



COPPE
UFRJ

PHYSICS-INFORMED NEURAL NETWORK MODELS FOR A DIGITAL TWIN
OF AN OFFSHORE GAS SEPARATION UNIT

Pedro Ricardo Cardoso Gonçalves

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Química, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Química.

Orientadores: Argimiro Resende Secchi
Maurício Bezerra de Souza
Júnior

Rio de Janeiro
Abril de 2023

PHYSICS-INFORMED NEURAL NETWORK MODELS FOR A DIGITAL TWIN
OF AN OFFSHORE GAS SEPARATION UNIT

Pedro Ricardo Cardoso Gonçalves

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO
GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA QUÍMICA.

Examinada por:

Prof. Argimiro Resende Secchi, D.Sc.

Prof. Maurício Bezerra de Souza Júnior, D.Sc.

Prof. Fabio Pereira dos Santos, D.Sc.

Prof. Natan Padoin, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
ABRIL DE 2023

Gonçalves, Pedro Ricardo Cardoso

Physics-Informed neural network models for a digital twin of an offshore gas separation unit/Pedro Ricardo Cardoso Gonçalves. – Rio de Janeiro: UFRJ/COPPE, 2023.

XXIII, 140 p.: il.; 29,7cm.

Orientadores: Argimiro Resende Secchi

Maurício Bezerra de Souza Júnior

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Química, 2023.

Referências Bibliográficas: p. 109 – 123.

1. Model development. 2. Artificial neural network. 3. Digital twin. 4. Gas separation. 5. Offshore. 6. Membrane. I. Secchi, Argimiro Resende *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química. III. Título.

*Para a indústria brasileira de óleo e gás. Que seja um
pontapé inicial para a era digital.*

Agradecimentos

Primeiramente gostaria de agradecer a Deus pela oportunidade da jornada acadêmica e ao meu guia e protetor espiritual pela proteção e esclarecimento ao longo dos meus caminhos, confortando-me nos momentos tempestuosos e dando forças nas etapas em que precisava me reerguer e continuar a trilhar o caminho para a realização dos meus sonhos.

Ao Professor Argimiro, gostaria de agradecer por todo apoio emocional, e técnico, com muita paciência, sabedoria e auxiliando, tirando dúvidas e me ajudando a trilhar o caminho das pedras em me tornar mestre. Também agradeço pela dedicação em corrigir meu trabalho e apresentações de maneira minuciosa sempre apontando o porque das mudanças. Muito obrigado.

Ao professor Maurício, por todo o incentivo, apontando as direções mais interessantes e desafiadoras. Por aguçar a minha curiosidade em descobrir sobre o mundo da Inteligência Artificial e deixar esse trabalho mais rico e atualizado com as mudanças da indústria. Muito Obrigado.

Agradeço à Universidade Federal do Rio de Janeiro e seu corpo docente de professores que não mediram esforços em repassar a dádiva do conhecimento, não apenas racional, mas também moral e ético nesta nova etapa do meu processo de formação profissional.

Agradeço à banca julgadora por terem aceito o convite e por avaliar nossa aptidão ao exercício da profissão.

O meu agradecimento aos meus colegas de faculdade que me acompanharam durante esta jornada que me auxiliaram nos estudos e que compartilharam dos momentos de descontração.

Por fim, mas não menos importante, à instituição de fomento à pesquisa do Brasil, CAPES, que financiou meus estudos para a realização deste trabalho.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MODELOS POR REDE NEURONAL FÍSICO-INFORMADAS PARA UM
DIGITAL TWIN DE UMA UNIDADE OFFSHORE DE SEPARAÇÃO DE GASES

Pedro Ricardo Cardoso Gonçalves

Abril/2023

Orientadores: Argimiro Resende Secchi
Maurício Bezerra de Souza Júnior

Programa: Engenharia Química

A nova revolução industrial, conhecida como Indústria 4.0, têm atraído empresas a investir em representações digitais, ou gêmeos digitais, de seus processos reais utilizando-se dados de planta para melhor modelar, controlar e detectar possíveis falhas de operação.

Neste trabalho foi desenvolvido um gêmeo digital de um sistema de desacidificação de gás natural em plataformas *offshore*, por meio de separação por membranas. Para modelagem matemática do sistema, foram combinadas duas abordagens distintas: por via fenomenológica, com base na literatura, e por aprendizado de máquina, constituída de redes neuronais. O trabalho visou a junção destas duas vertentes em duas formas de modelo híbrido, a primeira na qual os balanços de massa e quantidade de movimento são desenvolvidos por meio fenomenológico e a modelagem da transferência de massa através das membranas foi feita via redes neuronais; e a segunda em que a rede neuronal é informada pela física do sistema de forma a atender às leis de conservação. Como resultado principal, ambos os modelos se mostraram bastante acurados, com erros relativos inferiores a 2%, e podem ser usados como uma abordagem alternativa para quando se desconhecem os parâmetros dos equipamentos utilizados ou para facilitar a inclusão dos modelos em abordagens de otimização e controle, e assim contribuir para digitalização da indústria de óleo e gás.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

PHYSICS-INFORMED NEURAL NETWORK MODELS FOR A DIGITAL TWIN
OF AN OFFSHORE GAS SEPARATION UNIT

Pedro Ricardo Cardoso Gonçalves

April/2023

Advisors: Argimiro Resende Secchi
Maurício Bezerra de Souza Júnior

Department: Chemical Engineering

The new industrial revolution, known as Industry 4.0, has attracted companies to invest in digital representations, or digital twins, of their actual processes using plant data to better model, control, and detect possible operational failures.

In this work, a digital twin of a system for deacidification of natural gas in *off-shore* platforms, by membrane separation, was developed. For the mathematical modeling of the system, two distinct approaches were combined: phenomenological, based on the literature, and machine learning, consisting of neural networks. The work aimed at joining these two approaches in two forms of a hybrid model, the first in which the mass and momentum balances are developed by phenomenological means and the modeling of the mass transfer through the membranes was done via neural networks; and the second in which the neural network is informed by the physics of the system in order to meet the conservation laws. As a main result, both models have proven to be very accurate, with relative errors lower than 2%, and can be used as an alternative approach when the parameters of the equipment used are unknown or to facilitate the inclusion of the models in optimization and control approaches, and thus contribute to the digitalization of the oil and gas industry.

Contents

| | |
|--|-------------|
| Agradecimientos | v |
| List of Figures | xi |
| List of Tables | xiv |
| List of Symbols | xvi |
| List of Abbreviations | xxii |
| 1 Introduction | 1 |
| 1.1 Natural Gas | 1 |
| 1.2 Motivation | 3 |
| 1.3 Objectives | 5 |
| 1.4 Structure | 5 |
| 2 Background and Literature Review | 8 |
| 2.1 Gas Separation Processes | 8 |
| 2.1.1 Cryogenic Method | 9 |
| 2.1.2 Amine Absorption Method | 12 |
| 2.1.3 Membrane-based Method | 15 |
| 2.2 Digital Twin Concepts | 22 |
| 2.2.1 Building a Digital Twin | 24 |
| 2.3 Membrane Phenomenological Model | 26 |
| 2.3.1 Hollow-Fiber model | 26 |
| 2.4 Artificial Neural Network | 33 |
| 2.4.1 Neural Cell structures and model | 33 |
| 2.4.2 Advantages of Neural Network | 37 |
| 2.4.3 ANN training process | 38 |
| 2.4.4 Multi-Layer-Perceptron Network | 41 |
| 2.5 Physics-Informed Neural Networks | 44 |
| 2.5.1 PINNs Structure and Logic | 45 |

| | | |
|----------|---|-----------|
| 2.6 | Literature Review | 46 |
| 3 | Phenomenological model | 52 |
| 3.1 | Introduction | 52 |
| 3.2 | Methodology | 53 |
| 3.2.1 | Hollow fiber modules model and discretization | 53 |
| 3.2.2 | Model Discretization | 58 |
| 3.2.3 | Molar volume calculation | 60 |
| 3.2.4 | Diffusion Coefficient calculation for low-density gases | 61 |
| 3.2.5 | Viscosity calculation for binary mixture gases | 62 |
| 3.2.6 | Permeance calculation | 64 |
| 3.3 | Results and Discussion | 65 |
| 3.3.1 | Diffusion Coefficient Contribution Evaluation | 65 |
| 3.3.2 | Model Validation 1 | 68 |
| 3.3.3 | Model Validation 2 | 69 |
| 3.4 | Conclusion | 70 |
| 4 | Multi-layer Perceptron Network development | 71 |
| 4.1 | Introduction | 71 |
| 4.2 | Methodology | 72 |
| 4.2.1 | Data evaluation and treatment | 72 |
| 4.2.2 | Neural Network modeling | 74 |
| 4.2.3 | Architecture search | 75 |
| 4.2.4 | The Hybrid Model | 76 |
| 4.3 | Results and Discussion | 76 |
| 4.3.1 | Data Treatment | 76 |
| 4.3.2 | Neural network Parameter Study | 78 |
| 4.3.3 | Neural network architecture search | 81 |
| 4.3.4 | Hybridization performance evaluation | 84 |
| 4.4 | Conclusions | 85 |
| 5 | Physics-Informed Neural Network | 86 |
| 5.1 | Introduction | 86 |
| 5.2 | Methodology | 87 |
| 5.2.1 | The loss function | 87 |
| 5.2.2 | Data generation | 88 |
| 5.2.3 | PINN and Data-Driven models structure and parameters | 90 |
| 5.3 | Results and Discussion | 90 |
| 5.3.1 | One module data-driven model | 90 |
| 5.3.2 | One module PINN model | 94 |

| | | |
|----------|--|------------|
| 5.3.3 | Digital Twin data-driven model | 97 |
| 5.3.4 | Digital Twin PINN model | 99 |
| 5.4 | Conclusions | 102 |
| 6 | Conclusion | 104 |
| 6.1 | Developments Overview | 104 |
| 6.2 | Work Contributions | 106 |
| 6.3 | Proposal for Continuation | 107 |
| | Bibliography | 109 |
| A | Membrane diffusion equation | 124 |
| B | MLP network back-propagation algorithm | 125 |
| C | Larger real FPSO Gas separation process diagram | 129 |
| D | Fugacity calculation | 131 |
| E | Model discretization procedure | 134 |
| F | Raw dataset X Treated dataset chart comparison | 137 |

List of Figures

| | | |
|------|---|----|
| 1.1 | Gap between natural gas production and demand in Brazil. | 2 |
| 2.1 | Vapor pressure of main NG components, using Antoine equation. . . | 10 |
| 2.2 | NG cryogenic separation block diagram. | 11 |
| 2.3 | Generic steam deacidification process using absorbents. | 13 |
| 2.4 | Main amines compounds molecular structures. | 14 |
| 2.5 | Structure differences between isotropic and anisotropic porous dis- position. | 16 |
| 2.6 | Structure of a hollow-fiber module. | 18 |
| 2.7 | Structure of a spiral wound module. | 18 |
| 2.8 | Gas separation process scheme. | 21 |
| 2.9 | The evolution of industrial eras. | 24 |
| 2.10 | Co-current hollow-fiber module representation. | 28 |
| 2.11 | Counter-current hollow-fiber module representation. | 31 |
| 2.12 | Biological neural cell structure | 33 |
| 2.13 | Activation function profiles comparison | 35 |
| 2.14 | Artificial neural cell structure. | 36 |
| 2.15 | Error-correction-based artificial neural cell structure. | 40 |
| 2.16 | Multi-Layer Perceptron Network Scheme | 41 |
| 2.17 | PINN scheme. | 45 |
| 3.1 | A real FPSO gas separation process diagram. | 53 |
| 3.2 | Boundary conditions representation from the three different mod- ules found in this process, where 1 specified entrance condition, 2 is zero flux, 3 represents the continuity of flux and variables, and 4 is where the gradient is null. | 57 |
| 3.3 | General 1D mesh representation of finite-volume method, with the central volume P and the neighboring volumes W and E. The border midpoints are represented by w and e, respectively. | 59 |
| 3.4 | Average Péclet number over operation conditions for model tube side. | 66 |

| | | |
|-----|--|-----|
| 3.5 | Average Péclet number over operation conditions for model shell side. | 67 |
| 3.6 | Model validation. | 70 |
| 4.1 | Normalized raw data sheet from feed variables. Sampling time of 1 hour. | 73 |
| 4.2 | Roadmap for the search of key parameters non-structural from the manually-set Neural Network. | 75 |
| 4.3 | Comparison between the proposed pure phenomenological modeling workflow and the Hybrid process. | 77 |
| 4.4 | Normalized treated data dispersion from feed variables. | 78 |
| 4.5 | Loss behavior as the network is trained. | 81 |
| 4.6 | Permeance prediction result for the model with two output layers. | 82 |
| 4.7 | Permeance prediction result for the two models with one output layer each. | 83 |
| 5.1 | Roadmap for the PINN and data-driven model development and performance evaluation. | 89 |
| 5.2 | First results for molar flowrates, permeabilities, and side pressure evolution for the data-driven model using one input and one output for each network. | 91 |
| 5.3 | Best results for molar flowrates, permeabilities, and pressure evolution for the data-driven model using one input and one output for each network. | 92 |
| 5.4 | Best results for molar flowrates, permeabilities, and side pressure evolution for the data-driven model using four inputs and one output for each network. | 94 |
| 5.5 | First results for molar flowrates, permeabilities, and side pressure evolution for the PINN model using four inputs and one output for each network. | 95 |
| 5.6 | Best results for molar flowrates, permeabilities, and cavities pressure evolution for the PINN model using four inputs and one output for each network. | 96 |
| 5.7 | Best results for molar flowrates, permeabilities, and side pressure evolution for the Digital Twin data-driven-modeled. | 98 |
| 5.8 | First results of molar flowrates, permeabilities, and side pressure evolution for the PINN model using four inputs and one output for each network. | 99 |
| 5.9 | Best results for molar flowrates, permeabilities, and side pressure evolution for the digital twin PINN-modeled. | 100 |

| | | |
|------|---|-----|
| 5.10 | Best results for shell pressure after the application of the error adjustment function. | 102 |
| C.1 | Enlarged operational FPSO gas separation process diagram. | 130 |
| E.1 | General 1D mesh representation of finite-volume method, with the central volume P and the neighboring volumes W and E. The border midpoints are represented by w and e, respectively. | 134 |
| F.1 | Normalized raw data sheet from feed variables. Time step=1hour. | 138 |
| F.2 | Normalized treated data dispersion from feed variables. | 138 |
| F.3 | Normalized raw data sheet from retentate variables. Time step=1hour. | 139 |
| F.4 | Normalized treated data dispersion from retentate variables. | 139 |
| F.5 | Normalized raw data sheet from permeate variables. Time step=1hour. | 140 |
| F.6 | Normalized treated data dispersion from permeate variables. | 140 |

List of Tables

| | | |
|-----|--|----|
| 2.1 | Atmospheric concentration and energy reflective capacity from main GHG. | 9 |
| 2.2 | selectivity of different membrane compounds. | 20 |
| 2.3 | Seperex UOP membrane benchmarks around the globe. | 21 |
| 3.1 | Hollow fiber membrane module specifications comparison between the patent from CASKEY (1993) and those used for the digital-twin structure proposal. | 58 |
| 3.2 | Peng-Robinson thermodynamic parameters for the fugacity calculation | 62 |
| 3.3 | Chapman-Enskog parameters for the Diffusion Coefficient and viscosity calculation | 63 |
| 3.4 | Parameters for cellulose acetate hollow-fiber membranes | 65 |
| 3.5 | Parameters for permeation evaluation considering binary mixture CO_2/CH_4 and linear approximation. | 65 |
| 3.6 | CHU <i>et al.</i> (2019) Scenario 1 parameters and feed conditions. | 68 |
| 3.7 | Model validation. | 69 |
| 3.8 | Parameters and feed conditions for model validation 2. | 69 |
| 4.1 | Model adjustment comparison between raw data and normalized data training sets. | 79 |
| 4.2 | Model adjustment comparison between ADAM and ADAGRAD optimizer algorithms. | 79 |
| 4.3 | ADAGRAD performance with a hyperbolic tangent activation function. | 80 |
| 4.4 | Model adjustment comparison between Tanh and Sigmoid activation functions. | 80 |
| 4.5 | ADAM optimizer performance with a sigmoid activation function. | 81 |
| 4.6 | Hybrid models and their comparison with the purely phenomenological model proposed on Chapter 3 | 85 |

| | | |
|-----|--|-----|
| 5.1 | Curve adjustment comparison between the data-driven and PINN models' KPIs for normalized features. | 97 |
| 5.2 | Digital-Twin curve adjustment comparison between the data-driven and PINN models' KPIs for normalized features. | 101 |
| 5.3 | Digital-Twin curve adjustment comparison between the data-driven and adjusted PINN (aPINN) models' KPIs for normalized features. | 101 |
| D.1 | Peng-Robinson thermodynamic parameters for the fugacity calculation | 133 |

List of Symbols

| | |
|-------------|--|
| ρ^m | Mixture molar density [mol/m^3], p. 31 |
| ξ | Error value for the PINN shell pressure prediction, p. 101 |
| s_p | Tube side flow speed [m/s], p. 55 |
| s_s | Shell side flow speed [m/s], p. 55 |
| A_G | General phenomenological coefficient, p. 17 |
| A_m | Total membrane area [m^2], p. 32 |
| A_o | Single fiber outside area [m^2], p. 29 |
| C_P | Fluid's calorific capacity at constant pressure [J/K], p. 32 |
| D | Flowline diameter [m], p. 31 |
| D_i^{ABs} | i^{th} Shell side component diffusion coefficient [m^2/s], p. 28 |
| D^{AB} | Diffusion coefficient parameter [m^2/s], p. 124 |
| D_i^{AB} | i^{th} Component diffusion coefficient parameter [m^2/s], p. 124 |
| D_i | Membrane inside diameter [m], p. 28 |
| D_o | Membrane outside diameter [m], p. 28 |
| D_{xh} | Shell side hydraulic diameter [m], p. 55 |
| J_G | General flux, p. 17 |
| J_i | i^{th} compound molar flux [$mol/s \cdot m^2$], p. 20 |
| KPI | Key Process Indicator, p. 53 |
| K | Thermal conductivity [$J/m \cdot K \cdot s$], p. 32 |
| K_w | Membrane fiber thermal conductivity coefficient [$J/m \cdot K \cdot s$], p. 33 |

| | |
|-------------|---|
| L | Membrane Length [m], p. 28 |
| N | Module's total fiber amount, p. 28 |
| N^c | Flow's total compound amount, p. 31 |
| O | Next layer's neural cell amount, p. 41 |
| Obj | Neural cell learning objective function, p. 40 |
| P_i | i^{th} component vapor pressure [Pa], p. 30 |
| P_i^c | i^{th} component critical pressure [Pa], p. 30 |
| P_i^{sat} | i^{th} component saturation pressure [Pa], p. 30 |
| P_x | Shell side pressure [Pa], p. 29 |
| P_y | Tube side pressure [Pa], p. 29 |
| Q_{mem} | Heat flux [$J/m^2 \cdot s$], p. 32 |
| R | Universal gas constant [$m^3 \cdot Pa/K \cdot mol$], p. 30 |
| S_i | i^{th} coefficient solubility into the membrane sheet [$m_{gas}^3/m_{mem}^3 \cdot Pa$], p. 124 |
| T | Temperature [K], p. 30 |
| T_f^s | Shell side final temperature [K], p. 32 |
| T_i^s | Shell side initial temperature [K], p. 32 |
| T_f^t | Tube side final temperature [K], p. 32 |
| T_i^t | Tube side initial temperature [K], p. 32 |
| T_i^c | i^{th} component critical temperature [K], p. 30 |
| T_i^r | i^{th} component reduced temperature, p. 30 |
| T_e | Dimensionless temperature regarding molecules characteristic energy, p. 62 |
| U_0 | Global heat transfer coefficient [$J/m^2 \cdot K \cdot s$], p. 32 |
| V_i | i^{th} component volume [m^3], p. 30 |
| V_m | Total membrane volume [m^3], p. 32 |

| | |
|-----------------|---|
| V_o | Single fiber outside volume [m^3], p. 29 |
| W_{omn} | m^{th} hidden layer synaptic weight from n^{th} neural cell to next layer's O^{th} neural cell, p. 41 |
| X_i | Shell side i^{th} component molar fraction, p. 28 |
| Y_i | Tube side i^{th} component molar fraction, p. 28 |
| Z_x | Peng-Robinson Compressibility factor for the shell side composition, p. 56 |
| Z_y | Peng-Robinson Compressibility factor for the tube side composition, p. 56 |
| Δw_{ki} | k^{th} neural cell i^{th} previous input weight increment value, p. 39 |
| Φ | Radiation-based heat generation rate [J/s], p. 32 |
| α_{ij} | i^{th} component selectivity to j^{th} component respect, p. 20 |
| β | Isobaric thermal expansion coefficient [K^{-1}], p. 32 |
| δ | Local synaptic weight increment gradient, p. 42 |
| ϵ | Viscous heat increment due wall friction [$J/m \cdot s$], p. 32 |
| η | Neural cell learning rate parameter, p. 39 |
| γ | Hyperbolic tangent adjust parameter, p. 35 |
| μ | Fluid/gas dynamic viscosity [$Pa \cdot s$], p. 30 |
| ∇ | Laplace operator, p. 28 |
| ω_i | i^{th} component acentric factor, p. 30 |
| ω_{AB} | The dimensionless collision integral for viscosity, p. 62 |
| ρ^m | Molar density [mol/m^3], p. 30 |
| σ_{AB} | The average molecule diameter from a binary mixture, p. 62 |
| ϵ | Back propagation algorithm <i>momentum</i> constant, p. 43 |
| φ | Activation function notation, p. 34 |
| ζ | Discrete time indicator, p. 40 |

| | |
|----------------|---|
| ζ^{-1} | Discreet time operator, p. 40 |
| a | Ramp angular coefficient, p. 35 |
| a_{mix} | Van der Walls mixture parameter for quadratic approximation, p. 60 |
| b | Logistic equation adjust parameter, p. 35 |
| b_k | k^{th} neural cell bias, p. 36 |
| b_{mix} | Van der Walls mixture parameter for linear approximation, p. 60 |
| c_i | i^{th} component molar concentration [mol/m^3], p. 28 |
| dX_G/dy_G | General driving force, p. 17 |
| e_k | Error value, p. 40 |
| f | General representation of a process variable, p. 56 |
| f_{in}^{i+1} | General representation of an inlet variable value of $i^{th} + 1$ module, p. 56 |
| f_{max} | Axon maximum pulse frequency [Hz], p. 34 |
| f_{out}^i | General representation of an outlet variable value from i^{th} module, p. 56 |
| $g(t)$ | Mean axon pulse frequency for t^{th} time input signal values [Hz], p. 34 |
| g_t | Earth's gravity acceleration [m/s^2], p. 30 |
| h | Membrane thickness [m], p. 20 |
| h^s | Shell side fluid/gas heat transfer coefficient [$J/m^2 \cdot K \cdot s$], p. 33 |
| h^t | Tube side fluid/gas heat transfer coefficient [$J/m^2 \cdot K \cdot s$], p. 33 |
| i | Input signal index, p. 34 |
| j | Previous layer indicator index, p. 42 |
| k | Last layer's neural network response signal total amount, p. 41 |

| | |
|----------------|---|
| k_i | i^{th} compound permeability coefficient [$mol \cdot m/s \cdot m^2 \cdot Pa$], p. 20 |
| k_j | j^{th} component permeability coefficient [$mol \cdot m/s \cdot m^2 \cdot Pa$], p. 20 |
| m | Total neural network hidden layers, p. 41 |
| n | Inputs signal number amount, p. 34 |
| p_i | i^{th} compound partial pressure [Pa], p. 20 |
| r_k | k^{th} cell expected response signal value, p. 39 |
| s | Membrane fiber side average velocity vector [m/s], p. 28 |
| s_r | Fiber molar flow vector radial velocity contribution [m/s], p. 28 |
| s_z | Fiber molar flow vector axial velocity contribution [m/s], p. 28 |
| t | Time variable [s], p. 28 |
| t_p | Axon pulsing period [s], p. 34 |
| t_r | Axon resting period [s], p. 34 |
| u | Shell side molar flow vector [mol/s], p. 28 |
| u_{zi} | i^{th} element shell molar flow vector axial contribution [mol/s], p. 55 |
| v | Tube side molar flow vector [mol/s], p. 28 |
| v_{zi} | i^{th} element tube molar flow vector axial contribution [mol/s], p. 55 |
| w^T | Transposed weight vector, p. 36 |
| w_i | i^{th} input synaptic weight, p. 34 |
| w_{ki}^{new} | k^{th} neural cell i^{th} input weight new value, p. 39 |
| w_{ki}^{old} | k^{th} neural cell i^{th} input weight previous value, p. 39 |
| x | Input signal vector, p. 36 |
| x_i | i^{th} input signal, p. 34 |

| | |
|----------|---|
| x_k | k^{th} neural cell input signal, p. 36 |
| x_{ki} | k^{th} neural cell i^{th} input signal, p. 39 |
| y_k | k^{th} neural cell response signal, p. 36 |
| z | Membrane axial coordinate variable [m], p. 28 |

