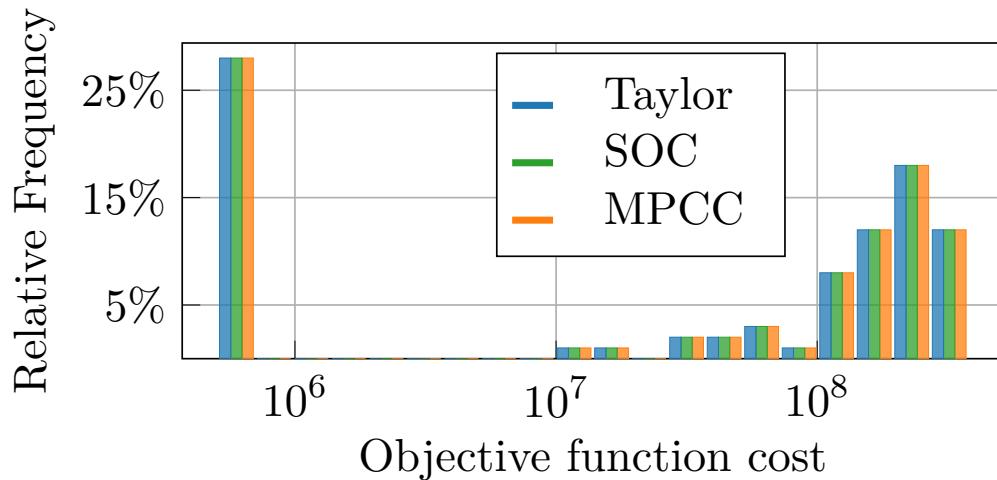


Figure 3: Cost function histogram attained for the Taylor, SOC, and MPCC Weymouth approximation approaches.

## Results - Case 1

The boxplot examination reveals significantly lower approximation errors on the proposed complementarity constraints approach over key network arcs.

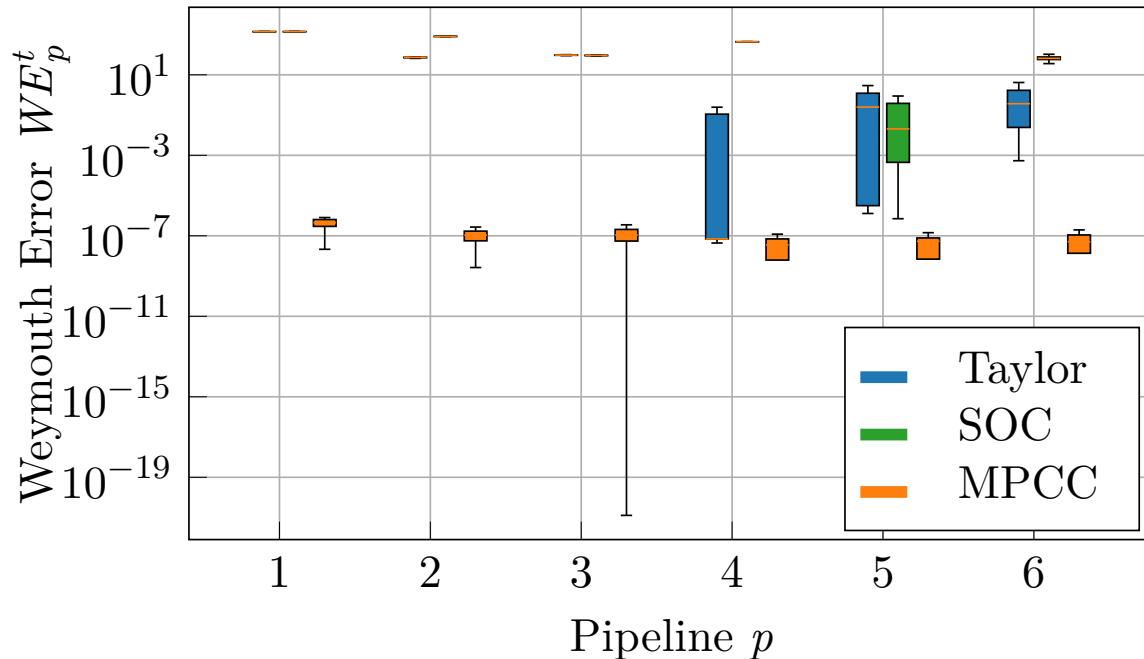


Figure 4: Boxplot of Weymouth error distribution for each pipeline in the 8/9 system attained by contrasted approximation approaches.

