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 10, 2025

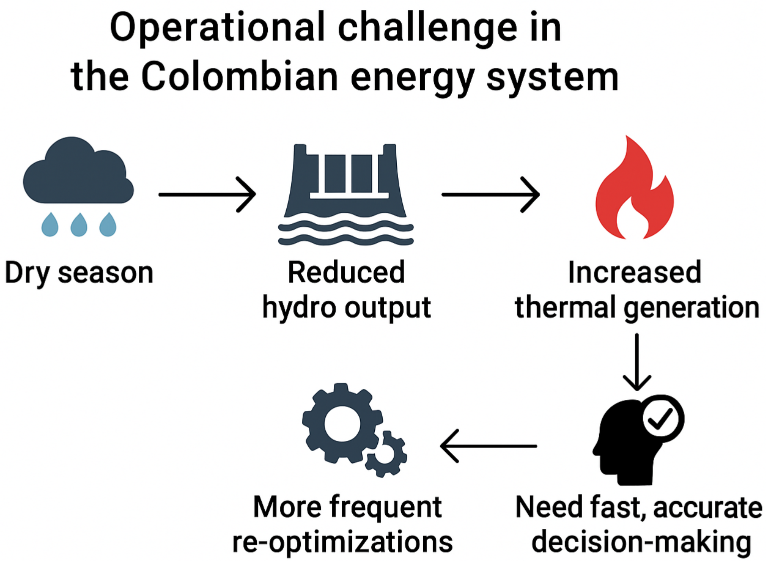
Energy Context & Relevance of Natural Gas

Natural gas plays a crucial role in the global energy mix, offering a cleaner alternative to other fossil fuels by producing lower CO_2 emissions. It supports both industrial processes and electricity generation, acting as a reliable energy source in diverse contexts.

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.

Operational Challenge in the Colombian Energy System

Colombia's electricity supply relies heavily on hydroelectric plants, which are highly dependent on rainfall patterns. During dry seasons, reduced water availability forces the system to increase the participation of thermal plants powered by natural gas.
This seasonal shift places significant operational stress on the gas transport network.



Motivation & Problem Statement

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

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.

Objectives

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 Prediction Using Graph Neural Networks

- Traditional optimizers are **accurate but slow** for large-scale or real-time scenarios.
- Gas networks are inherently **graphs**: nodes = wells, users, storage; edges = pipelines, compressors.
- GNNs exploit **topology-aware learning** for faster prediction.
- **Goal:** Achieve near-optimizer accuracy with 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:

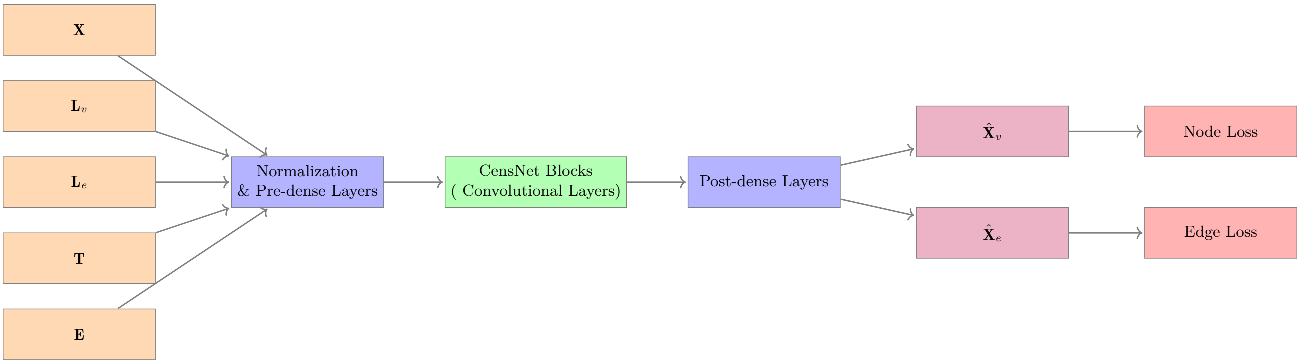
$$\mathbf{H}_e^{(l+1)} = \sigma\left(\mathbf{T}^\top\Phi(\mathbf{H}_v^{(l)}\mathbf{p}_v)\mathbf{T} \odot \tilde{\mathbf{A}}_e\mathbf{H}_e^{(l)}\mathbf{W}_e\right)$$

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, lengths, compressor ratios.
- Incidence matrix ($\mathbf{T}$): explicit mapping between nodes and edges.

CensNet-based Model Architecture

- **Processing:**
  - Normalization and pre-dense layers transform heterogeneous inputs into a common latent space.
  - *CensNet convolutional 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 injected 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.

Task-Dependent Loss Functions

- Regression task ⇒ Regularized MSE loss:

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

- Supports:
  - **Nodal losses** – Injected flows
  - **Edge losses** – Transported flows
  - **Gas balance** – Physical consistency
- Regularization avoids overfitting.

Test Systems

- **8-node network:** Small-scale testbed.
- **63-node network:** Realistic Colombian gas infrastructure.
- Data: Generated via optimization model (**??**) with physical constraints (**??**)–(**??**).

Figure 2: Case study networks.

Runtime vs Accuracy Trade-off

- Optimizer: High accuracy, high runtime.
- CensNet: Comparable accuracy, much faster inference.
- Statistical validation: t-test confirms no significant loss in accuracy (p-values $\ll 0.05$).

Figure 3: Prediction accuracy vs computation time.