List of Abbreviations

| | |
|------------------------------------|--|
| <i>Ar</i> | Argon, p. 10 |
| <i>CO₂</i> | Carbon dioxide, p. 1 |
| <i>C₂H₆</i> | Ethane, p. 1 |
| <i>C₃H₈</i> | Propane, p. 1 |
| <i>C₄H₁₀</i> | Butane, p. 1 |
| <i>H₂S</i> | Hydrogen sulfide, p. 1 |
| <i>N₂</i> | Nitrogen, p. 1 |
| <i>N₂</i> | Nitrogen, p. 10 |
| <i>O₂</i> | Oxygen, p. 10 |
| <i>SO₂</i> | Sulphur Dioxide, p. 12 |
| AI | Artificial Intelligence, p. 5 |
| ANN | Artificial Neural Network, p. 4 |
| ANP | Brazilian National Agency of Petroleum, Natural Gas, and Biofuels, p. 11 |
| ASU | Air Separation Unit, p. 25 |
| AdaGrad | Adaptive Gradient Algorithm, p. 44 |
| Adam | Adaptive Moment Estimation, p. 44 |
| <i>CH₄</i> | Methane, p. 1 |
| DT | Digital Twin, p. 22 |
| FPSO | Floating Production Storage and Offloading, p. 21 |
| GHG | Greenhouse Gases, p. 3 |

| | |
|------|---|
| GOR | Gas-Oil Ratio, p. 3 |
| IPCC | Intergovernmental Panel on Climate Change, p. 8 |
| IoT | Internet of Things, p. 23 |
| LMS | Least-Mean-Square, p. 39 |
| LMTD | Logarithmic Mean Temperature Difference, p. 32 |
| LNG | Liquefied Natural Gas, p. 11 |
| MDEA | Methyldiethanolamine, p. 14 |
| MLP | Multi-Layer Perceptron, p. 4, 5 |
| NG | Natural gas, p. 1 |
| NOAA | National Oceanic and Atmospheric Administration, p. 8 |
| O&G | Oil & Gas, p. 3 |
| PZ | Piperazine, p. 13 |
| VLSI | Very-large-scale-integration, p. 38 |

Chapter 1

Introduction

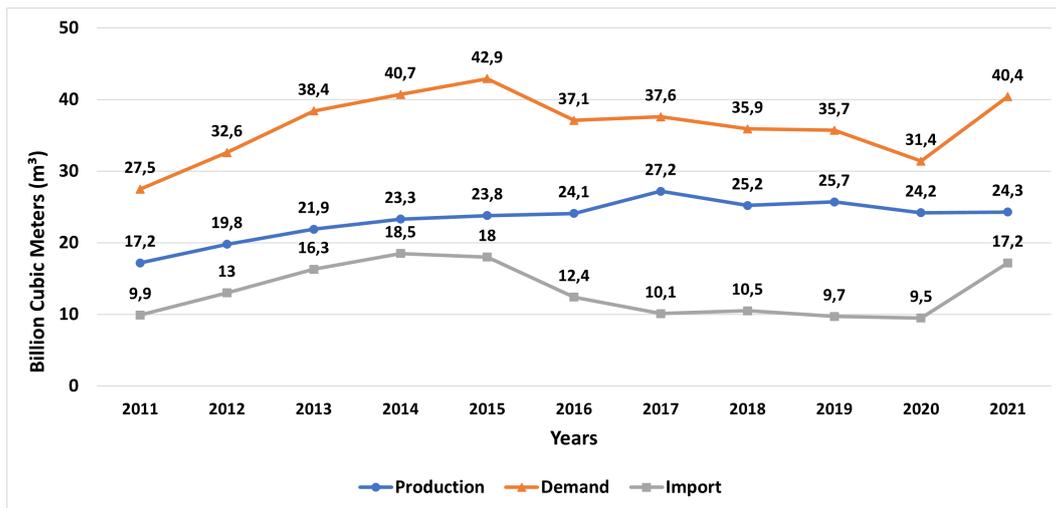
1.1 Natural Gas

Natural Gas (NG) is a fossil fuel composed majorly of light-weighted hydrocarbons such as methane (CH_4), ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}) and some acid gases. Acid gases are compounds that, once in water solution, act like acid elements and lower the solution pH, for instance, carbon dioxide (CO_2), hydrogen sulfide (H_2S), and nitrogen (N_2) also known as sour gases or contaminants (LIMA and GONÇALVES, 2016; PARO, 2005).

Like petroleum, NG is formed through organic material decomposition at high temperatures and pressures over millions of years. It is not only available at offshore reserves but also onshore. It is most commonly found on gas reserves where gas-oil rates are higher than the usual oil reserve. This form is denominated as not-associated gas as it is found in its free form in a gas phase layer. NG can also be found in oil reserves, in which the oil phase is the majority, or NG is dissolved in the oil phase. This form is an associated gas reservoir (FERREIRA, 2006; PARO, 2005).

Nevertheless, NG can also be found onshore inside low-porosity rocks such as shale, low-permeability sandstone, coal, and naturally formed NG hydrates (COLOMER and ALMEIDA, 2015). NG can also be produced through organic decomposition using aerobic or anaerobic bacteria or coal decomposition (VIEIRA *et al.*, 2005). Of all oil-derived fossil fuels, NG is considered the cleaner one. That is because NG composition is mainly composed of a lower carbon chain, so the amount of CO_2 exhausted as a combustion product is lesser than other fossil fuels such as gasoline or diesel.

According to FERREIRA (2006) and PARO (2005), the first Brazilian NG reserve was discovered in 1940 in Bahia. Historical events suggest that NG was first used by the population between 6000 and 2000 BC in the region now known as Iran



Source: adapted from BP (2022)

Figure 1.1: Gap between natural gas production and demand in Brazil.

and discovered only in 1659 in Europe. Since its discovery, Brazilian NG natural reserves have increased. By the end of 2018, Brazil had 0.4 trillion cubic meters of NG proven reserves and 0.2% share of all NG reserves in the world. The amount of NG available in these reserves could last for 15 years if the production rate were maintained stable. However, Brazil's NG production is not auto-sufficient, as it cannot suffice the whole country's growing demand. Therefore, the country still relies on LNG imports (BP, 2019, 2022). Figure 1.1 shows the Brazilian gap between production and demand.

This gap is covered through importation mainly from Bolivia, using Brazil-Bolivia pipelines, and by LNG sea transportation from the US, Trinidad and Tobago, Niger, Qatar, and other countries. These five are the leading Brazil suppliers (BP, 2019, 2022). Despite this notorious gap, Figure 1.1 shows a significant increment in NG production in the last decade. A recent study shows that Brazil is the 27th country that most consumes NG, with average annual demand growth of 5.7% in the last decade (ANP, 2019a). This consumption increment is higher than the optimistic growth predictions from Empresa de Pesquisa Energética (EPE) for whole energy demand until 2030 (2.6% of average annual growth), including NG (EPE, 2018). NG is increasing its participation in the country's energy mix.

This enhancement in production is mainly due to new oil wells found in the Santos Basin and Campos Basin in Brazil's southeast area. Both fields are composed majorly of pre-salt reservoirs. The operation began in 2010 and, since the end of 2017, has been the country's primary source of oil and gas. Nowadays, pre-salt oil corresponds to 65% of Brazil's total production (ANP, 2019b).

Although pre-salt reservoirs hold high quality (higher API degree) and high

volumes of oil, these also come with loads of associated gas. For instance, while the Gas-Oil Ratio (GOR) of a standard offshore oil field is about 150, pre-salt reservoirs reach a GOR of 600 at the Libra field without CO_2 re-injection (ARINELLI *et al.*, 2015). That rate makes exploration even more challenging, which now has to deal with separating these phases, associated with higher pressures from ultra-deep waters and even lower temperatures in Brazil's pre-salt wells (PASQUALETTE *et al.*, 2017).

Besides, the pre-salt gas phase also contains from 50% to 80% molar of CO_2 in its composition, higher than 10% to 30% molar post-salt offshore wells. Apart from being increasingly required, NG cannot be commercialized as it is removed from petroleum reservoirs (crude form). It needs to be treated through a gas separation process, which is of primary economic importance in the Oil & Gas (O&G) industry field. Without it, NG could not be commercialized as fuel and probably would be re-injected into an oil reservoir or sold for a small price to gas refineries. The presence of acid gases, such as CO_2 and H_2S , reduces its calorific power, thus quality and market price (ATCHARIYAWUT *et al.*, 2007).

In addition, those components also hinder process pipelines. That is because of their corrosion properties, not to mention that they are also precursors to hydrate formation (a low porosity solid structure, which is a significant risk to transmission lines as it can block stream flow completely) (HAMMERSCHMIDT, 1934; POBEREZHNY *et al.*, 2019). Furthermore, the gas separation process prevents more CO_2 is released into the atmosphere since this component leads to a list of Greenhouse Gases (GHG) that most contribute to its intensification (IPCC, 2014). Because of that importance, countless studies are found on enhancing CO_2 and H_2S separation. However, until this work's bibliography revisions, no study with this purpose was found using industry 4.0 tools such as machine learning, AI, or big data. That gap in the literature triggered the primary motivation for this work. To propose one or more of those tools to help the oil gas industry achieve its digital maturity.

1.2 Motivation

Due to the high Gas-Oil Ratio (GOR) and high amount of CO_2 in the pre-salt fields, new approaches to gas treatments are mandatory. Once offshore gas processing needs to be redimensioned according to reservoir composition, this includes several gas separator modules, which is a significant concern due to physical space limitations at offshore platforms (BELTRAO *et al.*, 2009).

The Natural Gas deacidification process is an essential step of offshore oil processing. Its acid gases, such as carbon dioxide (CO_2) and hydrogen sulfide (H_2S),

reduce natural gas heating power. That quality loss also means income loss for the industry. Besides, both compounds can cause further problems as they contribute to hydrate formation in the transmission lines, which can cause not only pipeline damage (MAKOGON, 2010), due to hydrate obstruction and pressure increase but also promote pipeline corrosion (POBEREZHNY *et al.*, 2017). Furthermore, CO₂ is pressurized back into the reservoir to improve oil extraction, boosting production and generating more profit.

Therefore, some models once used to simulate, control and predict GN separation must be updated. Those models that predict process transient and steady-state operation are essential for advanced process control, real-time optimization, and risk mitigation. Until recently, these models were purely phenomenological and based on thermodynamic, physical, and chemical properties. These models seek to respect the physical laws and result in detailed comprehension and how each variable interacts. However, it has some restraints when facing nonlinearity problems, mainly predicting and analyzing mass transportation through different membrane fibers in a gas separation process (ASGHARI *et al.*, 2018).

Machine learning-based models, though, could deal with this nonlinearity with less modeling effort. These models use industry databases and attempt to learn the relation between input and output variables using linear or nonlinear regression. They are also called "black box" models. The most common machine learning algorithm is the Artificial Neural Network (ANN) (ASGHARI *et al.*, 2018). However, as an empirical model, it is also uncertain when describing a complex process. Furthermore, it does not take into account any thermodynamic or physical laws. Thus, combining both models into a hybrid one can result in a powerful approach to respect the physical laws and deal with nonlinearities faster than phenomenological-based models, which enables more accurate real-time control and optimization algorithms (ASGHARI *et al.*, 2018; QUIZA *et al.*, 2012).

Some authors already studied machine learning advantages on O&G segments, such as: AHARI *et al.* (2011) used a Multi-Layer Perceptron (MLP) to model the catalytic oxidative coupling of methane; BALABIN *et al.* (2015) tested some methods to predict biodiesel properties, and ANN was one of them; PEER *et al.* (2008) used an ANN to predict module gas separation permeability and the required membrane area to process different natural gas compositions. However, no record of a hybrid model on a gas separation subject was found in the literature.

The use of machine learning, big data, AI, and digital twins are attributes of the new digital era in the industry. Many manufacturing segments are turning their attention to this new era to make their business more profitable and reliable. The O&G segment has a modest digital maturity achievement, with a grade of

4.68 points out of 10 (KANE *et al.*, 2015). That score suggests plenty of room for the O&G industry to explore this digital transformation.

With that motivation, this work intends to create a digital twin model for an offshore gas separation unit using a hybrid phenomenological-artificial model to predict mass, momentum, and energy transfer inside each membrane module with an Artificial Neural Network (ANN) to learn and predict permeability factor variations from each compound. That way, it will rely on thermodynamic, physical, and chemical properties while dealing with permeation nonlinearities faster. Because of that, faster process prediction is expected, compared with purely phenomenological, and also serves as a precursor for real-time optimization algorithms development and predictive control.

1.3 Objectives

The main objective of this work is to develop a digital twin for a membrane-driven natural gas deacidification unit to deal with processes nonlinearities and predict process states more efficiently.

To accomplish those objectives, the following specific objectives are persecuted.

- Develop a phenomenological model to describe each membrane module's mass, momentum, and heat transfer.
- Develop and evaluate a Multi-Layer Perceptron network to predict compound permeability.
- Determine the best hybridization structure to predict the output with significant accuracy and efficiency to be implemented in real-time.
- Create and evaluate the performance of a digital twin with the best alternative model developed for the third objective.

1.4 Structure

Chapter 2 introduces a background and literature review regarding the NG deacidification process. It begins with an overview of the NG separation process in Section 2.1, presenting the fundamentals, objectives, and discussion regarding the commonly used method and their benefits, including cryogenic, absorption, and membrane-based separation processes. A brief overview regarding the concepts of a digital twin is presented in Section 2.2. Then, Section 2.3 shows

the phenomenological background for mathematically modeling these membrane modules regarding mass, energy, and momentum conservation. The section 2.4 presents a background regarding the MLP network structures and how they can be integrated into this work. Finally, Section 2.5 points out this work's innovative level as it presents some literature published regarding these topics mentioned above and how this work can fill some of the science's gaps.

Chapter 3 describes a case study regarding the first objective, where a phenomenological-based model is proposed and tested with the results available in the literature. It begins in Section 3.1, making an overview of the process that will serve as the base for the digital-twin development. Then, in Section 3.2 some decision-making regarding each membrane module's hypothesis, material, and dimensions are carried out. Also, in this section, the mass, momentum, energy, and thermodynamic models are proposed. Section 3.3 adds an assertiveness comparison between the proposed model with the ones already consolidated in the literature. Section 3.4 summarizes the knowledge gathered regarding the phenomenological-based model and exposes the decision-making regarding its components.

Chapter 4 presents a case study regarding achieving the second specific objective identified in this work, where a neural network structure is proposed and studied deeply to verify its accuracy and performance in regard to the prediction of the permeability factors and also is used to determine the permeability factor as input for chapter 3 phenomenological-based model. Section 4.1 gives an overview of what variables will be used, how the search for the best MLP structure will be done, and how the hybridization will undergo. Section 4.2 presents the data treatment process and the search for the best MLP configuration to produce the finest adjustment. Section 4.3 adds an assertiveness comparison between the results from the model of the Chapter 3 and the Neural Network-proposed model developed and presents the results of a comparison between the results from the model of the previous Chapter 3 and the hybridization model proposed. Section 4.4 summarizes the knowledge gathered regarding the Hybrid Network-Phenomenological model and exposes the decision-making regarding its usability.

Chapter 5 intends to address the last part of the third and fourth specific objectives proposed in this work, where the physic-informed neural network model is proposed. Section 5.1 overviews how the best structure will be searched. Section 5.2 presents the methodology used to get to the best Network configuration. Section 5.3 adds an assertiveness comparison between the results from the model of the Chapter 3 and the network model proposed. Section 5.4 summarizes the knowledge gathered regarding the network model and exposes the decision-

making regarding its usability.

Chapter 6 presents a summary of the developments and findings from this work, proposes the future works that could be an outcome of those findings, and concludes the dissertation.

Appendix A presents the membrane diffusion coefficient calculation method used. Appendix B shows the detailed backpropagation training algorithm. Appendix C shows a scaled-up figure from the real process used to create the digital twin. Appendix D presents some thermodynamic parameters calculation strategy. Finally, Appendix E presents the phenomenologic-based model finite volume discretization strategy.

Chapter 2

Background and Literature Review

2.1 Gas Separation Processes

The gas separation process for CO_2 removal is in high demand in the oil and gas industry. Whether in the purification of natural gas or biogas or to reduce the environmental impacts caused by releasing GHG gas into the atmosphere. Besides, the presence of carbon dioxide in natural gas reduces its calorific value, thus acting as an impurity and a highly corrosive agent to transport and storage structures.

In addition, CO_2 is the gas that most causes the greenhouse effect according to the National Oceanic and Atmospheric Administration (NOAA) database presented in the fifth assessment report of the Intergovernmental Panel on Climate Change (IPCC) that is summarized in Table 2.1. For that reason, this process has become of significant importance for the industry as it allows high-value-added products with higher purity, prevents GHG emission, and enhances productivity by removing process contaminants (ATCHARIYAWUT *et al.*, 2007; IPCC, 2021).

Table 2.1 also shows that, besides CO_2 not having much energy reflective capacity, it is the most contributor to the greenhouse effect due to its abundance in the atmosphere, and so to the global warming (IPCC, 2021). If an offshore platform cannot separate all CO_2 , NG must be treated in an onshore gas process plant before commercialization. Not treating that gas offshore would cause the greenhouse effect to worsen as, once onshore, excessive CO_2 can no longer be re-injected inside a gas reservoir. Also, the presence of carbon dioxide and sulfidric acid increases the probability of hydrate formation (ATCHARIYAWUT *et al.*, 2007; POBEREZHZHNY *et al.*, 2019), which would imply constant equipment maintenance, as mentioned in the last Section. Besides, once the untreated gas is onshore, the options are to purify and sell it to the gasified beverages industry or to liberate it into the atmosphere. Eventually, this excessive CO_2 will be part of the atmosphere composition.

Table 2.1: Atmospheric concentration and energy reflective capacity from main GHG.

| GHG | Atmospheric Concentration (ppt) ¹ | Total energy Reflection Contribution ² (W/m ²) |
|---------------------|--|---|
| CO ₂ | 410 ± 0.2 (ppm) | 2.156 ± 0.259 |
| CH ₄ | 1867 ± 1 (ppb) | 0.544 ± 0.109 |
| N ₂ O | 332 ± 0.2 (ppb) | 0.208 ± 0.031 |
| CFC | 822 ± 0.8 | 0.28 |
| HCFC | 293.2 ± 0.7 | 0.06 |
| HFC | 212.3 ± 0.8 | 0.03 |
| Others ³ | 178.71 | 0.025 |

Source: adapted from IPCC (2021)

¹ Data from 2019.

² Total heat reflection capabilities; or the cumulative greenhouse effect contribution.

³ Joint contributions from: sulphur hexafluoride (SF₆), tetrafluormethane (CF₄), esafluoroethane (C₂F₆) and carbon tetrachloride (CCl₄).

Studies on gas separation processes began more than a century and a half ago when GRAHAM (1866) first studied the differences between distinct compounds permeating through a membrane film. Since then, many other authors have proposed various approaches to address gas separation processes. Some authors researched gas absorption using monoethanolamine (MEA) and diethanolamine (DEA) or their derivatives process, such as AL-BAGHLI *et al.* (2001), SIPÖCZ *et al.* (2011) and TAHERI *et al.* (2016). Others, for instance, focused their research on the membrane-based gas separation process, for example, MARZOUK *et al.* (2012), CHU *et al.* (2019) and HASEGAWA *et al.* (2017). Some other authors yet, studied gas separation through cryogenics, when compounds are separated by differences in dew points (ANSARINASAB *et al.*, 2017; DEBNATH *et al.*, 2019; EBRAHIMI and ZIABASHARHAGH, 2017).

These three separation methods are the most commonly cited in the bibliography as efficient and economically feasible depending on what type of gas separation is dealt with, its effluent purity grade, or treated gas destination. Each process has its vantages and disadvantages. The following sections discuss where each one is recommended to be used.