## Results - Case 2

It is worth noting that both baselines yielded the same objective function values. The consistently positive relative difference indicates that the complementarity constraints formulation always yields larger cost values than Taylor and SOC for the 118/48 system.

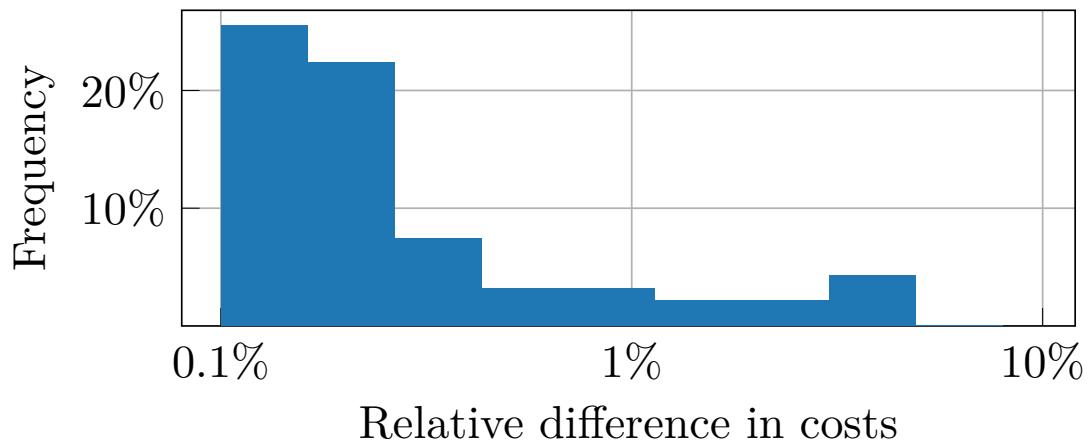
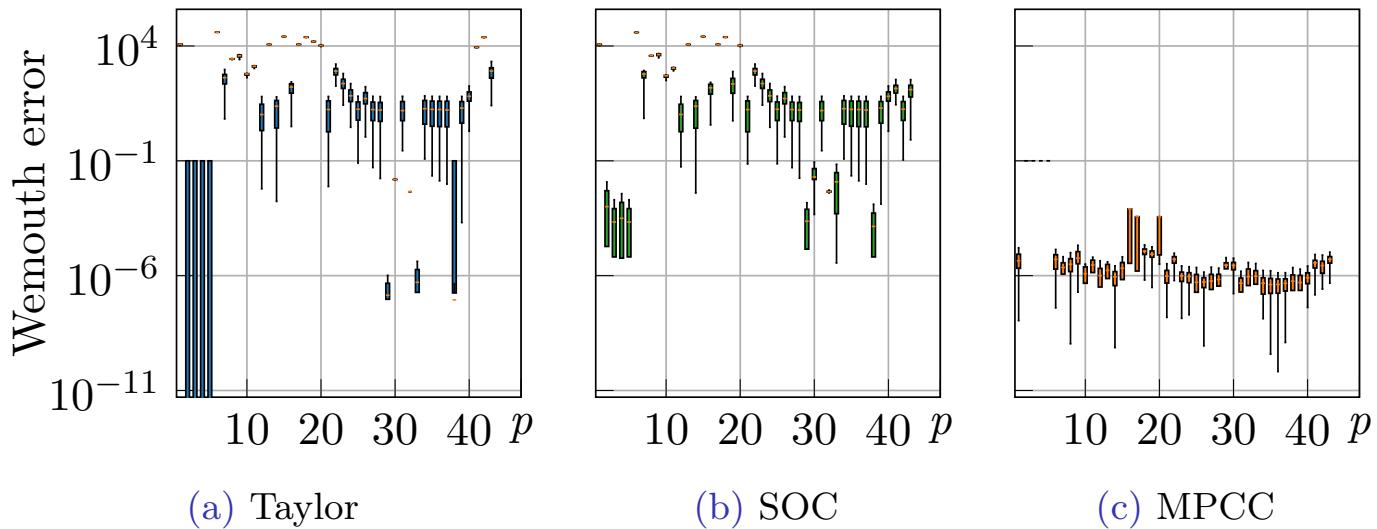


Figure 5: Histogram depicting the relative frequencies of cost differences obtained between MPCC and the other approaches in the 48-node 118-bus system.

## Results - Case 2

Contrarily to cost function analysis, Weymouth approximation results reveal a significant error reduction of about seven orders of magnitude (from  $10^1$  to  $10^{-6}$ ) under the proposed complementarity constraints.



**Figure 6:** Weymouth approximation errors for each pipeline  $p$  reached by Taylor (left), SOC (center), and MPCC (right) approaches in the 118/48 system.

## Results - Case 3

Instead of estimating the distributions of the objective function and Weymouth error as in cases 9/8 and 118/48, the 96/63 case validates the Weymouth approximations in an operation case of ten consecutive days ( $|\mathcal{T}| = 10$ ) with randomly changing gas extraction costs.

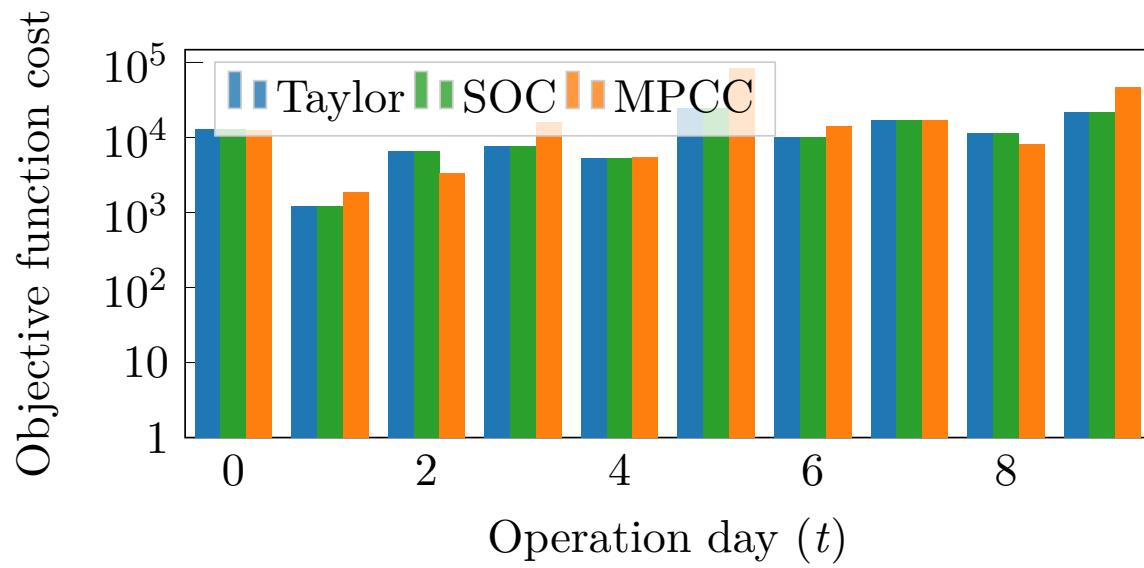


Figure 7: Daily operating cost obtained with each of the approaches in the 63-node 96-bus system.

# Results - Case 3

Regarding the Weymouth approximation analysis, the figure presents the error distribution and its relationship with the gas flow and the scheduled day for Taylor, SOC, and MPCC