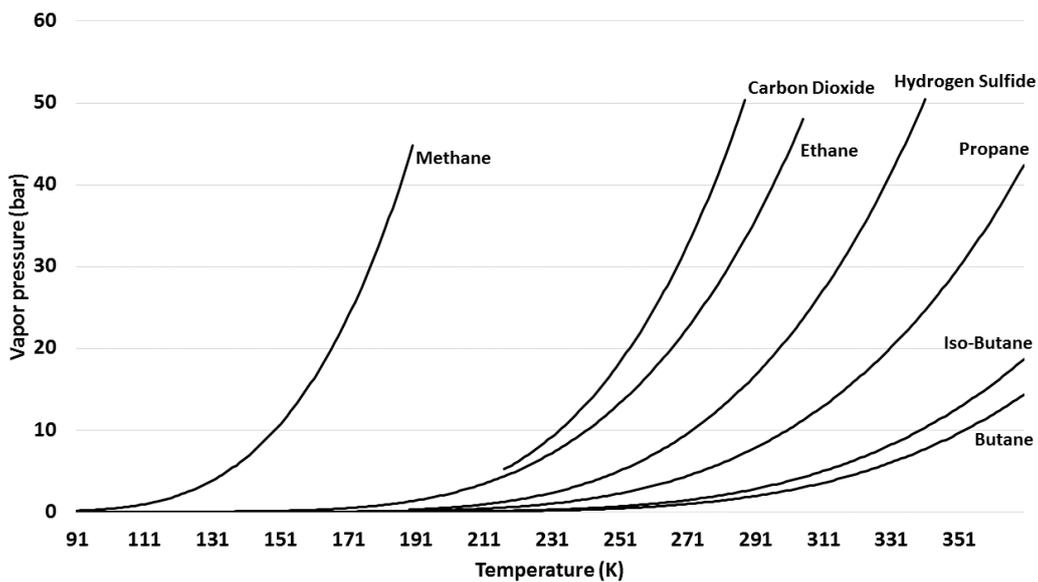
2.1.1 Cryogenic Method

Cryogenic separation involves lowering stream temperature until the desired compound reaches its dew point so it can be separated from the rest of the stream. It demands much energy as gases usually have negative temperature dew points. This process requires a large amount of physical space to fit all required equip-

ment, as it depends on robust cooling systems and columns (KNAPIK *et al.*, 2018). Although, depending on compounds and stream properties, this process can produce streams purities near 100% and is commonly used to produce industrial oxygen (O_2), nitrogen (N_2), and argon (Ar) (SMITH and KLOSEK, 2001).

Besides, when looking at NG component's vapor pressure, as shown in Figure 2.1, this gives a false impression that cryogenic separation would suffice the project's objective. However, this method faces some issues when considering a gas mixture due to CO_2 and NG mixture properties. For instance, $CO_2 - CH_4$ and $CO_2 - C_2H_6$ interactions make liquid CO_2 phase to solidify as it concentrates, hindering productivity and phase separation. $CO_2 - C_2H_6$ also produces an azeotrope at $-53\text{ }^\circ\text{C}$ and 0.6 molar fraction of CO_2 , which also compromises process' efficiency (KIDNAY and PARRISH, 2006). HOLMES *et al.* (1983) proposed feeding NG mixed with heavy hydrocarbon streams to avoid CO_2 solidification while in a distillation column. However, this process still relies on plenty of physical space available and a great amount of energy.

Figure 2.1 also shows that the main compounds' vapor pressure tends to pull away from each other, making the separation process easier for high pressures and high temperatures.



Source: adapted from NIST (2019)

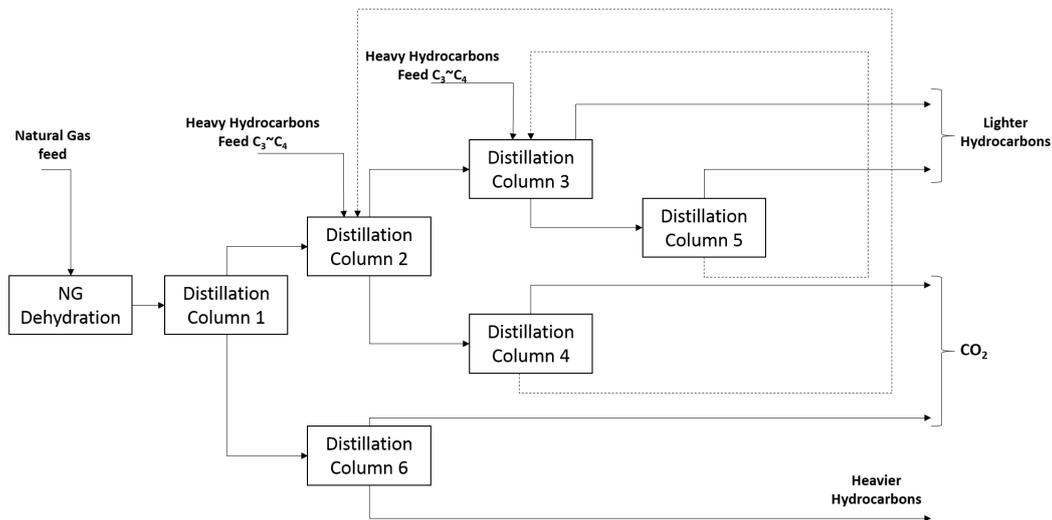
Figure 2.1: Vapor pressure of main NG components, using Antoine equation.

A recent study regarding energetic efficiency made by KNAPIK *et al.* (2018) showed that it is possible to integrate Liquefied Natural Gas (LNG) re-gasification and cryogenic air separation units to make NG stream reach CO_2 dew point and

that way, reduce the required energy amount. In that work, 83.7% of CO_2 in a flue gas stream was removed using this process, which is an expressive positive result.

According to the Brazilian National Agency of Petroleum, Natural Gas, and Biofuels (ANP), to be commercialized, NG needs to match certain specifications regarding its calorific power, Wobbe index, minimum hydrocarbon composition, complies with less than 3% molar of CO_2 specification, and a maximum of 10ppm of H_2S (ANP, 2008).

Furthermore, offshore platform physical space availability is scarce. When taken into consideration, it makes this cryogenic method unfeasible in practical terms as offshore platforms do not rely on much free space for cryogenic machinery (MAQSOOD *et al.*, 2014). Figure 2.2 presents a schematic block diagram of a sour gas cryogenic separation standard process from NG using Ryan/Holmes process proposal (HOLMES *et al.*, 1983), which requires a significant amount of distillation columns and, therefore, ample space.



Source: adapted from MAQSOOD *et al.* (2014)

Figure 2.2: NG cryogenic separation block diagram.

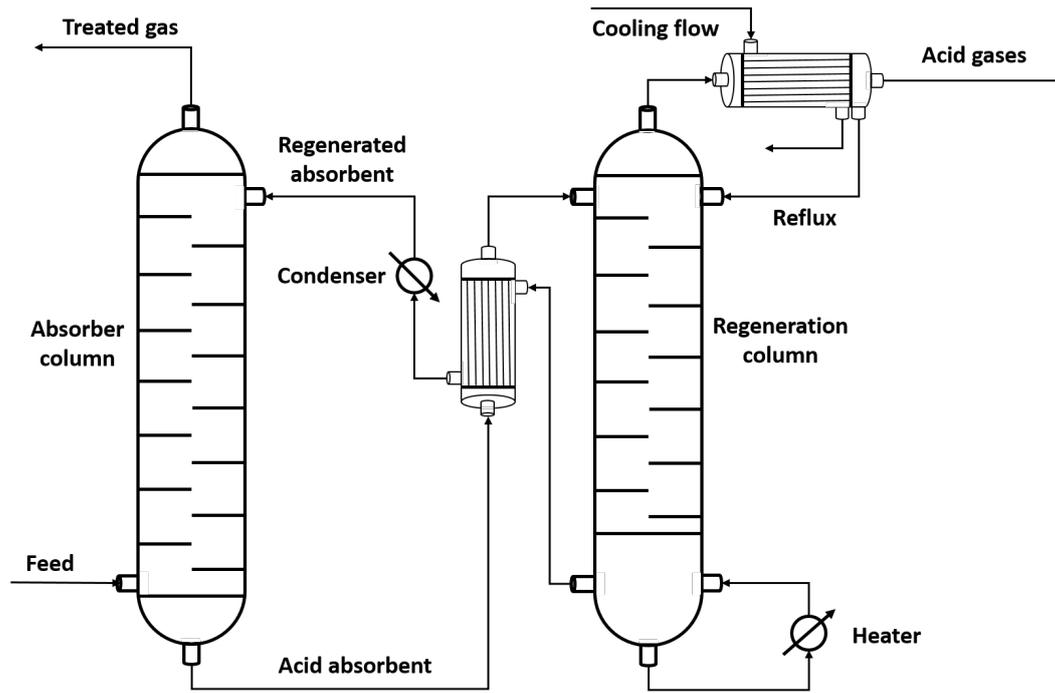
Although this process seems simple to understand, its operation is more complex. Distillation columns require high control systems and operate in a short gap of temperature and pressure. To optimize efficiency, as suggested by MAQSOOD *et al.* (2014), six columns would be needed, which implies six complex control systems, making the process more prone to failure. Nevertheless, its efficiency regarding the purity of effluents is unquestionable.

2.1.2 Amine Absorption Method

Gas separation by absorption method uses liquid or solid compounds with a high affinity to one or a few compounds and removes them from a vapor stream. The most used process is the gas-liquid based, which uses an aqueous solution (usually an amine-based compound) to remove undesirable compounds. This section discusses the commonly used amine absorption process, its benefits, its disadvantages, and why it is more challenging to integrate it into an offshore platform.

Amine absorption technology is a well-known gas separation process in many industry segments. The first process recorded by literature is a patent made by Bottoms in the late 1930s, which has two towers (one for absorption and another for regeneration) for CO_2 , H_2S , and sulfur dioxide (SO_2) separation using an absorbent agent. Amine compounds are explicitly mentioned as good options for removing acid gases. The mathematical model proposed by BOTTOMS (1930), for instance, is still used nowadays for process simulations and studies regarding new absorbents or optimizing their regeneration process (XUE *et al.*, 2017).

Figure 2.3 presents a simple absorption process. The acid stream enters at the bottom of the first column while the absorbent is loaded at the top. Both streams contact each other on a counter-current pattern to maintain a high gradient concentration between both phases (gas and liquid). After that, the sweetened stream gets out of the first column by its top while the absorbent, full of acid compounds, is pre-heated and enters the second column by its top. On the bottom of this second column is a heater that regenerates the absorbent. Then, the absorbent is cooled and goes back to the first column, and starts over the process. A stream rich in acid gases is collected on the second-column condenser at the top.



Source: developed by the author.

Figure 2.3: Generic steam deacidification process using absorbents.

The absorption method consists of using a solvent with a great affinity to non-desired compounds in a process stream. On NG process plants, either onshore or offshore, to remove CO_2 and H_2S from NG stream, carbon-amine compounds are well known to be reliable and efficient for reaching these goals (XUE *et al.*, 2017).

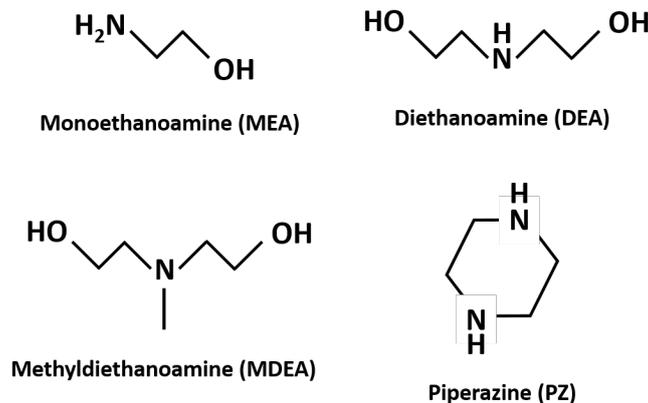
The most common absorbents are aqueous solutions of Monoethanolamine (MEA) and Diethanolamine (DEA) due to their higher benefit-cost ratio for the industry. According to XUE *et al.* (2017), of the two compounds, MEA has the best cost-efficient ratio, providing a good absorption efficiency with lower regeneration costs than DEA. Also, HUERTAS *et al.* (2015) studied the absorption ratio of MEA and concluded that aqueous solutions of 2.5 wt% have better performance regarding CO_2 removal, reaching an average of 580 g CO_2 /kg MEA approximated.

Many other absorbents are also studied in the literature, with efficiencies ranging from 85% up to 99% of CO_2 removal. For instance, Piperazine (PZ), also an amine-based compound studied by GAO *et al.* (2019), ROCHELLE *et al.* (2011) and FREEMAN *et al.* (2010), which resulted in a 99% CO_2 removal according to GAO *et al.* (2019) work, PZ can also support temperatures up to 150 °C without thermal degradation, has a better resistance against oxidation, and less volatility than MEA.

Another amine-based compound also widely studied as an alternative to MEA and DEA is methyl diethanolamine (MDEA), a tertiary amine compound which,

according to KOHL and NIELSEN (1997), has lower volatility when compared to MEA and DEA, which allows being used in higher concentrations without losses, for instance, 60 wt%. MDEA also demands lower energy to react with CO_2 and H_2S , which means that the system can operate at a lower temperature. Besides, MDEA is an almost non-corrosive compound with a higher absorption capacity. Its primary disadvantage, though, is that, due to absorption energy low rates, this process occurs at an equally lower rate, so MDEA is rare to be used alone in a process. It is usually combined with other amines, such as MEA, DEA, or PZ, to increase absorption speed.

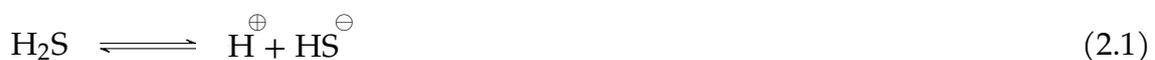
Figure 2.4 illustrates the molecular structure of these main compounds. It is easy to identify that MEA, DEA, and PZ are not tertiary amines and have at least one hydrogen bonded to nitrogen. That bound allows CO_2 to react straight to the amine compound forming carbamates, while H_2S is ionized and carried away in the aqueous solution by alkanolamine protonation. Carbamates are stable compounds that carry away the CO_2 (KOHL and NIELSEN, 1997). Some authors propose a formation of an intermediary compound yet before carbamate named zwitterion (CAPLOW, 1968; DANCKWERTS, 1979).

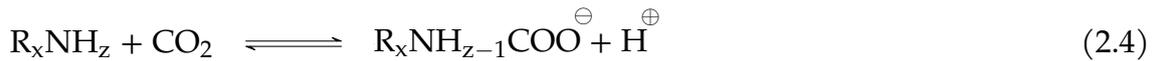


Source: developed by the author.

Figure 2.4: Main amines compounds molecular structures.

A generic absorption mechanism is showed below where $0 < x \leq 3$, $0 \leq y \leq 2$, and $0 < z$. Note that for primary and secondary amines, CO_2 can bond directly to the absorber creating a carbamate. However, this does not occur for tertiary amines, such as MDEA. In that case, all mechanism from reaction 2.1 to reaction 2.3 happens naturally, but not on reaction 2.4 as z cannot be zero.





For that reason, tertiary amines rely on H_2S and CO_2 ionization on an aqueous solution to be carried away dissolved. This makes the absorption process slow as deionization is not as fast as if carbamate was formed. That is why primary or secondary amines are added to tertiary amine solutions to improve process speed and thus its efficiency (CAPLOW, 1968; KOHL and NIELSEN, 1997).

Furthermore, some disadvantages make the absorption process unfeasible to use in an offshore plant. According to GABELMAN and HWANG (1999), these processes require a high investment, and operation costs are also elevated. Much physical space is needed to fit the adsorption towers, heat exchangers, condensers, and heaters. Also, they are very limited in operational conditions because each tower cannot overflow, drag the aqueous solution, or form foam. If any of these occurs, production will be compromised as well as processes efficiency and, consequently, a significant loss of money will result for the offshore industry.

The next topic analyzes the gas separation process by selective membranes, a process to avoid the disadvantages of amine scrubbing as an alternative to CO_2 capture with conventional methods. As discussed in the next section, membrane permeation separation emerges as a less costly alternative regarding physical space and implementation and promotes optimum performance.

2.1.3 Membrane-based Method

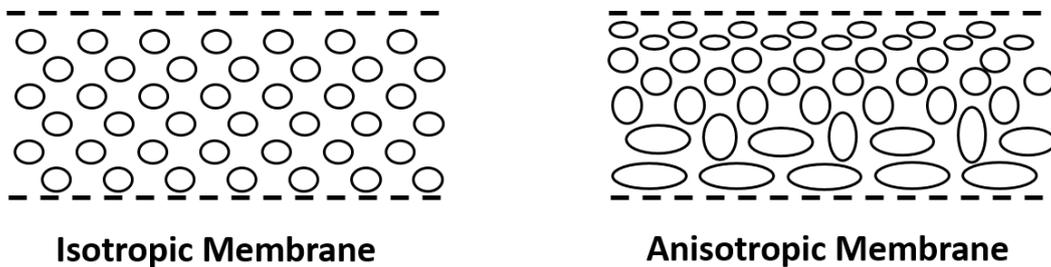
Membranes are porous film-shaped structures made usually of different polymers with selective capabilities. That means it allows certain compounds to pass through its porous while blocking others when in contact with any stream. In other words, membranes act as filters (BAKER, 2004).

The first identification of membrane properties regarding selectivity remounts the year of 1861 when Thomas Graham published his article identifying that a French letter paper sheet embedded with a film of a jelly made of starch could

separate sugar from gum in an aqueous solution (GRAHAM, 1861). Five years later, this same author studied gas penetration into a vulcanized rubber and evaluated its selectivity properties, marking the first paper produced to identify and explore these structures' properties (GRAHAM, 1866).

With the development of new structures and technological advances, membranes are widely used in several areas of operation, from simple filtration and industrial separation to medical applications such as artificial kidneys and lung structures. This is because the pore diameter of a membrane film can vary from $100\ \mu m$ at standard filtration to $3\ \text{\AA}$ as in reverse osmosis processes (BAKER, 2004).

Membranes are characterized according to the type of structure: symmetric (isotropic) or asymmetric (anisotropic). Isotropic membranes have a homogeneous or quasi-homogeneous distribution of pores with similar diameters. Examples of isotropic membranes are micro-porous, dense (nonporous), or electrically charged. While anisotropic ones have heterogeneous structures, often composed of layers made of different materials, but with a high mass diffusive, promoting high desired substrate removal flow. The market widely chooses them for their high performance in terms of efficiency and productivity (BAKER, 2004; MULDER, 1996). Figure 2.5 illustrates the difference between isotropic and anisotropic. Note that isotropic membranes have symmetric porous schematics, while anisotropic membranes do not.



Source: developed by the author.

Figure 2.5: Structure differences between isotropic and anisotropic porous disposition.

This difference in material disposition inside the membrane gives it one of its best properties: selectivity. That is because compounds larger than the largest membrane pore do not percolate through the fibers, acting like a compound filter, while other compounds smaller than the largest pore are partially separated. Also, materials with a better affinity to one compound than the others influence separation processes.

Membrane module development for the separation process is flexible enough

to adapt each film according to the process it will be used. It can be thick or thin and molded in any shape: flat, cylindrical, or spiral-wound-shaped. It can also be electrically charged or neutral and even solid or liquid. That includes metal, polymer, zeolites, and glassy, among others. The transport method can be active or passive, and if passive, it is proportionally driven by concentration, temperature, pressure, or even electric potential gradient (MULDER, 1996; PANDEY and CHAUHAN, 2001). Mathematically speaking, this permeation dependence can be described according to Equation 2.5, where a flux of mass, energy, or momentum (J_G) is proportional to a driving force (dX_G/dy_G) weighted by a coefficient called phenomenological coefficient (A_G). X_G is the state variable which the flux evaluation is needed and y_G the physical space which that occurs (MULDER, 1996).

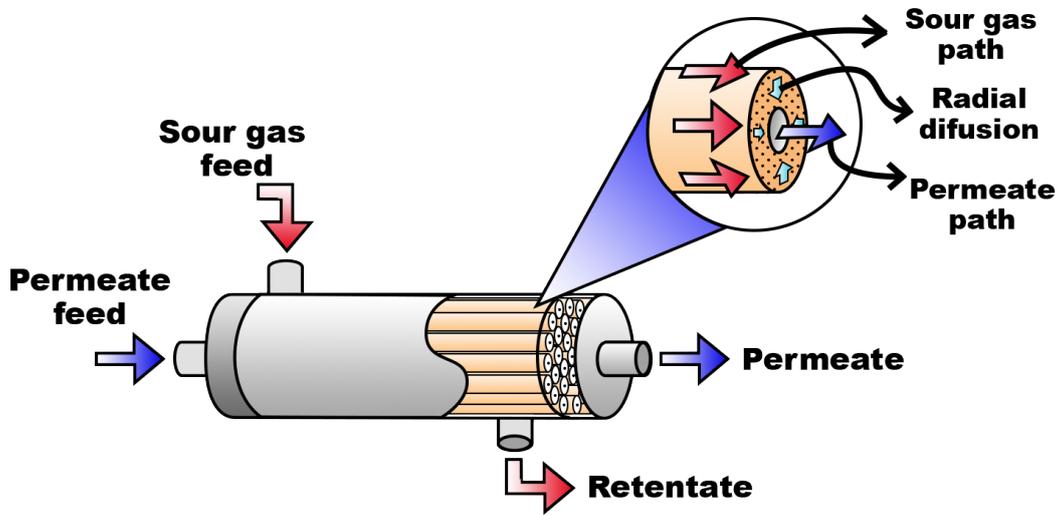
$$J_G = -A_G \frac{dX_G}{dy_G} \quad (2.5)$$

Regarding gas separation for natural gas, cellulose acetate or polyimide membranes are generally used in these processes (HE and HÄGG, 2011). However, several studies of membranes based on carbon fibers and inorganic membranes based on zeolites are gaining notoriety among researchers (CHU *et al.*, 2019; HASEGAWA *et al.*, 2017; HE and HÄGG, 2011). Commercially, only cellulose acetate and polyamides are widely used in several companies worldwide. Market leaders are Honeywell UOP with Separex™ and Cynara-NATCO with cellulose triacetate technology (CHU *et al.*, 2019).

Using membranes is a less costly alternative regarding physical space, implementation, and optimum performance. However, this process is subject to CH_4 load loss, does not have optimal performance for low feed pressure, and is not an efficient solution for the removal of H_2S (BHIDE *et al.*, 1998). Typically, membranes used for gas separation have high density, tiny pore diameter, and anisotropic structure to increase systems productivity (BAKER, 2004).