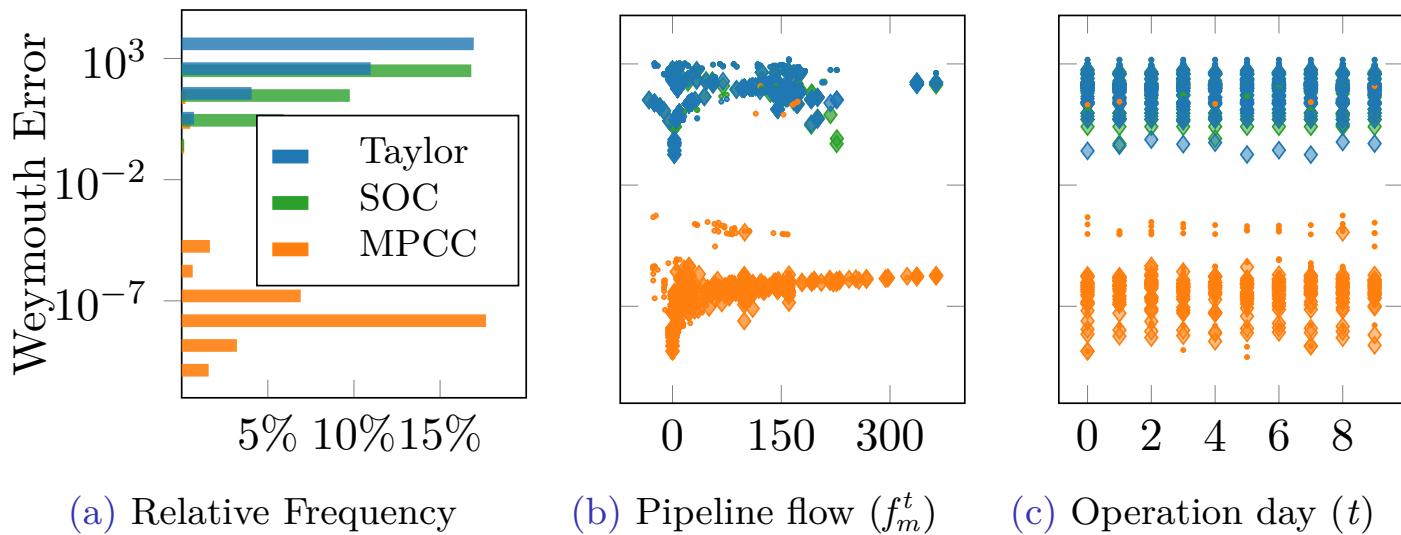


Figure 8: Errors in the Weymouth equation expressed in three different formats- general error, error per flow, and error per day.

# Physics-Guided Neural Networks

## Motivation:

- Traditional neural networks: fit data, ignore physics → risk of unrealistic predictions.
- In gas networks, physical constraints (mass balance, pressure-flow laws) are critical.

## Physics-Informed Neural Networks (PINNs):

- Incorporate governing equations directly in the loss function.
- Penalize deviations from:
  - **Gas balance constraint** – nodal mass conservation.
  - **Weymouth equation** – relation between flow and pressure differences in pipelines.
- Acts as a *regularizer* for better generalization under uncertainty.

# Methodology: Physics-Guided CensNet

**Overall loss function:**

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

**Gas balance loss:**

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

- Enforces nodal inflow = outflow + demand.

**Weymouth loss:**

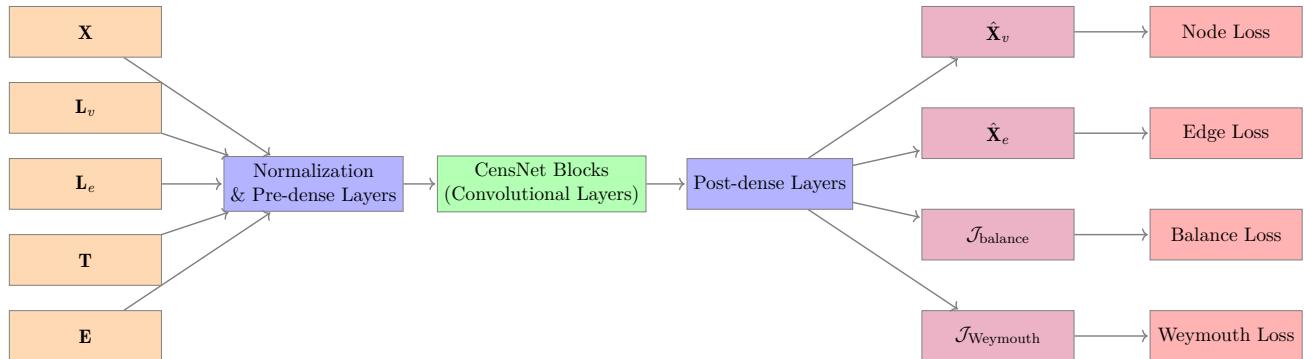
$$\mathcal{J}_{\text{weymouth}} = \mathbf{M}_{\mathcal{P}} \left( \hat{\mathbf{f}}_e^{\circ 2} - \mathbf{K} \circ \left( \mathbf{T} \cdot \hat{\boldsymbol{\pi}}^{\circ 2} \right) \right)$$

- Relates squared flows to squared pressure differences.
- Applies only to pipelines (not compressors).

# Methodology: Physics-Guided CensNet

## Implementation:

- Built on CensNet architecture from deterministic setup.
- Training data from nonlinear gas network optimization (MPCC).
- Noise (5%–25%) injected to emulate operating uncertainty.



# Case Study I: 8-Node Network – Performance Comparison

Method	Node Error	Edge Error	Balance Error	Time (s)
CensNet (N)	$0.00 \pm 17.99$	$22.76 \pm 15.43$	$-0.01 \pm 17.45$	$0.86 \pm 0.50$
CensNet (N+E)	$-0.11 \pm 18.27$	$0.22 \pm 21.65$	$-0.12 \pm 1.70$	$0.85 \pm 0.50$
CensNet (N+E+B)	$-0.02 \pm 18.00$	$-0.02 \pm 21.30$	$-0.03 \pm 0.90$	$0.85 \pm 0.50$
CensNet (N+E+W)	$-0.07 \pm 17.56$	$2.60 \pm 20.03$	$-0.07 \pm 2.21$	$0.86 \pm 0.50$
CensNet (N+E+B+W)	$0.05 \pm 17.91$	$0.25 \pm 21.14$	$0.05 \pm 1.69$	$0.85 \pm 0.50$

Table 3: Performance of CensNet-based models vs. IPOPT benchmark.

- Using only node loss leads to accurate nodal flows but large errors in edge predictions.
- Adding edge loss (**N+E**) reduces edge error from 22.76 to 0.22 and improves global balance.
- Including balance loss (**N+E+B**) further minimizes balance error, achieving best overall consistency.
- Weymouth loss (**N+E+W**) introduces slight trade-offs, slightly increasing edge/balance errors but retaining nodal accuracy.
- All CensNet variants are over an order of magnitude faster than IPOPT, with **N+E+B** offering the best accuracy-speed compromise.

# Stochastic Analysis: Methodology

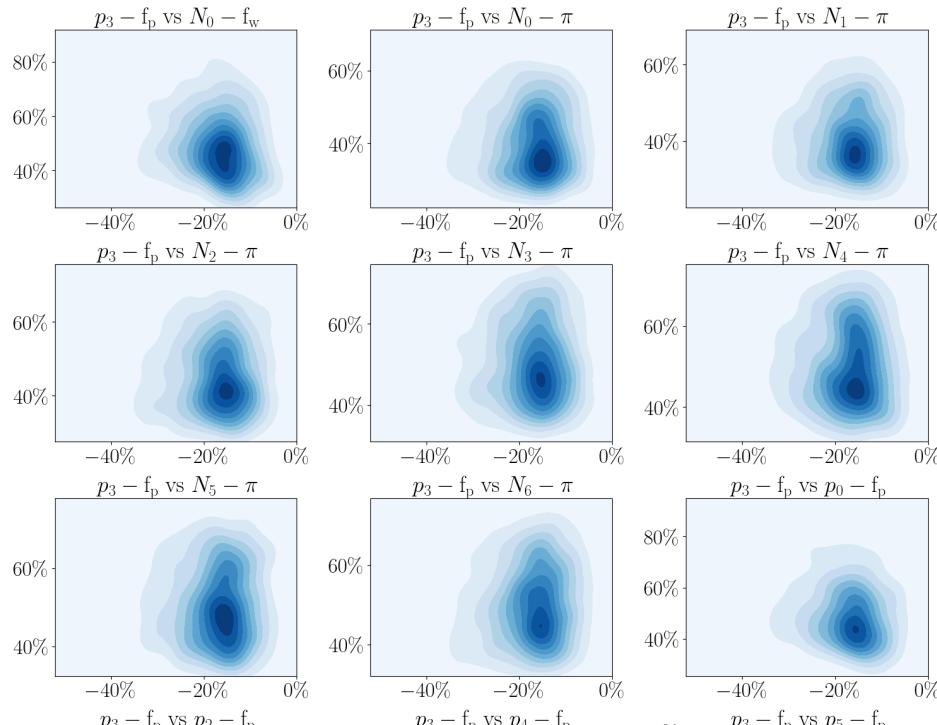
- Kernel Density Estimate (KDE) fitted to training inputs.
- Synthetic samples generated and propagated through CensNet.
- Log-likelihoods compared for  $y_{\text{sample}}$  and  $\bar{y}_{\text{train}}$ .
- Kolmogorov–Smirnov (K–S) test used for distributional similarity.