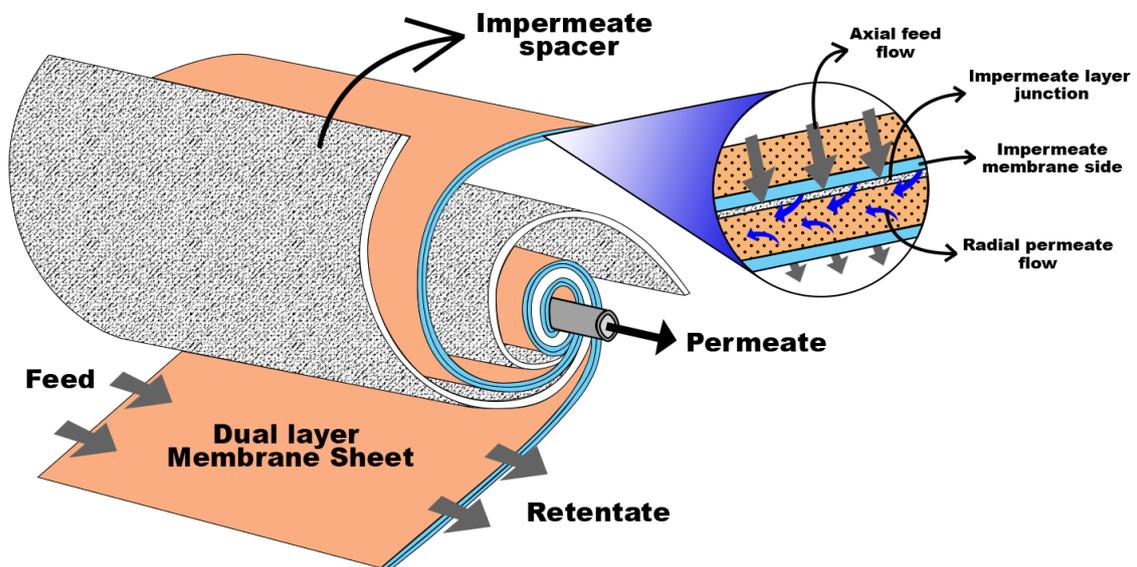
However, not only materials make a difference regarding productivity. The gaseous separation systems structure greatly influences systems performance, especially when productivity is the leading business driving force. Therefore, an optimized structure must be used to promote an excellent benefit-cost ratio. Literature reports that structures that used to be installed on offshore gas treatment are generally hollow-fiber or spiral wound structures. These module's characteristics maximize the contact area between the feed stream and membranes in a short physical space and allow easy maintenance and operation (CHU *et al.*,

2019; HE and HÄGG, 2011). Figure 2.6 and Figure 2.7 show the characteristics of hollow-fiber membranes and spiral-wound modules, based on OKUBO *et al.* (2018) patent, respectively.



Source: developed by the author.

Figure 2.6: Structure of a hollow-fiber module.



Source: developed by the author.

Figure 2.7: Structure of a spiral wound module.

Hollow-fiber membrane modules presented in Figure 2.6 present a structure quite similar to a heat exchanger, though the tubes are replaced by hollow

cylindrical-shaped membranes instead. The permeation process is also like energy transport, as material permeation happens radially from the membrane's external layer toward its internal cavity. Hollow-fiber membranes have an advantage over spiral-wound as each cylindrical structure can be easily replaced by another, if it stops operating or its lifecycle ends.

Spiral-wound membranes, though, as shown in Figure 2.7, have different approaches regarding the structure. Unlike hollow fiber, a spiral wound consists of a few dual-layer membrane sheets wrapped in a hollow tube between spacers. There is a path in which the permeate flows towards the hollow metal tube of each membrane sheet layer, where it leaves the module. Mass transfer is also radial. However, sour gas feed is made in axial form from module entrance to exit.

For that reason, regarding this process's physical effects, the diffusion of compounds on dense and anisotropic membranes has to be done by a driving force that promotes compound passage from one membrane side to another, selectivity though is given by membrane type or affinity to compounds as already discussed before. In the case of diffusion, these driving forces are the presence of gradients, that is, pressure, concentration, or chemical potential differences between the feed stream and permeate stream.

Regarding gas separation, concentration differences, or more specifically, partial pressure differences, due to the gas phase, are the driving force that makes compound permeation. Normally a gas separation happens at high pressures and flows by a very selective surface to the compounds whose objective is to separate (BAKER, 2004).

With the knowledge of partial pressure gradient between membrane sides, it is not difficult to recognize that another important deterministic factor for process performance is the permeate flow pressure. This stream will pass through the membrane removing compounds that have diffused through its pores to an appropriate treatment unit. This flow may be arranged in two different ways: in the same direction as the feed stream (co-current process), where partial pressure gradient becomes lower as sour gas flows inside each module, or it may be crossing in the opposite direction from the feed flow (counter-current process), where partial pressure gradient increases as feed percolate through membrane length (BAKER, 2004; CHU *et al.*, 2019; GABELMAN and HWANG, 1999; HE and HÄGG, 2011; MULDER, 1996; PAN, 1986; XUE *et al.*, 2017).

For anisotropic and dense membranes, Equation 2.5 changes to the following Equation 2.6 form, where k_i is the permeability of component i through a membrane sheet thickness (h) under a gradient of partial pressure (p_i).

$$J_i = -k_i \frac{dp_i}{dh} \quad (2.6)$$

Table 2.2: selectivity of different membrane compounds.

| Polymer | P_{CO_2}/P_{CH_4} selectivity |
|---------------------------|--|
| Polytrimethylsilylpropane | 2.0 |
| Silicone Rubber | 3.4 |
| Natural Rubber | 4.6 |
| Polystyrene | 8.5 |
| Nylon 6 | 11.2 |
| Poly(Vinyl Chloride) | 15.1 |
| Polycarbonate | 26.7 |
| Polysulfone | 30.0 |
| Polyethyleneterephthalate | 31.6 |
| Cellulose acetate | 31.0 |
| Poly(ether imide) | 45.0 |
| Poly(ether sulfone) | 50.0 |
| Polyimide | 64.0 |

Source: MULDER (1996).

Equation 2.6 development can be found in Appendix A.

Regarding permeability, it is good to punctuate this parameter as the most important one when considering gas separation using membranes. That is because nonporous dense membranes use selectivity instead of pore diameter as a barrier for larger compounds. Selectivity is another parameter derived from permeability relations intrinsic to each membrane's material. Selectivity denotes how an compound can cross a membrane sheet, preferably more than others. In other words, it is a ratio of the compound's permeability defined in Equation 2.7 (AARON and TSOURIS, 2005; MULDER, 1996).

$$\alpha_P = \frac{k_i}{k_j} \quad (2.7)$$

To illustrate the importance of knowing the selectivity of a membrane, Table 2.2 shows how the CO_2/CH_4 selectivity can change regarding its material.

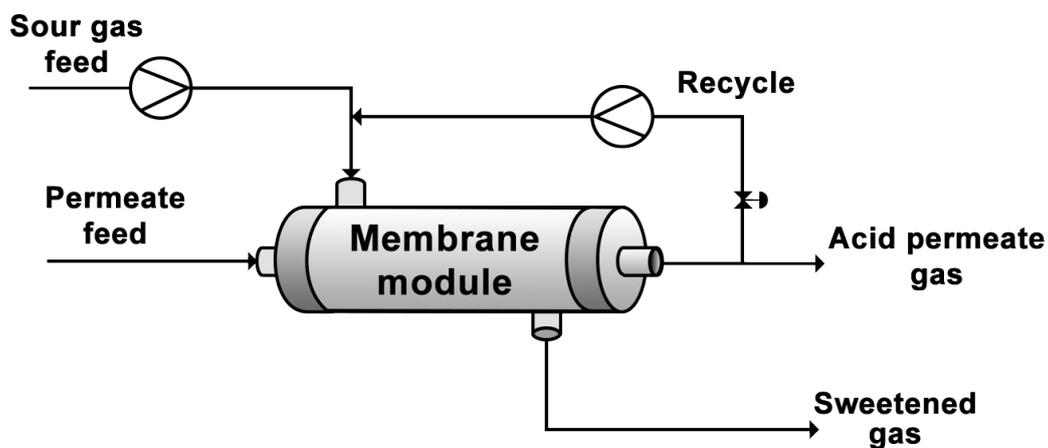
One significant advantage for offshore platforms and Floating Production Storage and Offloading (FPSO) ship regarding the membrane-based process is their compact structure. Unlike the other processes mentioned before, membrane separation requires not one module but a series of modules to achieve NG selling specifications. Still, because of its compact structure, it only needs a little physical space on a platform or an FPSO (AARON and TSOURIS, 2005; DALANE *et al.*, 2017). Another significant advantage is its simplicity of operation. Just a few valves and compressors complete the scheme and are used to control feed pressure to ensure

Table 2.3: Separex UOP membrane benchmarks around the globe.

| Location | CO ₂ feed molar % | CO ₂ production molar % |
|-----------------------|------------------------------|------------------------------------|
| Kandanwari - Pakistan | 12 | <3 |
| Quadipur - Pakistan | 6.5 | <2 |
| Mexico | 93 | <5 |
| Salam - Egypt | 12 | <3 |
| Tarek - Egypt | 6 | <3 |
| Texas - USA | 60 | Not specified |
| Indonesia | 40 | 20 |
| Thailand | 34 | 12 |
| Malaysia | 45 | 6 |

Source: CNOP *et al.* (2007).

that the pressure gradient does not go higher or lower than specified to operate. In general, it does not need heaters, condensers, or a robust temperature control unit, as shown in Figure 2.8 in a simple process design structure. In specific cases, it requires special treatment such as feed temperature control (ECHT and SINGH, 2008).



Source: developed by the author.

Figure 2.8: Gas separation process scheme.

Furthermore, membrane-based separation can achieve excellent efficiency in the field, as shown in Table 2.3 regarding a case study using UOP Separex modules around the globe. These numbers show that membrane-based separator modules can be flexible enough to deal with high concentrations of CO₂ and meet international natural gas commercialization standards.

Another advantage is the simplicity of replacing the membrane sheet when

needed, rather than if it is blown up or at the end of its lifespan, without stopping the whole process. Also, it does not need to be regenerated unless it is clogged by water or another impurity that is not treated before (KIDNAY and PARRISH, 2006).

This brings forth a few disadvantages of membrane-based gas separation. The first one is that water must be removed from the feed as it can clog or even dissolve a few membrane compositions, so crude NG must pass through a molecular sieve before separation to address this issue (KIDNAY and PARRISH, 2006). Besides, one module alone cannot reach the regulation's specifications. For that reason, a series/parallel structure is needed. Furthermore, the permeability/selectivity relation is inversely proportional, which means high selective membranes have low permeability while higher permeable ones do not have good selectivity (AARON and TSOURIS, 2005).

In addition, some polymers may be severely affected by high pressure CO_2 as they can plasticize. As a result, permeation capabilities are compromised, and membrane lifespan is critically reduced. That is why not every dense, anisotropic polymer can be used (CHIOU *et al.*, 1985). A glassy membrane, for instance, needs CO_2 to enhance its permeation properties. However, once conditioned, it can no longer return to its original form. The CO_2 changes its characteristics permanently (CHIOU and PAUL, 1987).

When using membranes, not only CO_2 is separated, but also a few hydrocarbons are lost from NG. That amount depends on many factors, such as membrane density, structure, length, and operational pressure gradient. Thus, the membrane process does not have the removal efficiency of an absorptive or cryogenic process. However, recycling can minimize this feed loss, not extinguish it. Furthermore, for all the pros and cons, the membrane process has a better benefit-cost ratio than other technologies (AARON and TSOURIS, 2005)

2.2 Digital Twin Concepts

The Digital Twin (DT) concept is attributed to Michael Grieves at the University of Michigan in 2003, according to GRIEVES (2014). At that time, Grieves referenced DT as a digital copy of an actual manufacturing process, just like a computational mirror. However, depending on the application, there needed to be more technological advances and a richness of information regarding the actual process to generate a reliable digital twin, as it could not represent the physical process at that time. In the chemical industry, this concept was known many years earlier, with applications in process monitoring, model predictive control, and real-time optimization, for instance.

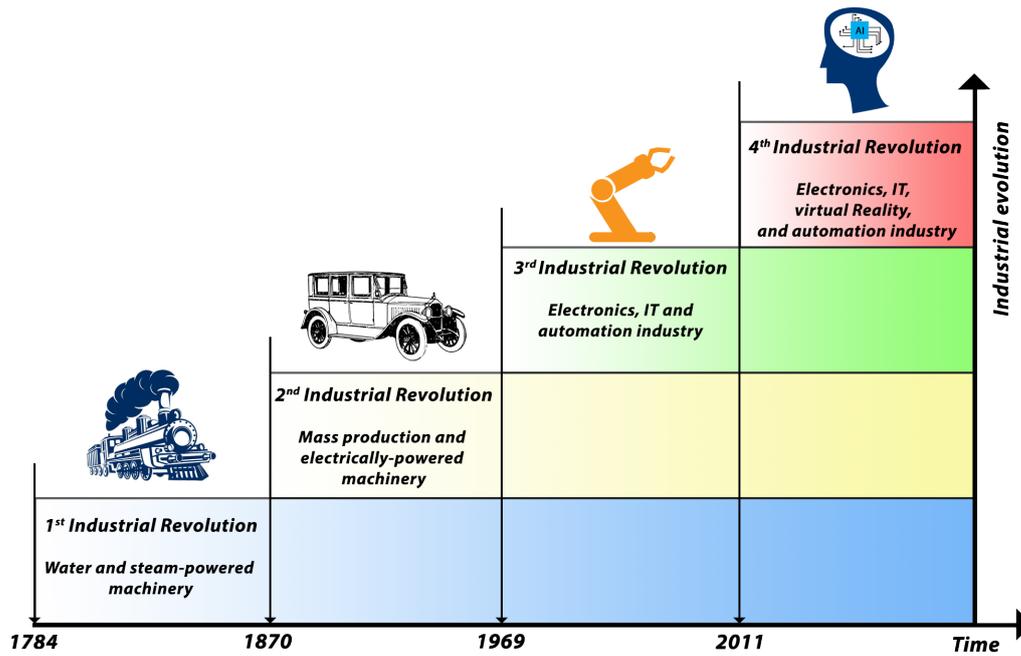
Since then, many advances in instrumentation, software development, and a higher amount of data have become available, so DT has become more representative incrementally. Nowadays, lightweight models, such as Artificial Neural Networks (ANN), has fast numerical processing capabilities, and, consequently, can simulate complex processes and their behaviors in real time. As a result, those advances make it possible to simulate, control, and even predict process behavior and operations from the most diverse manufacturing segment (GRIEVES, 2014).

With advances in the industry and recent developments in information technology, such as the Internet of Things (IoT), Big Data, and Artificial Intelligence (AI), integration between physical and digital environments has become even close. Brand new possibilities are available to explore, such as Manufacture, Oil & Gas, Medical, and Transportation, among other industries (TAO *et al.*, 2019).

That integration is creating a new era of industry, namely Industry 4.0, or the Industrial Internet. These are how highly integrated digital-physical industries are being called. This new approach has allowed processes to be monitored in real-time 24/7, which means better control systems coordination, and even enhanced predictive control management. Digital twin, in other words, is a digital mirror of an actual process that can analyze a massive amount of data, evaluate, optimize a process, predict results or system tendency and even actuate on a control system (TAO *et al.*, 2019).

According to KAGERMANN *et al.* (2013), it all began in 2011 when the term 4th industrial revolution emerged. In a German press release, it was said that an era of Cyber-Physical Systems would, not only strengthen the national industry but also drive new developments globally.

Figure 2.9 shows the industry's evolution from the steam revolution in 1784 until today with remote analyzers, controllers, and artificial intelligence.



Source: developed by the author.

Figure 2.9: The evolution of industrial eras.

Due to its young age, there are few pieces of research regarding digital twin applications in the literature regarding the O&G industry. However, some industries around the globe are known to be selling Digital Twin solutions to O&G companies, like the Norwegian FutureOn with its FieldTwin® and the British 2H Offshore with its Riser System Digital Twin services, for instance, (FUTUREON, 2020) (2H OFFSHORE, 2020).

KANE *et al.* (2015) conducted research on different employees, and how they see their company dealing with the industry 4.0 tools and their competition against other companies. This research resulted that the O&G industry was in its infancy regarding digital maturity, achieving a grade of 4.68 out of 10, where 10 is a fully mature industry, with a few metrics inside the bottom five of all companies researched. That study pointed out that there is a significant gap for improvement regarding O&G industry 4.0 turnover.

2.2.1 Building a Digital Twin

As mentioned above, a Digital Twin relies on an actual process or equipment. That said, the first step to creating a digital twin is to collect enough data from a physical process. That can be achieved by installing an appropriate and well-located amount of sensors to measure key input and output variables, and if the process

is affected by the environment, its surroundings also must be monitored in this process. The next step is to create communication between the industry ground and the digital environment. That must include reliable communication, including edge processing, security, and storage to keep all that data updated and saved in safe and easy-to-get cloud storage. After that, the process must be modeled just as it is in the field, using a software environment that can read and understand all that data that is constantly kept and saved (LU *et al.*, 2019).

Keep in mind that this digital model must be able to analyze all that data, and its parameters are adjusted to achieve a maximal similarity to physical process responses, just analyzing all its inputs and delivering its outputs in a graphic or table and the differences between the real-process responses to it predicted through the DT. That way, the DT can be trained to indicate if this difference is a normal process variation or if some failure is occurring. That gives one last, but not only, characteristic of a DT, which is a warning system when the physical process presents deviations more significant than usual, alerting its controller to verify what is happening on the field (DEON *et al.*, 2022; LU *et al.*, 2019).

Some DT is also used as an actuator once it is attached to the control system and its P&ID is well modeled; DT can also be trained to mitigate process deviations or failures. It can control a valve or even shut down the entire system in a critical emergency, acting as a predictive controller and detecting a failure sooner than it used to. Furthermore, with a digital model, optimizing algorithms can be used to improve real process productivity and performance (PARROTT and WARSHAW, 2017).

Additionally, it is easy to see that, in a 4th industrial revolution era, DT is a crucial tool to achieve better performance, monitoring, and safety processes through a product lifespan. KENDER *et al.* (2021) developed a digital twin-based approach for a cryogenic flexible Air Separation Unit (ASU). The model proposed used theoretical mass, energy, and momentum transport balances. The proposed DT includes a highly detailed dynamic virtual model of the ASU, which can be connected to historical or real-time plant data. The study justifies the high modeling effort through a detailed analysis of the shutdown scenario. The ASU's high degree of integration creates hazardous situations, such as low vacuum conditions in the SC, that may not be detected without simulation studies.

LIN *et al.* (2021) proposed a 3-dimensional digital twin model for particle tracing in a hollow-fiber ultra-filtration process using theoretical mass and momentum transport balances, and also analyzing pore patterns from membrane cavities. The study investigated the particle capture mechanism in detail through particle tracing imaging and analysis. The authors concluded that the simulation method presented a high level of accuracy, efficiency, and detailed information

involving the filtration mechanism, which is challenging to obtain through experiments. Through the combination of image analysis and multi-scale simulation, the study established a platform to evaluate membrane performance with high reliability and efficiency.

However, despite the growing interest in digital twin technology, there has been no research found on the development of a digital twin for an offshore hollow-fiber membrane-based gas separation unit that would use both theoretical equations and machine-learning-based models. Because of that, this work focuses on building a Digital Twin from an actual process of an offshore NG deacidification system using different models, comparing its performance to contribute to the O&G industry.

The next section show a background and how a membrane-based gas separation module can be modeled using 2 different approaches:

- Purely phenomenological model
- Purely artificial neural network (Black Box) model

2.3 Membrane Phenomenological Model

As previously discussed, gas separation through dense and anisotropic membranes is governed by the concentration gradient of the desired compounds between the sides of the membrane. In gaseous systems, it is common to represent this gradient as a function in terms of partial pressure. With a model in hand, it is possible to evaluate some properties, predict state variables tendency, separate system optimal design, predict maintenance costs, design control systems, and predict possible operational problems (CHU *et al.*, 2019).

As discussed in Section 2.1.3, there are two most common membrane module designs in offshore O&G industries: hollow-fiber-shaped and spiral-wound-shaped modules. In this research, an FPSO operated by SBM Offshore dataset, which uses Cynara's Natco® hollow fiber membranes made with cellulose acetate, was studied (SBM OFFSHORE, 2013). In the next section, phenomenological models of this system are presented in detail.

2.3.1 Hollow-Fiber model

Hollow-fiber membrane modeling was studied in detail by PAN (1986) under high pressures. According to the study, the model was a uni-dimensional mass balance and used the Hagen-Poiseuille equation to describe pressure permeate

stream pressure drop. The work of PAN (1986) added significant value to hollow-fiber membrane studies, serving as the basis for many subsequent works. For instance, KOVVALI *et al.* (1994) based their studies on the model proposed by PAN (1986) to study the performance of another numerical solving method. CHOWDHURY *et al.* (2005), based on PAN (1986) models but using Wilke's method for nonpolar gases to calculate the mixture viscosity, and solved the continuity equations using finite differences algorithm. Furthermore, CHU *et al.* (2019) also used the PAN (1986) approach and the Hagen-Poiseuille equation for predicting shell-side pressure drop. They studied the effects of packing factors on mass and momentum balances.

Hollow-Fiber membrane models are the ones that most vary due to their higher flexibility to process configurations. Permeation can flow parallel or cross-flow from the feed. Besides, if parallel, permeate can flow co-current or counter-current. The same variation is possible by feeding the sour gas from the shell or tube sides. In total, at least six different configurations influence phenomenological model development (MARRIOTT and SØRENSEN, 2003). For that reason, this work focuses on the process commonly found in the offshore O&G industry, which does not use permeate cross flow, only parallel. This reduces possibilities up to 4 different configurations that are similar to each other (CHU *et al.*, 2019).

Based on studies from PAN (1986), MARRIOTT and SØRENSEN (2003), CHU *et al.* (2019), HE *et al.* (2017), COKER *et al.* (1999), WHITAKER (1977), BIRD *et al.* (2004), COKER *et al.* (1998) and many others, it is reasonable to consider the following hypothesis for the hollow-fiber model:

- The membrane module arrangement is the hollow-fiber type with feed on the shell side and permeates flow on the tube side, as shown in Figure 2.10;
- Permeates configuration is counter-current;
- Although the co-current flow is not perfectly established near each module's header, radial mass change due to this contribution is negligible, when considering the process as a whole, and can be disregarded;
- Membranes are perfect. All of them have the same pore distribution, mean pore size, and thickness (h) throughout the length;
- Both sour and permeate gas have constant diffusion coefficients for all their components;
- Membranes do not deform under pressure or temperature differences;
- Gases does not condensate in any of the module's side flow.