# Model Robustness under Uncertainty

- Log-likelihoods:  $y_{\text{sample}} = -6.696 \times 10^6$ ,  $\bar{y}_{\text{train}} = -6.657 \times 10^6$ .
- K–S test:  $p > 0.44$  across all alternatives.
- ⇒ Synthetic outputs statistically consistent with training outputs.

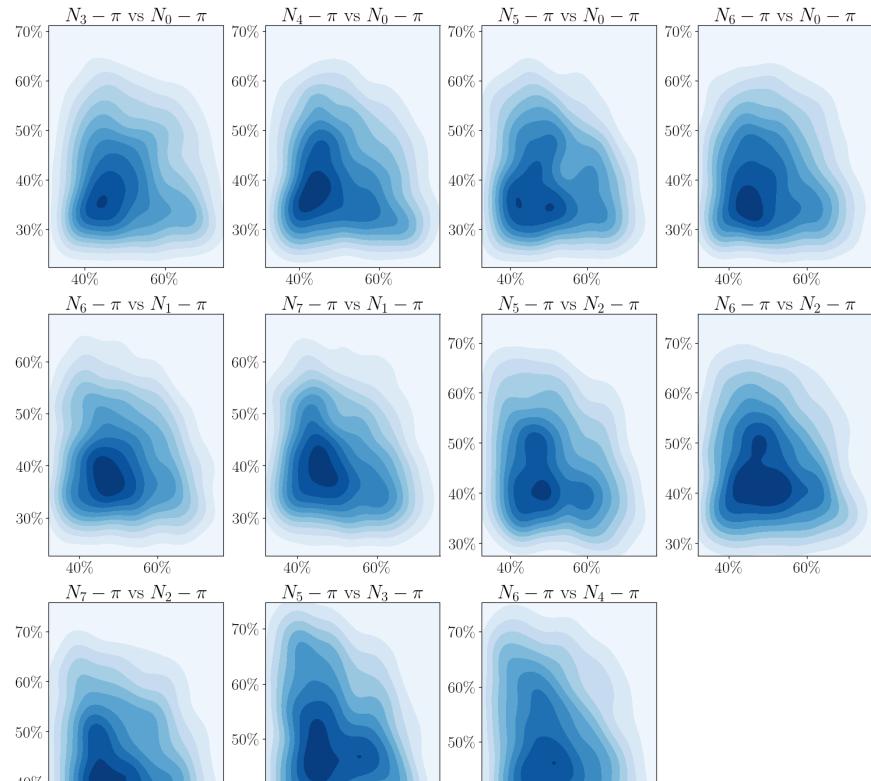
# Negative Flows in Pipeline $p_3$

- Flow in  $p_3$  consistently negative across scenarios and KDE samples.
- Caused by orientation assumption → optimizer adjusts direction.
- Operability unaffected, but  $p_3$  supports cost minimization.