Mass balance model

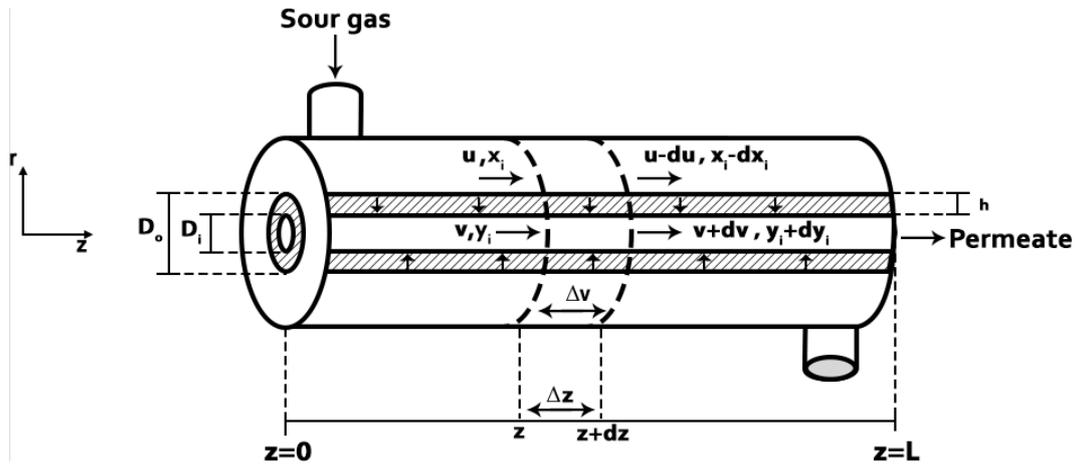
According to BIRD *et al.* (2004), the following equation can describe the equation of continuity for a multi-component mixture and constant diffusive coefficient:

$$\frac{\partial c_i}{\partial t} + (\nabla \cdot c_i s) = D_i^{ABs} \nabla^2 c_i \quad (2.8)$$

In which, $D_i^{ABs} \nabla^2 c_i$ is the diffusive term according to Fick's law of diffusion, D_i^{ABs} is the i^{th} shell side component diffusion coefficient for binary mixture, $(\nabla \cdot c_i s)$ is the advective term related to velocity pattern inside each fiber gap. Considering two dimensional model and cylindrical coordinates, Equation 2.8 can be rewritten into Equation 2.9:

$$\frac{\partial c_i}{\partial t} = D_i^{ABs} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial c_i}{\partial r} \right) + \frac{\partial^2 c_i}{\partial z^2} \right] - \left[\frac{1}{r} \frac{\partial (r s_r c_i)}{\partial r} + \frac{\partial (s_z c_i)}{\partial z} \right] \quad (2.9)$$

Furthermore, for a hollow-fiber module, it is convenient to establish a single fiber mass, momentum, and heat balance and then extrapolate for N number of fibers in a module, as shown in Figure 2.10



Source: developed by the author.

Figure 2.10: Co-current hollow-fiber module representation.

where u and v are sour gas and permeate flow speed vector, respectively, X_i and Y_i are shell side and tube side i^{th} component molar fraction, D_i and D_o are single hollow fiber inside and outside diameter, respectively, and the difference between

them is the membrane thickness represented by h . Also, L is accounted for the total hollow fiber membrane total length.

Besides, a mass flux through a membrane film is given by Equation 2.6 from Section 2.1.3 that relates molar flux through a membrane sheet as locally proportional to compound permeability and its partial pressure (p_i) gradient throughout membrane thickness (h).

$$J_i = -k_i \frac{dp_i}{dh} \quad (2.10)$$

For which, extrapolating to N membranes of A_o and V_o outside area and volume, respectively, and constant membrane thickness (h) gives the global molar flux equation for a single module as shown in Equation 2.11.

$$J_i = -\frac{NA_o k_i}{V_o h} (P_x X_i - P_y Y_i) \quad (2.11)$$

where P_x is shell side pressure, P_y is tube side pressure and A_o and V_o are given by the following equations:

$$A_o = \pi D_o L \quad (2.12)$$

$$V_o = \frac{\pi D_o^2 L}{4} \quad (2.13)$$

Besides, since it is a gaseous system susceptible to pressure influence, an equation of state is also required to connect molar concentration and its partial pressure value. For that, the Peng-Robinson equation of state can be used, due to its accuracy for hydrocarbon mixtures, and is represented by Equation 2.14 (PENG and ROBINSON, 1976).

$$\left\{ \begin{array}{l} P_i = \frac{RT}{V_i - b_1} - \frac{a_1}{V_i(V_i + b_1) + b_1(V_i - b_1)} \\ a_1 = 0.45724 \frac{R^2(T_i^c)^2}{P_i^c} (1 + (0.37464 + 1.54226\omega_i - 0.26992\omega_i^2) (1 - \sqrt{T_i^r}))^2 \\ \omega_i = -1 - \log_{10} \frac{P_i^{sat}}{P_i^c} \Big|_{T_i^r=0.7} \\ b_1 = 0.0778 \frac{RT_i^c}{P_i^c} \\ T_r = \frac{T}{T_i^c} \end{array} \right. \quad (2.14)$$

where P_i is i^{th} pure component vapor pressure, T stands for temperature, R is the universal gas factor, V_i is the i^{th} component volume, T_i^c stands for i^{th} component critical temperature, P_i^c is the i^{th} component critical pressure, T_i^r stands for i^{th} component reduced temperature, ω_i is PITZER *et al.* (1955) acentric factor for i^{th} component and P_i^{sat} is the i^{th} component saturation pressure.

Momentum balance model

According to BIRD *et al.* (2004), a general equation of motion applies to the system and is shown in Equation 2.15.

$$\frac{\partial(\rho s)}{\partial t} = -[\nabla \cdot \rho s s] - \nabla P + \mu \nabla^2 s + \rho g_t \quad (2.15)$$

where $[\nabla \cdot \rho s s]$ is the convective momentum flux rate per volume from feed across membrane sheet, $\nabla P + \mu \nabla^2 s$ is the molecular transport momentum rate per volume, and ρg_t is the momentum increment rate per volume due to an external force, here represented as an increment due to gravitational forces. In two dimension cylindrical coordinates, the following equations for the momentum balance model consider constant viscosity (μ) (BIRD *et al.*, 2004).

• Axial direction

$$\frac{\partial(\rho^m s_z)}{\partial t} = - \left[\frac{\partial(\rho^m s_r s_z)}{\partial r} + \frac{\partial(\rho^m s_z^2)}{\partial z} \right] - \frac{\partial P}{\partial z} + \mu \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial s_z}{\partial r} \right) + \frac{\partial^2 s_z}{\partial z^2} \right] \quad (2.16)$$

• **Radial direction**

$$\frac{\partial (\rho^m s_r)}{\partial t} = - \left[\frac{\partial (\rho^m s_r^2)}{\partial r} + \frac{\partial (\rho^m s_z s_r)}{\partial z} \right] \frac{\partial P}{\partial r} + \mu \left[\frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial r s_r}{\partial r} \right) + \frac{\partial^2 s_r}{\partial z^2} \right] \quad (2.17)$$

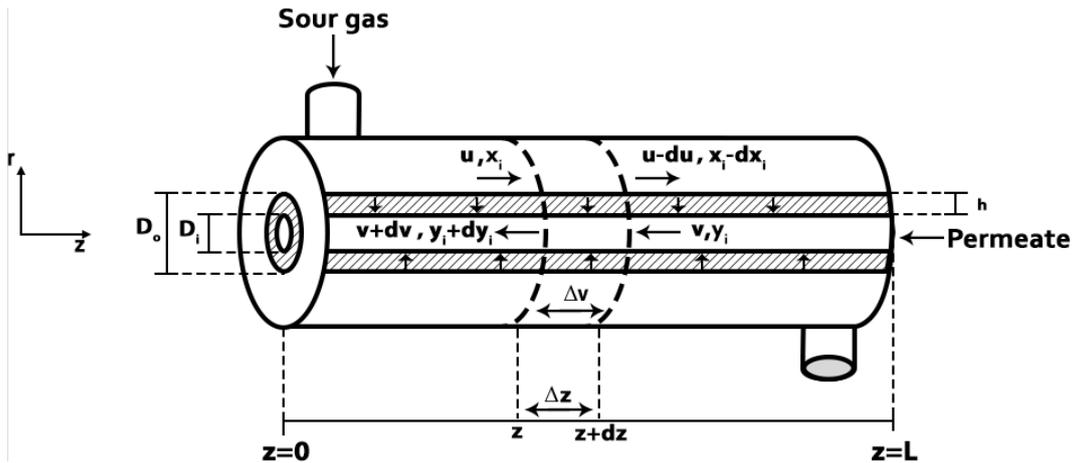
where:

$$\rho^m = \sum_{i=1}^{N^c} c_i \quad (2.18)$$

Many studies, though, go further into considering a few other terms disposable. PAN (1986), for instance, assumes that concentration polarization does not substantially affect model balance and that the axial modeling is sufficient to describe the system. The author considered the steady-state model and pressure drop inside each module cavity following the Hagen-Poiseuille equation, presented in Equation 2.19.

$$\frac{\partial P}{\partial z} = - \frac{128RT\mu v_z}{\pi D^4 NP} \quad (2.19)$$

For the counter-current parallel model, though, as described by Figure 2.11, the integration method still considers z varying from 0 to L . However, permeate speed will assume negative values rather than a co-current process (MARRIOTT and SØRENSEN, 2003).



Source: developed by the author.

Figure 2.11: Counter-current hollow-fiber module representation.

Energy balance model

According to WHITAKER (1977), a general energy balance equation can be described by the Equation 2.20.

$$\rho c_p \left(\frac{\partial T}{\partial t} + s \nabla T \right) = -\nabla(K \nabla T) + T \beta \left(\frac{\partial P}{\partial t} + s \nabla P \right) + \nabla s : \epsilon + \Phi \quad (2.20)$$

where the energy flux variation through time and space is proportional to energy flux per unity volume due to conduction ($\nabla(k \nabla T)$), the rate of reversible work per unity volume ($T \beta \frac{DP}{Dt}$) and viscous dissipation (irreversible work) per unity volume ($\nabla s : \epsilon$), which is always positive as it stands for irreversible work, and heat generation or consume per unity volume (Φ) (WHITAKER, 1977). The viscous dissipation ($\nabla s : \epsilon$), for Newtonian fluids, and isobaric thermal expansion coefficient (β), are described accordingly to Equations 2.21 and 2.22, respectively:

$$\nabla s : \epsilon = 2\mu \left[\left(\frac{\partial s_r}{\partial r} \right)^2 + \left(\frac{\partial s_z}{\partial z} \right)^2 \right] + \frac{\mu}{r^2} \left(\frac{\partial s_r}{\partial z} \right)^2 + \mu \left(\frac{\partial s_r}{\partial z} + \frac{\partial s_z}{\partial r} \right)^2 \quad (2.21)$$

$$\beta = -\frac{1}{\rho} \left(\frac{\partial \rho}{\partial T} \right)_p \quad (2.22)$$

Furthermore, for the consumption of heat equation (Φ), the hollow-fiber membrane module has a similar structure to shell-tube heat exchanger. The Logarithmic Mean Temperature Difference (LMTD) model can be used to predict heat flow from shell-side to bore-side for single fiber, which, according to BIRD *et al.* (2004), COKER *et al.* (1999) and WHITAKER (1977) can be described by Equation 2.23.

$$\Phi = Q_{mem} = \frac{U_0 A_m}{V_m} \left[\frac{(T_f^s - T_f^t) - (T_i^s - T_i^t)}{\ln \left(\frac{T_f^s - T_f^t}{T_i^s - T_i^t} \right)} \right] \quad (2.23)$$

where T_f^s and T_i^s are shell side final and initial flow temperature values, T_f^t and T_i^t are tube side final and initial flow temperature values, respectively, and U_0 is the overall heat transfer coefficient through membrane thickness. The total membrane area and volume are A_m and V_m , respectively. It can be described according to Equation 2.24 (COKER *et al.*, 1999):

$$U_0 = \left[\frac{1}{h^t} \left(\frac{D_0}{D_i} \right) + \frac{D_0}{2K_w} \ln \left(\frac{D_0}{D_i} \right) + \frac{1}{h^s} \right]^{-1} \quad (2.24)$$

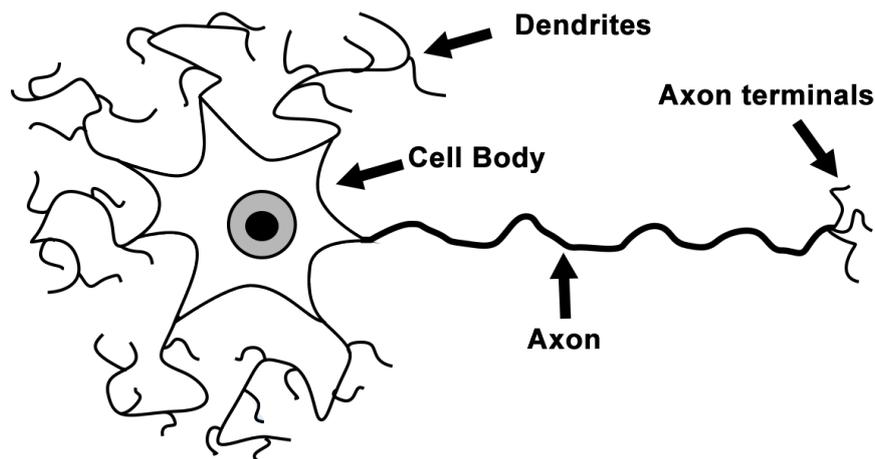
where h^s and h^t are shells and tube heat transfer coefficients, respectively, and K_w is the hollow fiber thermal conductivity coefficient.

2.4 Artificial Neural Network

2.4.1 Neural Cell structures and model

Biological neural Cells constitute three significant structures: dendrites, cell body, and axon. Each has a specific function: dendrites are receptors of the electric trigger from another neural cell in a process known as synapses. The cell body, for instance, translates this trigger into neural impulses, which is the information that the cell will pass forth to another neural cell through its axon structure until its terminals (HAYKIN, 2001; KOVÁCS, 1996). A biological cell structure is demonstrated in Figure 2.12.

This process begins when a series of neural impulses are transferred from one cell to another through neurotransmitters that provoke a change in electrical potential next to cell dendrites. These potential changes can be exciting or inhibitory. If a releasing threshold is achieved, all those triggers combined produce an action potential where the information is analyzed and codified into a frequency of pulses that is transferred through the axon to another synaptic area. Another piece of information passes on to another dendrite. The process starts all over again (KOVÁCS, 1996).



Source: developed by the author.

Figure 2.12: Biological neural cell structure

An artificial neural cell work quite similarly to a biological cell. The main idea is a structure capable of receiving a series of inputs and translating them into an

output or series of outputs. Broadly in the neural networks model, an artificial neural network tries to predict a system or process answer by analyzing its inputs and their relation with each output variable (KOVÁCS, 1996).

According to KOVÁCS (1996), a biological neural cell has a maximum axon frequency according to the Equation 2.25, which is a relation between the resting period (T_r) after a pulsing period (T_p). However, to achieve the maximum pulsing frequency, a crescent dendrite depolarization is directly related to the inputs' frequency and characteristics.

$$f_{max} = \frac{1}{t_r + t_p} \quad (2.25)$$

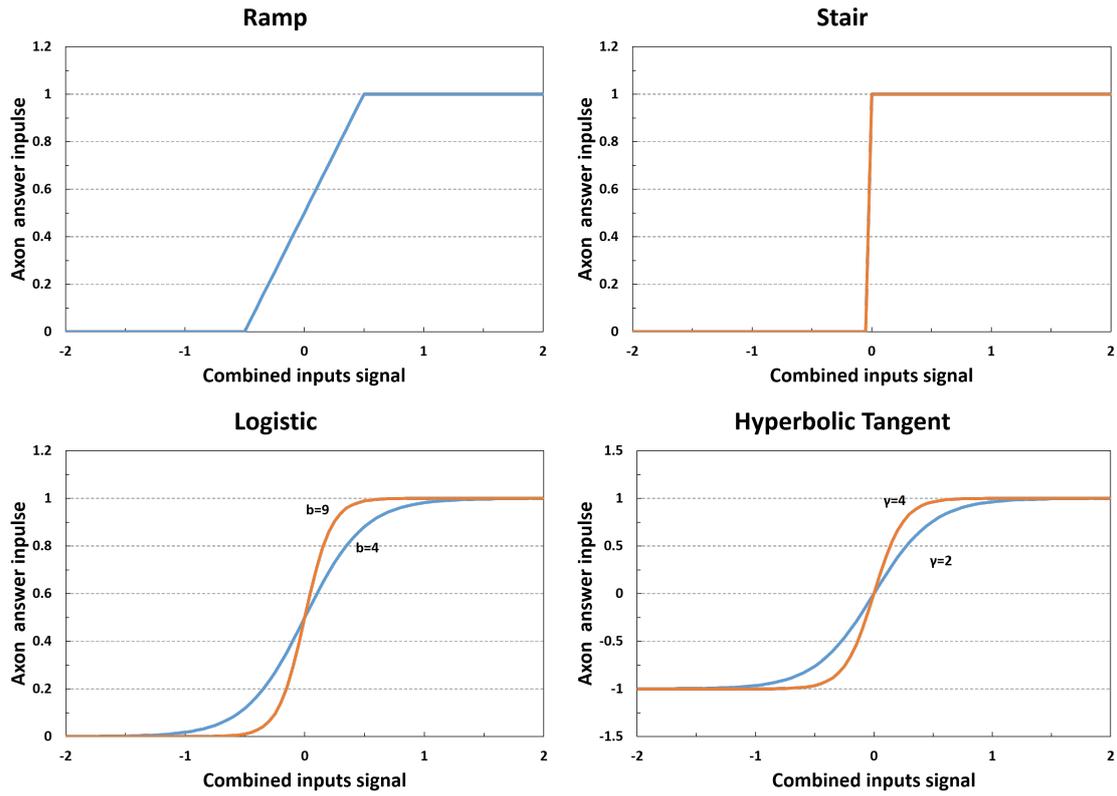
The relationship between dendrite depolarization due to inputs and axon pulse frequency is similar to a first-order equation. It can be described as a weighted input combination in Equation 2.26 (KOVÁCS, 1996).

$$g(t) = \varphi \left(\int_t \sum_{i=1}^N w_i(t) x_i(t) dt \right) \quad (2.26)$$

where w is the weight of input value x from i^{th} inputs, and $g(t)$ is the mean frequency of axon pulses under a certain amount of time t . That axon frequency is the neural cell's answer related to the inputs received during that period. Additionally, as mentioned before, there is a maximum frequency that each cell's axon is capable of producing, which has a first-order-equation shape, as described in Equation 2.26. $\varphi(\psi(t))$ denominated as activation function, which needs to have the same profile behavior as the relationship between pulse frequency and cell depolarization (HAYKIN, 2001; KOVÁCS, 1996).

That activation function can be modeled through a variety of different functions. However, the most commonly found in the literature are sigmoid functions, such as logistics or hyperbolic tangent functions, linear functions, or stair functions. They are mathematically described according to Equations 2.27 to 2.30, and their profiles can be visually compared in Figure 2.13.

$$\varphi(\psi) = \begin{cases} 1 & \text{if } \psi \geq 0; \\ 0 & \text{if } \psi < 0. \end{cases} \quad (\text{Stair}) \quad (2.27)$$



Source: developed by the author.

Figure 2.13: Activation function profiles comparison

$$\varphi(\psi) = a\psi \quad (\text{Ramp}) \quad (2.28)$$

$$\varphi(\psi) = \frac{1}{1 + \exp(-b\psi)} \quad (\text{Logistics}) \quad (2.29)$$

$$\varphi(\psi) = \tanh(\gamma\psi) = \frac{1 - \exp(-\gamma\psi)}{1 + \exp(-\gamma\psi)} \quad (\text{Hyperbolic tangent}) \quad (2.30)$$

However, in an actual process, time is considered discrete due to monitoring limitations. For that reason, only an instant evaluation is done. At that instant, a specific neural output (y_k) is evaluated from limited pulse inputs (x_k). That way, Equation 2.26 loses its time dependency and can be rewritten as described in

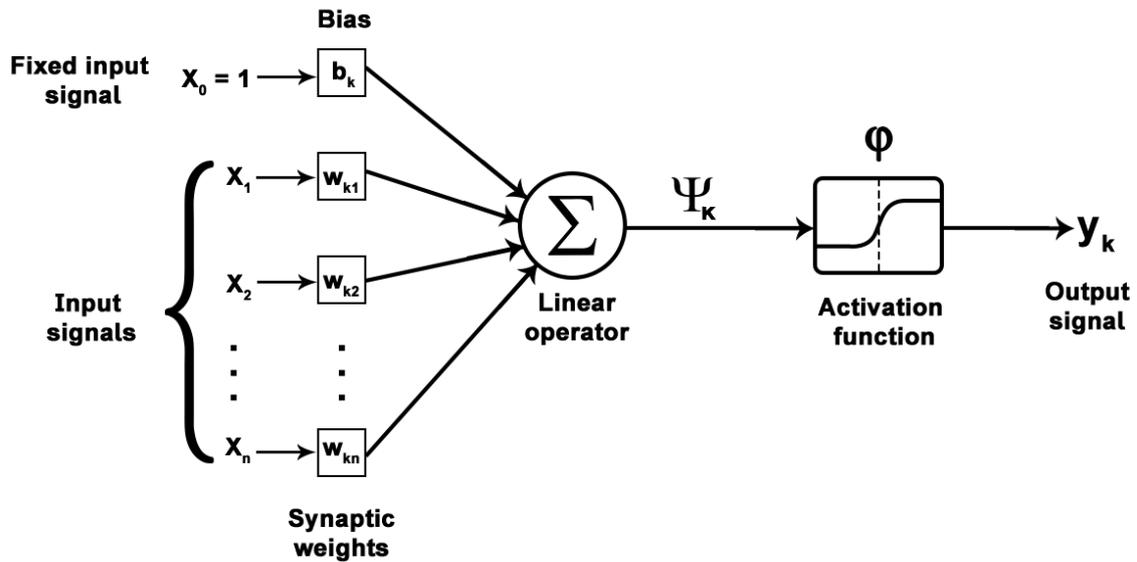
Equation 2.28 (HAYKIN, 2001).

$$y_k = \varphi \left(\sum_{i=1}^N w_i x_i \right) = \varphi \left(w^T x \right) \quad (2.31)$$

where " y_k " is the k^{th} neural cell output according to weighted inputs combination. Furthermore, HAYKIN (2001) proposes an existence of a bias parameter (b_k) which is nothing less than a linear operator that enhances or depletes the liquid neural input linear combination (u) from k^{th} neural cell. In mathematical terms, a bias can be integrated into Equation 2.31, which turns out the following:

$$y_k = \varphi \left(\sum_{i=1}^N w_i x_i + b_k \right) \quad (2.32)$$

Or yet, it can be assumed that bias is another weight from a unitary input and incorporated as a dendrite input as illustrated in Figure 2.14



Source: adapted from HAYKIN (2001).

Figure 2.14: Artificial neural cell structure.

Furthermore, this is a neural cell model using only equations. This neural cell model has been nominated *perceptron* and was first proposed by MCCULLOCH and PITTS (1943) with their Boolean model, and updated by ROSENBLATT (1958) as the first model to learn with a teacher, which will be shown later in this work.

However, the biological neural cell does not work alone. There are approxi-

mately 100 billion neurons only in the brain (VON BARTHELD *et al.*, 2016). That amount is connected in a complex network capable of reading, analyzing, thinking, retaining information, and responding correctly to all the exterior world stimuli. That is how ANN raised the interest to work with, as it can be a precursor of complex AI machinery (HAYKIN, 2001).

2.4.2 Advantages of Neural Network

To understand a neural network, we need to look more closely at how our brain processes all received information. According to HAYKIN (2001) a human brain can organize its 100 billion neural cells into processing highly complex information in a non-linear way or even in parallel processing. In the early years of born, a human brain can create its own rules regarding all the inputs received (vision, smell, taste, and feel are examples of inputs). These rules are usually denominated experiences. Once it experiences an input, the human brain starts to learn and understand its surroundings and adapt to them.

The artificial neural network structure is quite similar to a biological brain. One neural cell is relatively inefficient. For that reason, to achieve a good performance, an ANN is composed of some neural cells, as presented in Section 2.4.1, intertwined and connected to other similar structures in a distributed-parallel-nonlinear processing unit capable of learning, adapting, and storage knowledge from a series of environment experience. That learning process consists of changing synaptic weight and neural cell bias to best adapt its objective to each ANN stimulus (HAYKIN, 2001). In other words, an ANN structure is a high-performance adaptive processing unit (ALEKSANDER and MORTON, 1990).

An ANN can be non-linear or linear, depending on how each neural cell is modeled. A non-linear ANN means that this model can learn and understand non-linearity between inputs, outputs, or among them, which means that a non-linear ANN can be applied to a variety of cases. Besides, this characteristic enables another benefit from ANN. For instance, if there is a process in which the relation between inputs and output is not very clear in the literature, or even phenomenological models do not fit the data. ANN can map between inputs and outputs by "studying" the input-output relation, creating a functional black-box model that fits the process's data (HAYKIN, 2001).

Also, ANN can always be retrained to follow process modifications as time goes by. In other words, an ANN can be designed to adapt its synaptic weights at a real-time rate. Another robust ANN quality provides robustness to the model. Besides, each time a different input reaches the first neural cell's dendrites, this information is passed by the whole network as each neural cell is connected in

one way or another (HAYKIN, 2001).

Another essential benefit of ANN is its parallel-processing characteristic makes it computationally fast. This enables their use for very-large-scale integration (VLSI), which, according to MEAD (1989), is a powerful tool to analyze actual complex behaviors.

However, only some things are pro. There are a couple of cons: ANN could fail or even could not correctly describe what is happening. One of them is that a significant amount of data is needed to train an ANN, which relies on process instrumentation and data storage, and not all processes have the amount of data required to implement ANN. Another negative characteristic of ANN is that a data-based model does not rely on physical laws, only the interaction between input and output data, which is also called the black-box model. This is likely to be troublesome if one wishes to understand the process mechanisms further (ASGHARI *et al.*, 2018).

To overcome the lack of data many studies are using data augmentation techniques. This methodology is used to oversample a dataset by generating new and diverse examples based on the existing data. It is commonly used in machine learning and deep learning applications to overcome the problem of imbalanced datasets, where the number of samples in one class is much smaller than the other, or when the total data amount is not enough to make a model learn properly about the variables' relation. By applying data augmentation, the model is trained on a larger and more diverse dataset, which can improve its ability to generalize and recognize patterns in the data. Data augmentation can involve various techniques such as rotating, flipping, zooming, and adding noise to the original data to create new samples that are still representative of the original data (DING *et al.*, 2023; MIKOŁAJCZYK and GROCHOWSKI, 2018).

2.4.3 ANN training process

An ANN model's primordial and most crucial step is the learning phase, which consists of evaluation and weight adaptation to improve the model's performance. Much information regarding the exciting process needs to be available and constantly updated. This is because an ANN needs to analyze process inputs and adapt its synaptic weights to approach its results to actual process results. Furthermore, ANN weight optimization is called the learning phase, which is not unique. Different structure ANNs has their learning process. Studies regarding learning algorithms have been made since McCulloch developed the first neural cell model in 1943. In this section, some essential learning algorithms will be presented. Each one has its advantages and differs mainly in how a synaptic weight

is modified (HAYKIN, 2001).

The first learning algorithm was proposed by HEBB (1949), who proposed that a highly complex neural network learning behavior could be reduced to a local learning process. In this study, synaptic performance depended on a correlation function between inputs and outputs, and for that, a local error evaluation should suffice the network's learning procedure. Hebb's learning proposal is described on Equations 2.33 to 2.35.

$$w_{ki}^{new} = w_{ki}^{old} + \Delta w_{ki} \quad (2.33)$$

where Δw_{ki} is a function of inputs and outputs:

$$\Delta w_{ki} = F(y_k, x_{ki}) \quad (2.34)$$

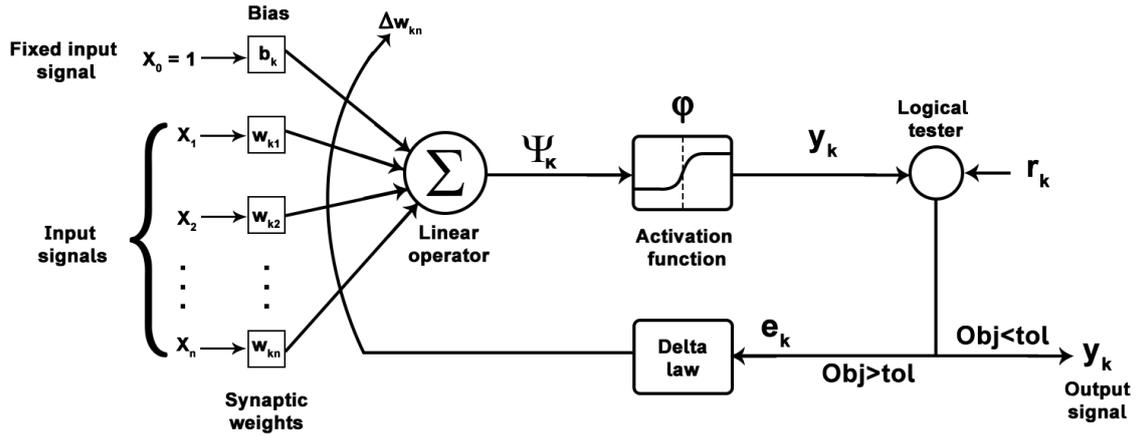
As example of correlation function we have:

$$\Delta w_{ki} = \eta (r_k - y_k) x_{ki} \quad (2.35)$$

where η is the learning rate, it modifies as the learning process advances. This parameter was deeply studied and modified throughout the years, and it is further discussed later in Section 2.4.4

Another simpler learning algorithm is *Error Correction Learning*, which is an adaptation of Hebb's postulate proposed by WIDROW and HOFF (1960) and consists of the Least-Mean-Square (LMS) theorem, which proposes another error signal output type. That signal triggers the adapting algorithm, which modifies the synaptic weights and bias from the k^{th} neural cell.

The adjustment is directed towards minimizing an objective function described by Equation 2.37 as the Widrow-Hoff LMS algorithm. The error signal can be described as Equation 2.36, and to illustrate that mechanism, Figure 2.15 demonstrates how the process is done. In other words, what an error correction learning algorithm does is add a logical evaluation between actual data and predicted exit and, depending on its performance, proposes a correction on the synaptic weights, which, according to WIDROW and HOFF (1960), is determined by Equation 2.38 and is identified as Widrow-Hoff law or Delta law (HAYKIN, 2001) (KOVÁCS, 1996).



Source: adapted from HAYKIN (2001).

Figure 2.15: Error-correction-based artificial neural cell structure.

$$e_k(t) = \sum_{i=1}^N ||r_k(t) - y_k(t)|| \quad (2.36)$$

$$Obj(t) = \frac{1}{N} e_k^2(t) \quad (2.37)$$

$$\Delta w_{ki}(t) = \eta e_k(t) x_i(t) \quad (2.38)$$

By the end of this process, each synaptic weight can be described as a function of an initial estimation plus the delta provided by Widrow-Hoff law, as Equation 2.39 describes. In computational terms, Equation 2.39 can be rewritten in a discrete time using zeta transformation, and the result is demonstrated in Equation 2.40 where ζ^{-1} is a unit delay operator.

$$w_{ki}(t+1) = w_{ki}(t) + \Delta w_{ki}(t) \quad (2.39)$$

$$w_{kn}(\zeta) = \zeta^{-1} [w_{ki}(\zeta + 1)] \quad (2.40)$$

In the next section, the present work will demonstrate how to model and how a Multi-Layer Perceptron (MLP) network works. There are many other network

structures in the literature. However, as this work will focus on MLP networks, an explanation of fundamentals will be equally specific.

2.4.4 Multi-Layer-Perceptron Network

A multi-layer perceptron network consists of perceptron-model neural cells connected with another series of neural cells and organized in a layer-shaped structure. In this ANN structure, a neural cell layer response generates an input signal for another neural layer, providing inputs for the ones that came forward. In other words, an MLP network consists of an initial perceptron layer that receives process input signals. Then this layer responds and feeds another perceptron's layer, which feeds another and another until the last perceptron layer responds to the desired variable. The signal always propagates in one direction (forward). That is why this process is called a feedforward network. (HAYKIN, 2001; KOVÁCS, 1996).

Figure 2.16 illustrates a structure of an MLP network. Each network node corresponds to a perceptron structure described in Figure 2.14 from Section 2.4.1.

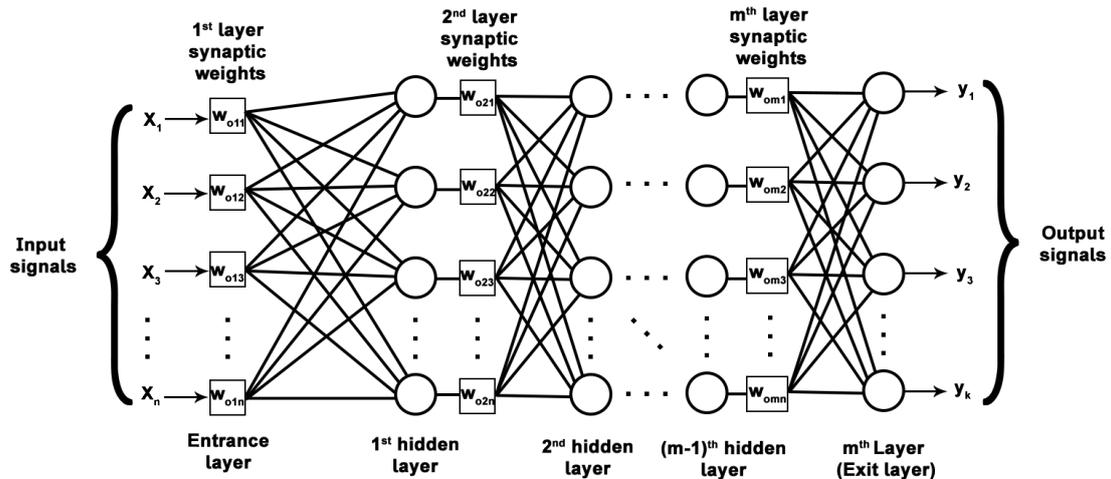


Figure 2.16: Multi-Layer Perceptron Network Scheme

Source: adapted from HAYKIN (2001).

Figure 2.16 illustrates that each synaptic input weights each neural cell node from the next layer. Under the same logic, W_{omn} means that the n^{th} neural cell input signal from m^{th} layer has a specific weight for the subsequent layers O^{th} neural cell node. Furthermore, there goes until the last node layer gives ANN's response. Keep in mind that output signals and layer's neural cells amount do not need to match (" O " does not strictly have to have the same dimension as " $n4$ "). Even layers do not need to have the same neural cell number.

- **The Backpropagation Algorithm**

As mentioned in Section 2.4.3, every ANN has its learning method. MLP network learning process is a mix of those three learning algorithms mentioned in this work. MLP uses an *error backpropagation*, commonly known as the Backpropagation algorithm only, which is an adaptation from Hebb's, Widrow-Hoff, and Rosenblatt algorithms studied and published by RUMELHART and MCCLELLAND (1986).

The backpropagation algorithm uses the same principles from Equations 2.36 and 2.37, with few differences. The only "visible layer" is the last one, so the error measure can only be calculated at this specific point. Then the main idea is to adjust synaptic weights from the MLP network from end to the beginning according to the LMS algorithm, which says that a synaptic weight correction $\Delta w_{mj}(t)$, where $w_{mj}(t)$ is the m^{th} exit layer's neural cell synaptic weight from an j^{th} previous layer neural cell response of t^{th} interaction time, is proportional to objective's function gradient described by $\partial Obj(t) / \partial w_{mj}(t)$.

Rewriting all the equations that describe an MLP network model gives the Equations 2.41 to 2.43 below.

$$\frac{\partial Obj(t)}{\partial w_{mj}(t)} = -e_m(t) \varphi'_m(u_m(t)) y_j(t) \quad (2.41)$$

$$\Delta w_{mj}(t) = \eta \delta_m(t) y_j(t) \quad (2.42)$$

This method δ is a local gradient that directs those synaptic weights correction, which in this case can be written as:

$$\delta_m(t) = -e_m(t) \varphi'_m(u_m(t)) \quad (2.43)$$

Moreover, this is a backpropagation weight update algorithm idea considering that the m^{th} layer is the exit layer. Furthermore, for an immediate exit previous layer "j", that $j = m - 1$ layer, the Equation 2.43 can be rewritten into Equations 2.45.

$$e_m(t) = ||r_m(t) - \varphi_m(u_m(t))|| \quad (2.44)$$

$$\delta_j(t) = \varphi'_j(u_j(t)) \sum^m \delta_m(t) w_{mj}(t) \quad (2.45)$$

where the weight correction $\Delta w_{j(j-1)}(t)$ keeps subjected to the delta law, which is described by Equation 2.46.

$$\left\{ \Delta w_{j(j-1)}(t) \right\} = \left\{ \begin{array}{c} \text{Learning Rate} \\ \eta \end{array} \right\} \cdot \left\{ \begin{array}{c} \text{Local gradient} \\ \delta_j(t) \end{array} \right\} \cdot \left\{ \begin{array}{c} \text{j-1 cell input} \\ \text{signal} \\ y_{(j-1)}(t) \end{array} \right\} \quad (2.46)$$

This gives the following prior-exit layer weight correction algorithm:

$$\Delta w_{j(j-1)}(t) = \eta \varphi'_j(u_j(t)) \left[\sum^m \delta_m(t) w_{mj}(t) \right] y_{(j-1)}(t) \quad (2.47)$$

Equations 2.41 to 2.45 and 2.47 detailed development can be found in the Appendix B. The Learning rate parameter plays a vital role in finding the gradient minimum and can be determined through different approaches. For instance, RUMELHART and MCCLELLAND (1986) proposed a fixed step (steepest descent) towards finding the negative gradient. However, DE SOUZA JÚNIOR (1993) showed that pursuing the optimum value using fixed steps is inefficient due to difficulties in choosing a fixed step that is not too large that overshoots the variable's optimum values or too small that hinders the method's convergence, especially near the optimum value.

To minimize the overshoot and oscillation possibilities, thus not giving up on convergence speed, MCCLELLAND and RUMELHART (1988) proposed a modification on the backpropagation algorithm to include a *momentum* term described by Equation 2.48. where ε is the *momentum* constant that correlates the past weight changes on the new values.

$$\Delta w_{mj}^{new}(t) = \eta \delta_m(t) y_j(t) + \varepsilon \Delta w_{mj}^{old}(t) = -\eta \nabla \text{Obj}(w_{mj}) + \varepsilon \Delta w_{mj}^{old}(t) \quad (2.48)$$

However, according to DE SOUZA JÚNIOR (1993) and LEONARD and KRAMER (1990), this method may not be enough to minimize the oscillation due to the fixed learning rate, and the use of fixed momentum parameters may also cause a lack of assertiveness as the algorithm may cycle the optimal point. LEONARD and KRAMER (1990) proposed a conjugate line search with gradient descent in both directions which η is updated using a line search algorithm and ε

is updated according to Equation 2.49 for the $q + 1^{th}$ interaction.

$$\begin{cases} \varepsilon = \frac{\nabla Obj(w_{q+1})^T \nabla Obj(w_{q+1})}{\nabla Obj(w_q)^T \nabla Obj(w_q)} \\ w_{q+1} = w_q + \eta s_q \\ s_{q+1} = -\nabla Obj(w_{q+1}) + \varepsilon s_q \\ s_0 = -\nabla Obj(w_0) \end{cases} \quad (2.49)$$

Recently, some mixed methods started attracting attention to parameter optimization. They use a combination of the stochastic objective function and evaluate its gradient behavior such as *Adam* (Adaptive Moment Estimation) algorithm proposed by KINGMA and BA (2015), which does a stochastic optimization using only first-order gradients and calculates adaptive learning rates according to the first and second moment gradients. Stochastic gradient-based algorithms are often used on deep learning applications where the input instances are high dimensioned or online learning is required (DUCHI *et al.*, 2011; KINGMA and BA, 2015). Another algorithm that has recently been attracting attention is the Adaptive Gradient Algorithm (AdaGrad) developed by DUCHI *et al.* (2011), which is considered as being an adaptive sub-gradient method that takes into account the function profile from previous interactions to adapt the learning rate. For that reason, many computational packages, such as Keras for Python or OptimLib for C++, choose to use these methods to solve their weight and bias optimization problems.

2.5 Physics-Informed Neural Networks

In this study, a Physics-Informed Neural Network (PINN) model is also proposed to monitor the membrane separation process without relying solely on data. At the beginning of 2019, an innovative data-driven approach using Neural Networks was developed by RAISSI *et al.* (2019). The authors have developed a neural network-based structure capable of solving data-driven problems while considering the governing laws by which the problem is bounded. The supervised-learning algorithm was called Physics-Informed Neural Network and is capable of learning not only the data behavior over a specific operating region but also the relationship that each output variable related to each other. This is possible because PINNs are trained using data, residuals from model equations, and boundary conditions.

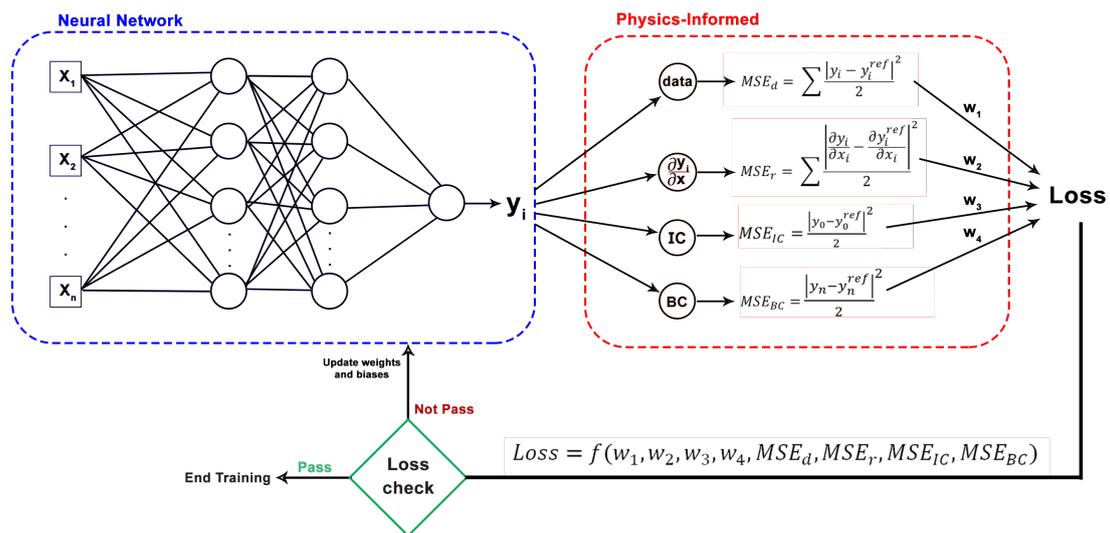
PINNs was developed to enhance the machine-learning-problems solution process to a level in which their physical laws should be respected, as close as

possible, regardless of the data behavior. Complex non-linear problems, such as fluid dynamics, will no longer need to be deeply tested for inconsistent results that violate mass or heat transfer principles. RAISSI *et al.* (2019) affirmed and showed how useful and accurate the PINN approach solves partial non-linear differential equations. They also used them for parameter inference, which comes in handy when dealing with processes without further knowledge of equipment parameters.

2.5.1 PINNs Structure and Logic

The first developed PINN structure was similar to an MLP structure elucidated previously in Section 2.4.4. However, the target data used for training does not consist solely of data. As output, a PINN is set to return the data, the residuals from the governing equations, and the boundary conditions. This also enhances the amount of information a PINN can provide, thus avoiding the black-box characteristics.

Figure 2.17 illustrates how the PINN is structured. At first, a neural network takes time-dependent or shape-dependent input values and predicts every substantial value. In Figure 2.17, the NN model is used to predict the Boundary Conditions, and governing equation residuals, and data, next an MSE algorithm, such as the one illustrated in Section 2.4.3, can be used to calculate a simple or weighted MSE that will be used for the training algorithm (such as ADAM backpropagation) that updates the NN weights and biases, and the results are validated again until the end of the training process.