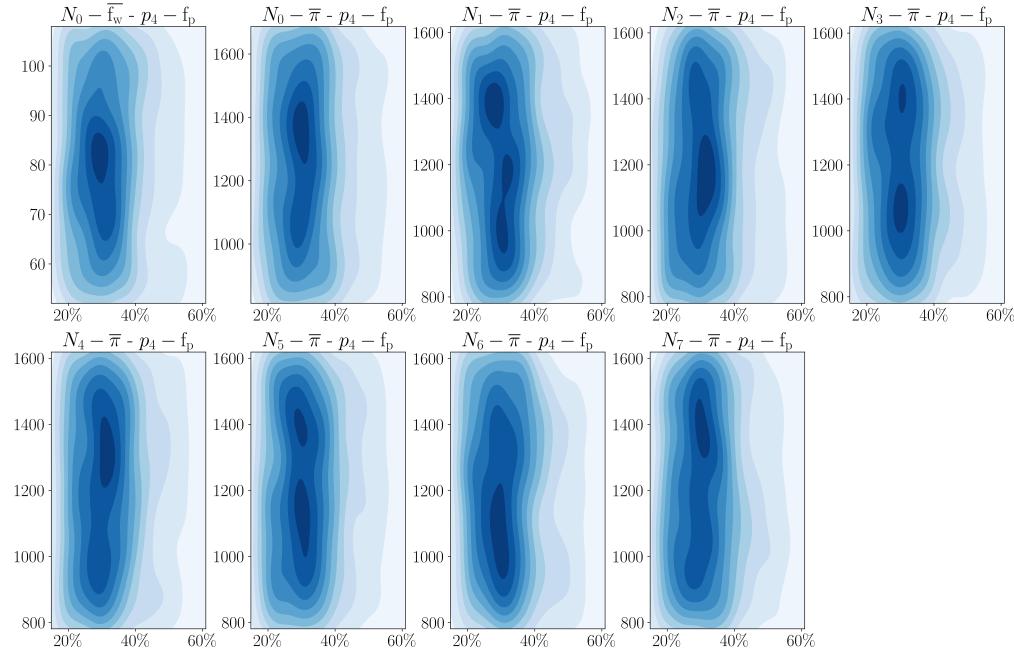
# Pressure Concentration in Mid-Range

- Nodal pressures cluster at 30–40% of normalized limits.
- Avoids extremes, ensuring reliable operation.
- Implies optimizer favors efficiency + stability.



# Stable Utilization of Pipeline $p_4$

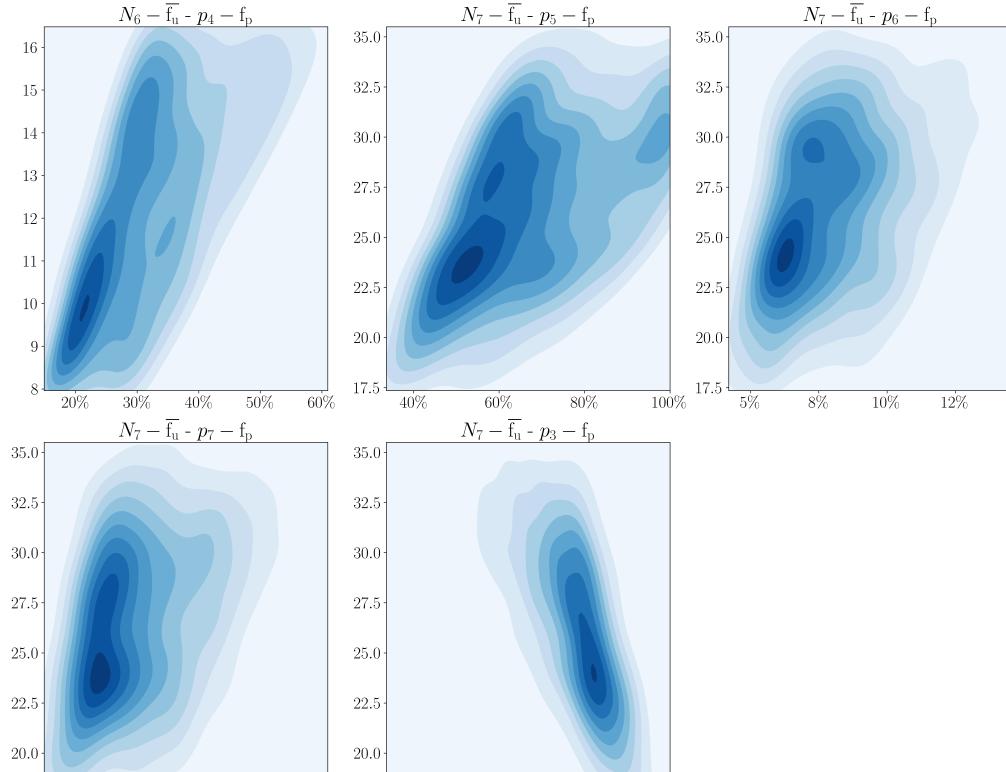
- Despite wide dispersion in inputs,  $p_4$  flow remains 20–40%.
- Reflects robustness: optimizer stabilizes this delivery path.
- Behavior consistently learned and reproduced in KDE samples.