Source: developed by the author.

Figure 2.17: PINN scheme.

The most common network used for PINNs are the MLP networks discussed previously in Section 2.4.4 (CARVALHO and BRAGA, 2022). Since the NN weights and biases update are a function of weighted individual loss functions, a great variety of configurations are possible, which enables the PINN approach to be used for many different problems, for instance: thermodynamics (CARVALHO and BRAGA, 2022), medicine (SAHLI COSTABAL *et al.*, 2020), electric and electronic (NELLIKATH and CHATZIVASILEIADIS, 2022), Civil Engineering (GOKHALE *et al.*, 2022) and O&G field (FRANKLIN *et al.*, 2022; MUDUNURU *et al.*, 2020; QUEIROZ *et al.*, 2021).

According to the developer, RAISSI *et al.* (2019), the PINN algorithm has higher sensitivity to data and residual deviation. For that reason, there is no standardized approach to treat the data prior to or after, even for solving the same problem. Also, a learning algorithm is susceptible to the loss function proposed by the developer. Training sessions for the PINN are the same exposed in Section 2.4.3 and 2.4.4 with only the difference of the objective function gradients calculation.

2.6 Literature Review

As already mentioned in Chapter 1 and the last sections, the present work intends to create a digital twin from an offshore gas deacidification unit while studying a new approach mixing phenomenological and data-driven models to respect physical laws while being faster enough to be implemented on real-time monitoring software. The proposed idea is to mix physical laws and data-driven structures to search for the best approach by using the data available to estimate the compound's permeability while relying on finite-volume phenomenological models to determine the other variables; resulting in a PINN structure capable of determining every variable for every discretization step in the digital-twin that could be faster and reliable enough.

To ensure that those approaches had never been used in the literature, a search using the most renowned research repository was used to verify this authenticity, such as Elsevier, Wiley Online Library, ACS Publications, Web of Science, and Google Scholar. As a result, similar research was found, but they had yet to use a hybrid model to create a digital twin from an existing process. This makes the present work the first of it until the date of its publication. Some similar works had though modeled the NG/CO₂ deacidification process using different approaches, as presented in the following:

PAN (1986) studied a high pressurized system of hollow-fiber membrane module for H₂, CO₂ and H₂S separation. The author considered an isotherm

model with a constant permeability coefficient. It was the first study to consider the permeate flow pressure build-up calculation using the Hagen-Poiseuille equation. Shell side pressure drop was still neglected. PAN (1986) also considered a steady-state process with plug flow on both sides (shell and tube). Because of that, a one-dimensional (axial) mass and momentum balance were used to model the problem, which was solved by an interactive shooting method. The author also validated the model using experimental data.

COKER *et al.* (1999) used a similar approach as PAN (1986) for a multi-component simulator for hollow-fiber membrane modules. However, it also included an energy balance model and more components from natural gas treatment. This study elaborated a pure phenomenological model where shell side pressure drop was neglected, and tube side pressure build-up was described by the Hagen-Poiseuille equation. Both flows were assumed to be in a plug-flow pattern and a steady-state process. The mathematical model used WHITAKER (1977) thermal energy balance for gas and used finite differences to solve the equations. The authors also used the Redlich-Kwong equation of state to predict gas mixture properties and neglected conductive heat transfer compared to the axial advective term. The objective was to evaluate the differences that the permeability coefficient suffered once membrane temperature changes on a variety of CO_2 and CH_4 gas mixtures. The authors concluded that temperature significantly changes the compound permeability coefficient associated with thermal gas expansion. COKER *et al.* (1999) did not validate their model with any experimental data; their work was to show the differences in permeability parameters whether considering temperature variation or not.

MARRIOTT (2001) and MARRIOTT and SØRENSEN (2003) evaluated many phenomenological models for hollow-fiber and spiral-wound membrane models for gas separation. In some cases, the authors considered mass, momentum, and energy balances and solved their equations using the orthogonal collocation method and finite-compounds method. Unlike the previously mentioned works, the authors also considered pressure drop on both module sides. However, this model challenge was to predict diffusion and dispersion coefficients among many other parameters such as permeability variation, fluid thermal conductivity, isobaric thermal expansion, and membrane fiber thermal conductivity. This model can be tricky if it does not consider some of these parameters from the literature. The authors validated their model using the data from PAN (1986) work.

The work of CHOWDHURY *et al.* (2005) was based on PAN (1986) model for hollow-fiber membranes, but using variable viscosity modeling. Gas mixture viscosity was calculated using Wilke's method for nonpolar gases and used component viscosity calculation as a temperature-dependent function. Their model was

validated with PAN (1986), SIDHOUM *et al.* (1988) and HARAYA *et al.* (1986) experimental data. Furthermore, the authors eliminated the required pressure initial guess required by the PAN (1986) model and incorporated their solution routine into *AspenPlus*TM as a useful model for further optimization and economic evaluation studies involving membrane separation for gases.

SHAHSAVAND and CHENAR (2007) study compared the ability of two different neural networks (Radial Basis Function - RBF, and MLP) with a multipurpose software for parameter estimation and equation solver (Table Curve ®) to predict the permeances and separation factors of hollow fiber membranes. Two experimental datasets were used to train the networks for separating CO₂ from CH₄, for the experimentation, both membrane types were used: polyimide-based and poly-phenylene-based. Both RBF and MLP networks outperformed the TC software. As result, the predictions of MLP networks were found to be highly influenced by the initial values of synaptic weights, which could result in overfitting, particularly when there was noise in the data. Conversely, RBF networks, when combined with suitable isotropic spread and regularization parameters, had the ability to effectively eliminate noise and detect the real underlying hyper-surfaces that may be obscured by noisy data sets.

PEER *et al.* (2008) proposed a different approach to model a hollow-fiber natural gas separation process. The authors used an Artificial Neural Network to predict module gas separation and the required membrane area to process different natural gas compositions. The authors modeled an MLP Network with two hidden layers with 5 and 4 perceptrons, respectively, three inputs not described by the authors and one output (retentate composition or membrane selectivity). They used hyperbolic tangent as an activation function, the Levenberg-Marquardt backpropagation training method with a mix of experimental data and the ones found in the literature. Their ANN model was validated using COKER *et al.* (1998) data. However, besides the conclusion that the ANN model can predict selectivity and retentate composition well, it has a strong possibility that their model is overfitted to their experimental data, that is because there are only a few experimental points over a limited operation region which, hardly their approach and MLP network structure can be applied to an actual steady-state gas separation process.

JUNG *et al.* (2010) developed and compared two models for CO₂ separation in hollow-fiber membrane modules: an explicit model based on mass balances and an MLP back-propagation neural networks model. Experimental data obtained from a single-stage module with recycle were used to validate the explicit model and train the MLP network. As result, the authors presented a small discrepancy between theoretical and computational model results. Finally, the authors con-

cluded that the proposed models could effectively be used in the analysis and operation of gas separation processes using hollow-fiber membranes.

AHMAD *et al.* (2012) proposed implementing an isothermic phenomenological two-dimensional cross-flow hollow-fiber CO_2/CH_4 separation unit model and implemented into *Aspen Hysys*TM for further optimization and economic studies. Furthermore, the study focused on simulating many different gas separation process configurations to evaluate the best configuration for CO_2 removal. Their model was validated with PAN (1986) and QI and HENSON (1998) experimental data on NG separation, which is odd, as PAN (1986) work relies on parallel permeation. In contrast, QI and HENSON (1998) studied an optimized configuration for a spiral-wound natural gas separation unit. They concluded that the gas processing cost is minimal when two modules are arranged in series in which the second permeate is recycled into the first feed flow.

LOCK *et al.* (2015) studied a CO_2 capture system from natural gas using a phenomenological model from different hollow-fiber membrane configurations. Shell side pressure variations were neglected, the plug-flow pattern was assumed on both module sides, the Hagen-Poiseuille equation for permeate pressure build-up was used, and the isothermic pattern was applied. The authors considered a radial cross-flow pattern for membrane gas diffusion and solved the equations using finite radial differences for co-current, counter-current, and cross-flow configurations. The authors' main idea was to compare the separation performance among these configurations and also do an economic evaluation once incorporated into *Aspen Hysys*TM. The study concluded that the counter-current pattern has a slightly higher performance than the cross-flow pattern, and both have much higher than the co-current separation pattern. The authors also noted that the most economical configuration does not always have the highest separation output. That is because different patterns require different auxiliary machinery, which impacts the gas processing cost.

CHU *et al.* (2019) used a similar approach as PAN (1986) did for an isothermic multi-component simulator for hollow-fiber membranes modules natural gas separation unit. However, it was also considered a pressure drop model for the shell side. This study elaborated a pure phenomenological model where shell side pressure drop and tube side pressure build-up were described by the Hagen-Poiseuille equation using hydraulic diameter concepts. Both flows were assumed to be in steady-state and as a plug-flow pattern. The authors introduced dimensionless variables to help the numerical algorithm and proposed an orthogonal collocation method. They also considered two different configuration patterns (co-current and counter-current). They implemented the model in MATLAB[®] and named as Mollocator, which was validated using an in-house membrane model, which

was validated using experimental data. The authors' main idea was to develop another approach for the hollow-fiber membrane model. However, they could have explained how the counter-current process was evaluated, and the Hagen-Poiseuille equation viscosity parameter value was not found in their work. They concluded that for higher packing factors, shell side pressure drop can significantly affect gas separation but needed to specify how higher it can be for not to influence much.

ULLAH *et al.* (2019) developed different silica-based membranes by coating silica solution into alumina-based support. They proposed a neural network approach for evaluating the effects of a silica-solution dip-coating time, inlet pressure, and inlet flow rate into different alumina-support structures to evaluate the permeability and selectivity factors, in pursuit of enhancing CO_2/CH_4 gas separation process. The authors used MLP neural network model with a backpropagation algorithm to predict the permeability factors using experimental data of pressure, inlet flow rate, and coating time as input variables. The predicted results from the MLP showed strong reliability, validity, and applicability. The accuracy of the predicted data regression was verified by achieving high levels of agreement with R^2 values of 0.999 for CO_2 and 0.998 for CH_4 . The trained neural network had mean absolute deviation errors of 10^{-6} . Also, the authors found an optimum alumina coating time to maximize the CO_2/CH_4 selectivity on silica membranes and concluded that the synthesized silica membrane was promising for CO_2/CH_4 separation under certain operating conditions.

NASIR *et al.* (2022) synthesized different types of membranes (polymeric, amine-based, and filler) and proposed a neural network approach for evaluating the permeance and selectivity factors for a binary mixture of CO_2/CH_4 gas separation process. The authors used MLP neural network model with backpropagation algorithm to predict the permeance factors and CO_2/CH_4 selectivity using experimental data of pressure, and membrane material concentration (wt%) as input variables. The MLP model was found to be flexible enough to be used for several outputs using only one MLP model. The accuracy of the predicted data regression was verified by achieving high levels of agreement with overall R^2 values higher than 0.96 and with an average relative error of 6.1% for CO_2 , 4.2% for CH_4 permeances and 3.2% for CO_2/CH_4 selectivity. The authors also concluded that gas permeance was found to decrease as the pressure increased, attributed to the glassy nature of the polymer. Conversely, an increase in amine concentration resulted in an increase in CO_2 permeance due to its affinity with amine-based membranes. Furthermore, the authors suggested that pressure and amine concentrations are the predominant variables affecting gas permeability in facilitated transport mixed matrix membranes.

In the natural gas sweetening field, several primary studies have been conducted. However, as of February 2023, no research has been found that employs a hybrid ANN and phenomenological approach or a PINN structure. Instead, many phenomenological approaches have been used, and only a few have utilized a black-box model, primarily to forecast permeability and selectivity factors. While neural networks have been employed to optimize membrane-based gas separation processes, no study has utilized them to predict the separation process and compare different approaches. This reinforces the innovative characteristic of the present work and its high contribution to the literature.

The next chapter of this work will present a case study aimed at modeling an existing and operational process to achieve the first and second specific objectives of this study.

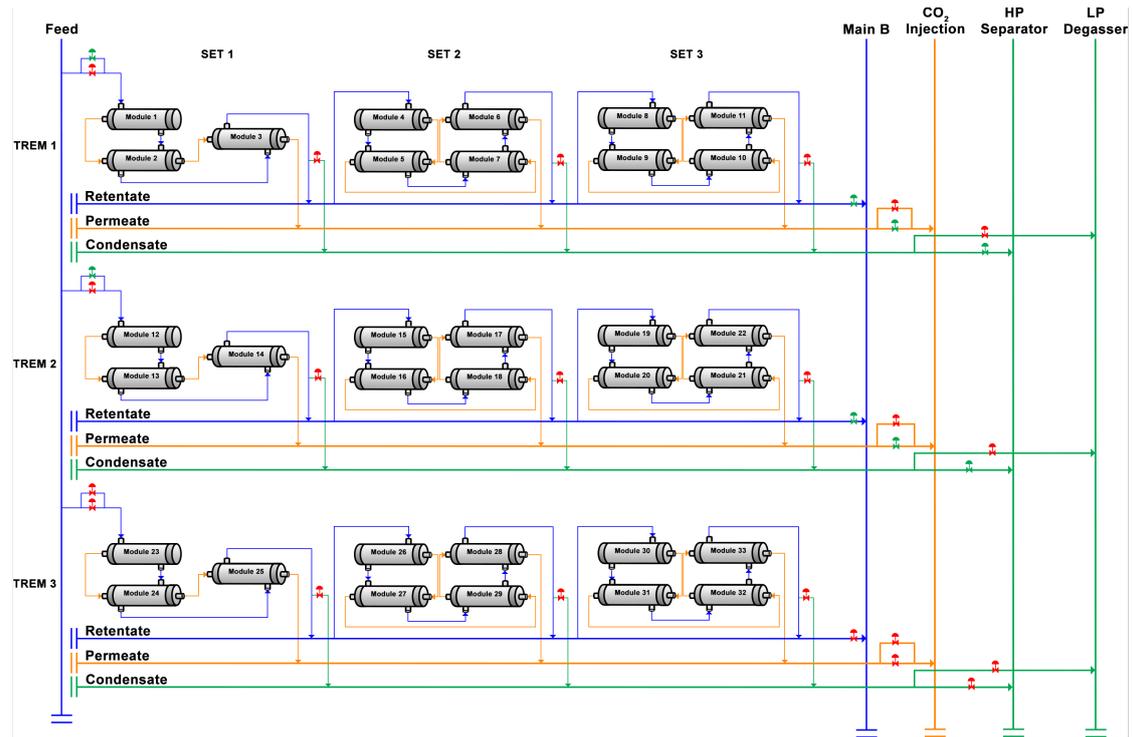
Chapter 3

Phenomenological model

3.1 Introduction

To derive a model for an actual gas separation unit, a detailed process flow diagram is required to determine what equipment and how many of each one are necessary. Also, it is imperative to know some key process inputs such as gas molar flowrate or velocity, feed temperature, pressure, and critical compound compositions. Besides, a real process generally has non-operating-parallel-ordered units, which are used while the other is under maintenance or had to be stopped operating for any other reason. That other unit must also be modeled accordingly to its components and equipment specifications.

This work is going to consider an FPSO gas separation configuration that is operating in Brazilian southwest oil fields whose process diagram is as shown by Figure 3.1. On that flow scheme, it can be seen three major parallel-organized gas separation units named "TREM" which are composed of 3 main separation structures each, the first with 3 membrane modules, the second with 4 membrane modules, and the third with 4 more membrane modules each. Also, Figure 3.1 shows 2 TREMs with open valves (green-colored), and the third TREM has all its valves closed (red-colored). The third TREM is the spare one. This FPSO always operates as shown in Figure 3.1: 2 TREMs operating while 1 is under maintenance. It can also be seen that the flow is a counter-current pattern on all modules. Besides, at the first modules, an inlet permeate flow also does not exist. Everything that flows into the permeate flowline comes exclusively from the feed stream. A larger copy from Figure 3.1 can be found in Appendix C for detailed evaluation.



Source: developed by the author.

Figure 3.1: A real FPSO gas separation process diagram.

3.2 Methodology

In this chapter, a phenomenological model is proposed for the process described by Figure 3.1, which mimics the one running the separation data used in this work. Depending on the model performance, another study regarding model dimension and accuracy is made to verify if dimension reduction should apply, thus lowering computational costs and model agility.

To achieve this work's first objective, model validation, some variables were considered constants and gathered from literature and shown as its methodology is developed. The algorithms were developed using Python programming language at Jupyter Notebook compiler.

3.2.1 Hollow fiber modules model and discretization

As mentioned in the last chapter, the model is proposed to consider the following hypothesis, which makes sense to the actual gas separation unit structure.

Hypothesis 1 (H1). *All membrane modules are identical, with identical membrane sheets. All of them have the same pore distribution, mean pore size, thickness, same packing factor and do not deform under pressure or temperature differences;*

Hypothesis 2 (H2). *The Peng-Robinson Equation of State describes the real gas mixture of CO₂ and CH₄;*

Hypothesis 3 (H3). *Radial temperature, mass, and momentum variation is considered uniform;*

Hypothesis 4 (H4). *Steady-state operation.*

Hypothesis 5 (H5). *Specific calorific capacity (C_p), thermal conductivity (k), and isobaric thermal expansion factor (β) are considered constant.*

Hypothesis 6 (H6). *Gases do not condensate in any of the module's side flows;*

Hypothesis 7 (H7). *The module is adiabatic;*

Hypothesis 8 (H8). *Reversible work and viscous dissipation are expected to be small even for a non-ideal gas mixture so that these terms can be neglected.*

Hypothesis 9 (H9). *Energy balance will not be considered. Instead, a constant temperature from the output flow from each side will be considered.*

Hypothesis 10 (H10). *The Hagen-Poiseuille equation will be used for both sides' pressure drop calculations and adapted to consider the non-ideal gas mixture.*

An evaluation of the process conditions on the steady-state condition as proposed by Hypothesis 4, both permeate and retentate feed temperatures are in equilibrium. So the majority of energy lost in the process is lost to the ambient in which the membrane modules are. Due to the lack of information about environmental conditions such as wind, temperature, and humidity, this study will consider the effects of temperature only for permeability calculations.

Furthermore, the Hypothesis 9 is based on the study of COKER *et al.* (1999), which concluded that, for low-density gases, the major impact of temperature was in the permeability factor. According to the authors, the temperature variation needed to double the permeability factor for CO₂ is 45 K, whereas for CH₄ is 20 K for a polyimide membrane. Supposing that this sensibility is also applicable to cellulose-acetate membranes, in this case study, the process conditions have an average temperature difference, between the inlet and outlet of 4.5 °C, for both inlet and outlet flows. Considering this, temperature variation would influence an average of 25% on CH₄ permeability and 11% on CO₂ permeability at the end of the process. For a single membrane module, this represents 2.3% for CH₄ and 1% for CO₂ of error. For that reason, the temperature variation was not taken into consideration.

Also, regarding Hypothesis 10, besides MULDER (1996) suggestion applying the Hagen-Poiseuille simplification for flows where the Reynolds number is

less than 2100, COKER *et al.* (1998) studied the application for low-density and high-pressure flows and how they influence the pressure profile for Hollow-Fiber membranes. The latest authors concluded that the pressure distribution could be well-represented by the Hagen-Poiseuille. Since that study, many other authors also rely on that conclusion, such as: AHMAD *et al.* (2012); CHU *et al.* (2019); COKER *et al.* (1999); LOCK *et al.* (2015) to propose variations or use it in their studies. Because of that, this equation should provide enough contribution towards the achievement of the first objective of this work.

The pressure drop model will be an adaptation from CHU *et al.* (2019). The authors used the Hagen-Poiseuille equation and considered ideal gas behavior, the shell side hydraulic diameter and the squared-pattern fibers arrange as mentioned in HE *et al.* (2017) as the most common pattern found. The hydraulic diameter considered by the authors is presented in Equation 3.1, and the friction factor used by the authors was the $24/Re$ instead of the original Hagen-Poiseuille tube flow friction factor of $16/Re$, which adaptation and rearranging the equations gives the Equation 3.2.

$$D_{xh} = \frac{4 \left(\frac{\pi D^2}{4} - \frac{\pi N D_o^2}{4} \right)}{\pi N D_o} \quad (3.1)$$

$$\frac{dP_x}{dz} = - \frac{192 N D_o (D + N D_o) R T \mu_m}{\pi (D^2 - N D_o^2)^3 P_x} \sum_{i=1}^2 v_{zi} \quad (3.2)$$

That being reinforced, considering all those hypotheses, and including the Peng-Robinson compressibility factor, the following equations from 3.3 to 3.6 are proposed to describe the membrane separation process.

$$\frac{d(u_{zi})}{dz} - \frac{D_i^{ABs}}{s_s} \frac{d^2 u_{zi}}{dz^2} = -\pi D_o N P (P_x X_i - P_y Y_i) \quad \text{Shell side} \quad (3.3)$$

$$\frac{d(v_{zi})}{dz} - \frac{D_i^{ABs}}{s_p} \frac{d^2 v_{zi}}{dz^2} = \pi D_o N P (P_x X_i - P_y Y_i) \quad \text{Tube side} \quad (3.4)$$

$$\frac{dP_x}{dz} = - \frac{192 N D_o (D + N D_o) R T \mu_m}{\pi (D^2 - N D_o^2)^3 Z_x P_x} \sum_{i=1}^2 v_{zi} \quad \text{Shell side} \quad (3.5)$$

$$\frac{dP_y}{dz} = \frac{128RT\mu_m}{\pi D_i^4 N Z_y P_y} \sum_{i=1}^2 u_{zi} \quad \text{Tube side} \quad (3.6)$$

The indexes 1 and 2 from the previous equations sum are the indexes that identify both elements from the binary mixture of CO_2 and CH_4 . Which, complies to the following boundary conditions visually explained in Figure 3.2 and itemized as follows:

1. Follows the Dirichlet boundary concept of a specified condition.

$$f = \text{constant} \quad (3.7)$$

2. Follows the Neumann boundary type of zero flux for the feed permeate condition.

$$\left. \frac{\partial f}{\partial z} \right|_{z=0} = 0 \quad (3.8)$$

3. Continuity of flux and variable.

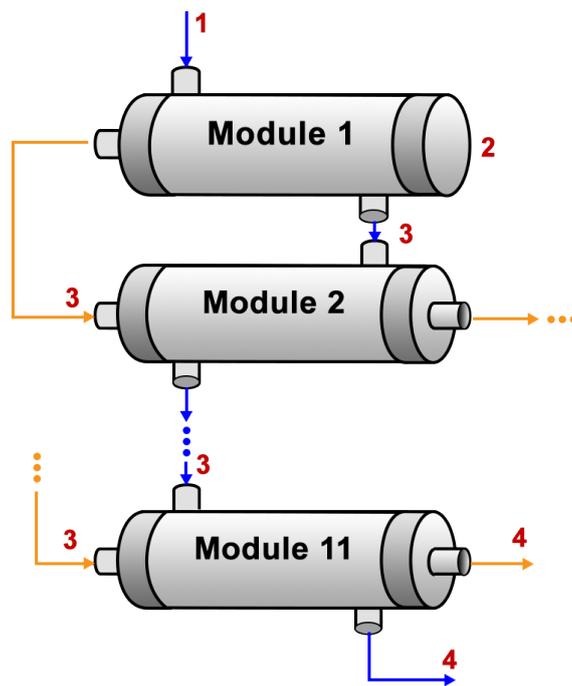
$$f_{out}^i = f_{in}^{i+1} \quad (3.9)$$

4. Null gradient for both models flow.

$$\nabla f = 0 \quad (3.10)$$

where f is a general representation of a state variable, f_{out}^i is a state variable value at the end of the i^{th} module and f_{in}^{i+1} is a state variable value at the beginning of the $i^{th} + 1$ module.