Optimization model: input variability vs. stable  $p_4$  flow.

# Linear Input–Output Relations

- Positive correlation: demand bounds ( $N_6, N_7$ ) vs.  $p_4$ – $p_7$  flows.
- Negative correlation:  $N_7$  demand bound vs.  $p_3$  flow.
- Shows network internalized optimizer's allocation strategies.



# Key Conclusions

- **GNN-based prediction:** CensNet achieved accurate and scalable predictions for natural gas networks, with significant reductions in computation time compared to traditional optimization.
- **Optimization with MPCC:** Proposed formulation rigorously approximated Weymouth constraints, improving accuracy of pressure–flow representation and showing robustness in real-world scheduling tasks.
- **Stochastic modeling:** Physics-guided neural networks preserved structural dependencies and generalized reliably under uncertainty, validated through KDE sampling and K–S testing.
- **Overall:** Integrating physical constraints with GNNs yields efficient, accurate, and robust surrogates for optimization in natural gas systems.

# Future Work

- Extend stochastic framework toward **full stochastic optimization** for gas–power coordination.
- Explore **physics-informed neural operators** to improve learning of nonlinear constraints (e.g., Weymouth).
- Incorporate **real operational data** from SCADA systems to validate scalability in live environments.
- Investigate **transfer learning across networks**, enabling models trained on one topology to adapt to others with minimal retraining.
- Apply developed methods to **decision-support tools** for dispatchers in interconnected energy systems.

# References I

-  Baumrucker, B., Renfro, J., and Biegler, L. (2008).  
Mpec problem formulations and solution strategies with chemical engineering applications.  
*Computers & Chemical Engineering*, 32(12):2903–2913.
-  Beal, L. D. R., Hill, D. C., Martin, R. A., and Hedengren, J. D. (2018).  
Gekko optimization suite.  
*Processes*, 6(8).
-  Bouza, G. and Still, G. (2007).  
Mathematical programs with complementarity constraints:  
Convergence properties of a smoothing method.  
*Mathematics of Operations Research*, 32(2):467–483.

## References II



- Duan, J., Liu, F., and Yang, Y. (2022).  
Optimal operation for integrated electricity and natural gas systems considering demand response uncertainties.  
*Applied Energy*, 323:119455.
- Fodstad, M., Midthun, K. T., and Tomasdard, A. (2015).  
Adding flexibility in a natural gas transportation network using interruptible transportation services.  
*European Journal of Operational Research*, 243(2):647–657.
- García-Marín, S., González-Vanegas, W., and Murillo-Sánchez, C. (2022).  
Mpng: A matpower-based tool for optimal power and natural gas flow analyses.  
*IEEE Transactions on Power Systems*, pages 1–9.

## References III

-  Jiang, T., Yuan, C., Zhang, R., Bai, L., Li, X., Chen, H., and Li, G. (2021).  
Exploiting flexibility of combined-cycle gas turbines in power system unit commitment with natural gas transmission constraints and reserve scheduling.  
*International Journal of Electrical Power & Energy Systems*, 125:106460.
-  Promigas (2021).  
Informe del sector gas natural 2021.
-  Ralph, D. and Wright, S. J. (2004).  
Some properties of regularization and penalization schemes for mpecs.  
*Optimization Methods and Software*, 19(5):527–556.

## References IV

-  Schwele, A., Ordoudis, C., Kazempour, J., and Pinson, P. (2019). Coordination of power and natural gas systems: Convexification approaches for linepack modeling.  
In *2019 IEEE Milan PowerTech*, pages 1–6.
-  Wächter, A. and Biegler, L. T. (2005). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming.  
*Mathematical Programming*, 106(1):25–57.
-  Yang, Z., Xie, K., Yu, J., Zhong, H., Zhang, N., and Xia, Q. (2019). A general formulation of linear power flow models: Basic theory and error analysis.  
*IEEE Transactions on Power Systems*, 34(2):1315–1324.