Next it is imperative to know module specification parameters such as internal diameter, fiber internal and external parameters, number of hollow fibers, length, and so on. Since this work objective is to make a digital twin, it relies on actual module specification data. However, due to intellectual property restrictions, the exact equipment parameters were not disclosed until the time of the development of this work. For that reason, and the effects of the study, a dummy model with parameters that complies with the maker model patent will be used instead and is resumed in Table 3.1 (CASKEY, 1993). The structure presented is enough to achieve the objectives proposed by this work without compromising the company's intellectual property.



Source: developed by the author.

Figure 3.2: Boundary conditions representation from the three different modules found in this process, where 1 specified entrance condition, 2 is zero flux, 3 represents the continuity of flux and variables, and 4 is where the gradient is null.

Table 3.1: Hollow fiber membrane module specifications comparison between the patent from CASKEY (1993) and those used for the digital-twin structure proposal.

| Hollow fiber module parameter | Patent | Used Values |
|---------------------------------|--------------------------|-------------------|
| Shell inner diameter (in) | 16 or 30 | 30 |
| Tube outer diameter (μm) | $80.5 \leq D_o \leq 240$ | 240 |
| Tube inner diameter (μm) | $70 \leq D_i \leq 130$ | 130 |
| Tube thickness relation | $h \geq 0.15D_i$ | $0.84D_i$ |
| Tube length (m) | 2 | 2 |
| Fibers amount on each module | $N \geq 50000$ | 2×10^6 |
| Membrane composition | Cellulose Acetate | Cellulose Acetate |

The values proposed to be used according to Table 3.1 are the ones that resulted in similar process results as the data gathered from January/2020 and January/2021. Since this work does not aim to estimate module parameters, it seems reasonable enough to move on with the development of this dissertation proposal with this module configuration.

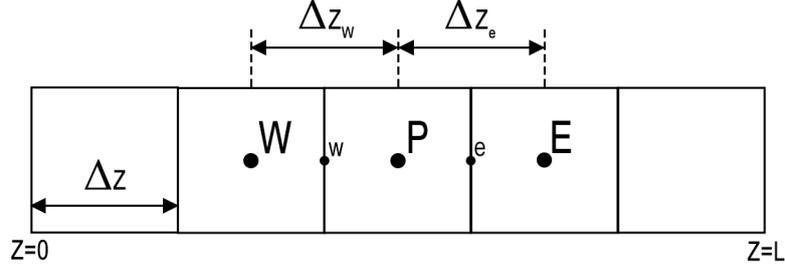
3.2.2 Model Discretization

For numeric implementation, models need to be discretized. In this work, the one chosen is the finite volume method because mass, momentum, and energy tend to conserve and because it is the most common method used on flow problems involving mass and energy transfer. It consists of creating an evaluation mesh inside each domain, which analyzes each property variation on each finite volume of the created mesh (MALISKA, 1995).

The discretization process of the finite volume consists of integrating Equations 3.3 to 3.6 on every finite volume, which yields the following equations:

$$\int_{\Omega_{VC}} \frac{d(u_{zi})}{dz} dVC - \int_{\Omega_{VC}} \frac{D_i^{ABs}}{s_s} \frac{d^2 u_{zi}}{dz^2} dVC = \int_{\Omega_{VC}} -\pi D_o NP (P_x X_i - P_y Y_i) dVC \quad (3.11)$$

$$\int_{\Omega_{VC}} \frac{d(v_{zi})}{dz} dVC - \int_{\Omega_{VC}} \frac{D_i^{ABs}}{s_p} \frac{d^2 v_{zi}}{dz^2} dVC = \int_{\Omega_{VC}} \pi D_o NP (P_x X_i - P_y Y_i) dVC \quad (3.12)$$



Source: developed by the author.

Figure 3.3: General 1D mesh representation of finite-volume method, with the central volume P and the neighboring volumes W and E. The border midpoints are represented by w and e, respectively.

$$\int_{\Omega_{VC}} \frac{dP_x}{dz} dVC = - \int_{\Omega_{VC}} \frac{192ND_o(D + ND_o)RT\mu_m}{\pi(D^2 - ND_o^2)^3 Z_x P_x} \sum_{i=1}^2 v_{zi} dVC \quad (3.13)$$

$$\int_{\Omega_{VC}} \frac{dP_y}{dz} dVC = \int_{\Omega_{VC}} \frac{128RT\mu_m}{\pi D_i^2 N Z_y P_y} \sum_{i=1}^2 u_{zi} dVC \quad (3.14)$$

where Ω_{VC} is each finite volume boundaries. Figure 3.3 represents a one-dimensional finite volume mesh of an arbitrary discretization. The central volume, represented by the letter P, is where the equations and fluxes will be evaluated. The letters W and E (West and East) represent the neighbor volumes, which share only one border with the P volume, whose midpoints are represented by letters w and e.

Applying Equations 3.11 to 3.14 under the finite volume represented by Figure 3.3, and using the central finite differences to approximate the first-order derivatives of the diffusive terms, gives the following discretized models:

$$\left(\frac{2D_i^{ABs}}{S_s \Delta z} \right) u_{ziP} = \left(\frac{D_i^{ABs}}{S_s \Delta z} - 1 \right) u_{zie} + \left(\frac{D_i^{ABs}}{S_s \Delta z} + 1 \right) u_{ziw} - \pi D_o N P_P (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta z \quad \text{Shell side} \quad (3.15)$$

$$\left(\frac{2D_i^{ABs}}{S_p \Delta z} \right) v_{ziP} = \left(\frac{D_i^{ABs}}{S_p \Delta z} - 1 \right) v_{ziw} + \left(\frac{D_i^{ABs}}{S_p \Delta z} + 1 \right) v_{zie} + \pi D_o N P_P (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta z \quad \text{Tube side} \quad (3.16)$$

$$P_{xe}^2 = P_{xw}^2 - \frac{384ND_o(D + ND_o)RT\mu_m}{\pi(D^2 - ND_o^2)^3Z_y} \sum_{i=1}^2 (v_{ziP})\Delta z \quad \text{Shell side} \quad (3.17)$$

$$P_{ye}^2 = P_{yw}^2 - \frac{256RT\mu_m}{\pi D_i^4 Z_x N} \sum_{i=1}^2 (u_{ziP})\Delta z \quad \text{Tube side} \quad (3.18)$$

One may notice that the signal changed for the tube side equations due to the way of integrating the equations for the counter-current pattern. While the integration from the shell side is done from $z=0$ to $z=L$ direction, for the tube side, the integration will take place from $z=L$ to $z=0$. At the end of the evaluation, the equation signals change. The values from each variable in the faces of each finite volume will be calculated as the average of the values found at w and e . In the first finite volume, w variable value will comply with the initial condition from Equation 3.7 and the e variable value will comply with the outlet condition from Equation 3.10 at the module's exit.

3.2.3 Molar volume calculation

To calculate phenomenological equation parameters, such as permeability, fugacity, molar flow, viscosity, and many others, the molar volume from each module side hydrocarbon mixture (permeate and retentate) will be calculated using Peng-Robinson thermodynamic model, described by Equation 3.19. According to BIRD *et al.* (2004). This equation of state can represent the binary mixture of CO_2 and CH_4 with high accuracy,

$$P = \frac{RT}{V - b_{mix}} - \frac{a_{mix}}{V(V + b_{mix}) + b_{mix}(V - b_{mix})} \quad (3.19)$$

where a_{mix} and b_{mix} for a mixture of compounds can be calculated using van der Waals mixture rules where a is quadratic approximation while b is linear MEHL (2009). The expression to calculate each parameter is as follows:

$$a_{mix} = \sum_{i=1}^n \sum_{j=1}^n y_i y_j a_{ij} \quad (3.20)$$

$$b_{mix} = \sum_{i=1}^n y_i b_i \quad (3.21)$$

where:

$$a_{ij} = (1 - k_{CO_2-CH_4}) \sqrt{a_i a_j} \quad (3.22)$$

where y_i is the molar fraction from i^{th} component, a_i and b_i are the properties of pure substances given by Equations 3.23 and 3.24.

$$a = 0.45724 \frac{R^2 (T^c)^2}{P^c} \left(1 + \left(0.37464 + 1.54226\omega - 0.26992\omega^2 \right) \left(1 - \sqrt{T^r} \right) \right)^2 \quad (3.23)$$

$$b = 0.0778 \frac{RT^c}{P^c} \quad (3.24)$$

where T^c stands for critical temperature, P^c is the critical pressure, T^r stands for reduced temperature described by Equation 3.26, ω is PITZER *et al.* (1955) acentric factor described by Equation 3.25 and P^{sat} is the saturation pressure.

$$\omega = -1 - \log_{10} \frac{P^{sat}}{P^c} \Big|_{T^r=0.7} \quad (3.25)$$

$$T^r = \frac{T}{T^c} \quad (3.26)$$

Table 3.2 presents the parameters used to calculate molar volume from this work.

3.2.4 Diffusion Coefficient calculation for low-density gases

To calculate the Diffusion Coefficient from each module side hydrocarbon mixture (permeate and retentate) at low density, the Chapman-Enskog model, described by Equation 3.27 for binary mixtures, will be used. Although it was first aimed to model monoatomic gases, according to BIRD *et al.* (2004), the model performed remarkably well for non-polar polyatomic gases.

Table 3.2: Peng-Robinson thermodynamic parameters for the fugacity calculation

| Parameter | CO ₂ | CH ₄ |
|---|-----------------|-----------------|
| P ^c (bar) ⁽¹⁾ | 73.9 | 46.3 |
| T ^c (K) ⁽¹⁾ | 304.1 | 190.4 |
| ω ⁽¹⁾ | 0.239 | 0.011 |
| k _{CO₂-CH₄} ⁽²⁾ | 0.103 | |

Source: (1)TERRON (2009); (2)LI (2008)

$$D_{AB} = 1.8583 \times 10^{-3} \sqrt{T^3 \left(\frac{1}{M_A} + \frac{1}{M_B} \right)} \frac{1}{P \sigma_{AB}^2 \Omega_{AB}} \quad (3.27)$$

where T is temperature, P is pressure, ω_{AB} is the collision integral, and it is given by Equation 3.28, σ_{AB} is the average molecule diameter and M_A and M_B the molecular mass from each component.

$$\Omega_{\epsilon} = \frac{1.06036}{T_{\epsilon}^0.15610} + \frac{0.1930}{\exp(0.47635T_{\epsilon})} + \frac{1.03587}{\exp(1.52996T_{\epsilon})} + \frac{1.76474}{\exp(3.89411T_{\epsilon})} \quad (3.28)$$

where T_{ϵ} is a dimensionless temperature given by Equation 3.29.

$$T_{\epsilon} = \frac{T}{\epsilon_{AB}} \quad (3.29)$$

$$\epsilon_{AB} = \sqrt{\epsilon_A \epsilon_B} \quad (3.30)$$

where ϵ_i is the molecule's characteristic energy. Table 3.3 presents the parameters used to calculate molar volume from this work.

3.2.5 Viscosity calculation for binary mixture gases

To calculate the viscosity from each module side hydrocarbon mixture (permeate and retentate) at low density, Wilke's model, described by Equation 3.31 for binary mixtures, will be used. Wilke's model is a less rigorous Chapman-Enskog-derived model for gas mixtures that provide sufficient accuracy without relying too much on computational cost (BIRD *et al.*, 2004; WILKE, 1950). Although it was first

Table 3.3: Chapman-Enskog parameters for the Diffusion Coefficient and viscosity calculation

| Parameter | CO ₂ | CH ₄ |
|--------------------------------|-----------------|-----------------|
| $M(\text{g/mol})$ | 44.01 | 16.04 |
| $\sigma(\text{Å})$ | 3.996 | 3.780 |
| $\epsilon(\frac{1}{\text{K}})$ | 190 | 154 |

Source: BIRD *et al.* (2004)

aimed to model monoatomic gases, according to BIRD *et al.* (2004), the model performed remarkably well for non-polar polyatomic gases at low density.

$$\mu_{mix} = \sum_{i=1}^n \frac{x_i \mu_i}{\sum_{j=1}^{n-1} x_j \Phi_{ij}} \quad (3.31)$$

where N is the number of chemical species in the mixture, x_i , is the mole fraction of i^{th} and j^{th} compounds for $i \neq j$, μ_i is the viscosity of pure compounds at a given temperature and pressure described by Equation 3.32, μ_{mix} the viscosity of the gas mixture and Φ_{AB} is the dimensionless quantities given by Equation 3.33.

$$\mu_i = 2.6693 \times 10^{-5} \frac{\sqrt{M_i T}}{\sigma_i^2 \Omega_{\mu i}} \quad (3.32)$$

$$\Phi_{ij} = \frac{1}{\sqrt{8}} \left(1 + \frac{M_i}{M_j}\right)^{-1/2} \left[1 + \left(\frac{\mu_i}{\mu_j}\right)^{1/2} \left(\frac{M_j}{M_i}\right)^{1/4}\right]^2 \quad (3.33)$$

where $\Omega_{\mu i}$ is the collision integral for viscosity, given by Equation 3.34.

$$\Omega_{\mu i} = \frac{1.16145}{T_\epsilon^0.14874} + \frac{0.52487}{\exp(0.77320T_\epsilon)} + \frac{2.16178}{\exp(2.43787T_\epsilon)} \quad (3.34)$$

in which T_ϵ is the same dimensionless temperature given by Equation 3.29. For binary mixtures, one can summarize the following couple of equations:

$$\mu_{AB} = \frac{x_A \mu_A}{x_A \Phi_{AA} + x_B \Phi_{AB}} + \frac{x_B \mu_B}{x_A \Phi_{BA} + x_B \Phi_{BB}} \quad (3.35)$$

$$\Phi_{AB} = \frac{1}{\sqrt{8}} \left(1 + \frac{M_A}{M_B}\right)^{-1/2} \left[1 + \left(\frac{\mu_A}{\mu_B}\right)^{1/2} \left(\frac{M_B}{M_A}\right)^{1/4}\right]^2 \quad (3.36)$$

where:

$$\Phi_{AA} = \Phi_{BB} = 1 \quad (3.37)$$

Table 3.3 presents the parameters used to calculate viscosity from this work.

3.2.6 Permeance calculation

To determine the change in the permeation values of CO_2 and CH_4 mixture along the membrane module, the mathematical functions developed by SABERI *et al.* (2016) were used. The presence of CO_2 in the feed/retentate stream promotes the plastification effect of cellulose acetate, which hinders compound transport and separation efficiency. SABERI *et al.* (2016) and co-workers assumed that the diffusivity across the membrane solely depends on the plasticizer component. To represent the sorption-diffusion model, the equations developed were based on: the dual sorption theory (Dual Mode Sorption), the partial immobilization sorption theory, and the competitive sorption theory.

The dual-mode sorption theory consists of the junction of two theorems. The first is Henry's law of diffusion, shown in Chapter 2, and the second from Langmuir's competition for permeation site theory ("hole-filling" theory). Equations 3.38 and 3.39 are SABERI *et al.* (2016) models for permeation in terms of component fugacity.

$$\mathbb{P}_{CO_2} = \frac{D_{CO_2 0/l}}{\beta_{CO_2} f_{CO_2}} \left\{ \exp \left[\beta_{CO_2} f_{CO_2} \left(k_{DCO_2} + \frac{F_{CO_2} C'_{HCO_2} b_{CO_2}}{1 + b_{CO_2} f_{CO_2} + b_{CH_4} f_{CH_4}} \right) \right] - 1 \right\} \quad (3.38)$$

$$\mathbb{P}_{CH_4} = D_{CH_4 0/l} \left\{ \exp \left[\beta_{CH_4} f_{CO_2} \left(k_{DCO_2} + \frac{F_{CO_2} C'_{HCO_2} b_{CO_2}}{1 + b_{CO_2} f_{CO_2} + b_{CH_4} f_{CH_4}} \right) \right] \left(k_{DCH_4} + \frac{F_{CH_4} C'_{HCH_4} b_{CH_4}}{1 + b_{CO_2} f_{CO_2} + b_{CH_4} f_{CH_4}} \right) \right\} \quad (3.39)$$

Table 3.4: Parameters for cellulose acetate hollow-fiber membranes

| Component | k_D ($cm^3/cm_{mem}^3 kPa$) | C'_H (cm^3/cm_{mem}^3) | b ($1/kPa$) |
|-----------------|---------------------------------|------------------------------|-----------------------|
| CO ₂ | 0.0143 | 37.29 | 1.32×10^{-3} |
| CH ₄ | 0.00151 | 37 | 1.32×10^{-4} |

Source: DONOHUE *et al.* (1989)

Table 3.5: Parameters for permeation evaluation considering binary mixture CO₂/CH₄ and linear approximation.

| Component | β | F | $D_{0,l}$ |
|-----------------|---------|-------|--------------------------|
| CO ₂ | 0.027 | 0.023 | 3.03257×10^{-3} |
| CH ₄ | 0.033 | 0.08 | 8.418×10^{-4} |

Source: Adapted from SABERI *et al.* (2016)

where \mathbb{P}_i is the component permeation factor, $D_{i0/l}$ is the diffusion coefficient of pure gas per membrane thickness, C'_{Hi} is the membrane site saturation constant β_i is an empirical constant that dependent on membrane material and thickness, and temperature, f_{CO_2} is the compound fugacity calculated using Peng-Robinson equation of state presented on Appendix D, k_{Di} is Henry's law solubility coefficient, b_i is the hole affinity constant and F_i is the fraction the compound absorbed.

Table 3.4 shows the Dual Mode Sorption parameters for the cellulose acetate membranes. In contrast, Table 3.5 parameters were calculated using linear approximation from SABERI *et al.* (2016) to match membrane process mean feed composition.

3.3 Results and Discussion

3.3.1 Diffusion Coefficient Contribution Evaluation

As the gases flow through a membrane module, a concentration profile may usually develop in specific points, especially in the initial stages of permeate side that do not have a feed sweep flow where a degree of back mixing may not be enough yet to form a uniform concentration. Some authors identified this profile as relevant in studies for low-pressure and high-density systems. For instance, SECCHI *et al.* (1999) with albumin ultrafiltration using reverse osmosis, VAN GAUWBERGEN and BAEYENS (2000) with their parametric study for reverse osmosis in spiral wound membranes, and AL-MUTAZ *et al.* (1997) with their study on radial effects contribution for a reverse osmosis desalination process in hollow fiber structures. MARRIOTT (2001) studied the effects of convection, diffusion, and

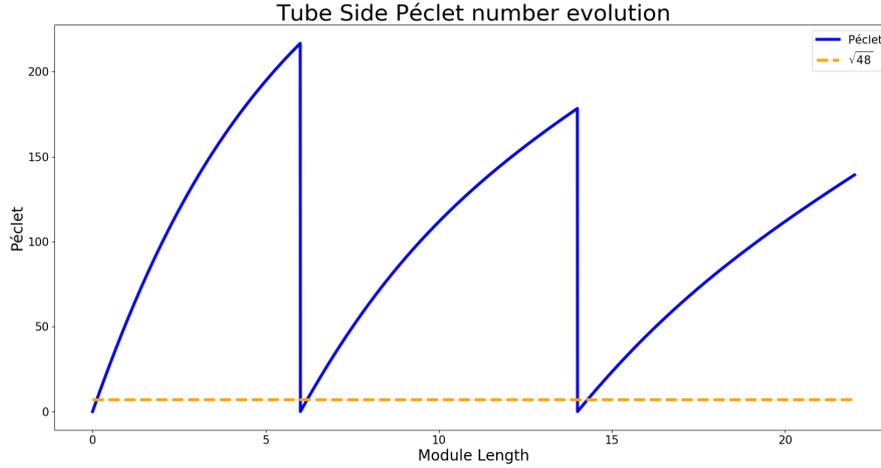


Figure 3.4: Average Péclet number over operation conditions for model tube side.

many other models for gas separation, but the diffusion coefficient was only considered for reverse osmosis structure.

This rate of mixing is measured using the Péclet number, a dimensionless parameter that represents the ratio of advective transport phenomena with diffusive transport, given by Equation 3.40 where v is the fluid speed, l is the characteristic length and D_{AB} the diffusion coefficient (MARRIOTT, 2001).

$$Pe = \frac{vl}{D_{AB}} \quad (3.40)$$

A large Péclet number indicates little mixing will occur, according to BIRD *et al.* (2004). When the Péclet number is less than $\sqrt{48}$ the molecular diffusion does take part in axial mass transport. A Péclet variation study alongside each theoretical module was conducted using Table 3.1 order of magnitude parameters, alongside the operational conditions described by the gathered data.

Figures 3.4 and 3.5 shows the average Péclet number variation from the beginning to the end of each set of modules considering more than 3 thousand feed conditions given by the gathered data. As one can see, it is notable that the Péclet number exceeds by a grand marge the threshold of $\sqrt{48}$ cited by BIRD *et al.* (2004). This suggests that the diffusion coefficient initially intended to be part of this study proposal will not contribute much to the phenomenological model structure. For that reason, another study regarding the actual contribution of the diffusive transport coefficient was carried out.

This change in significance is expected due to the absence of feed permeate flow, where the mass flow is solely provenient from the compound's permeation process. Fluid velocity is too low, thus enhancing mass diffusion significance on

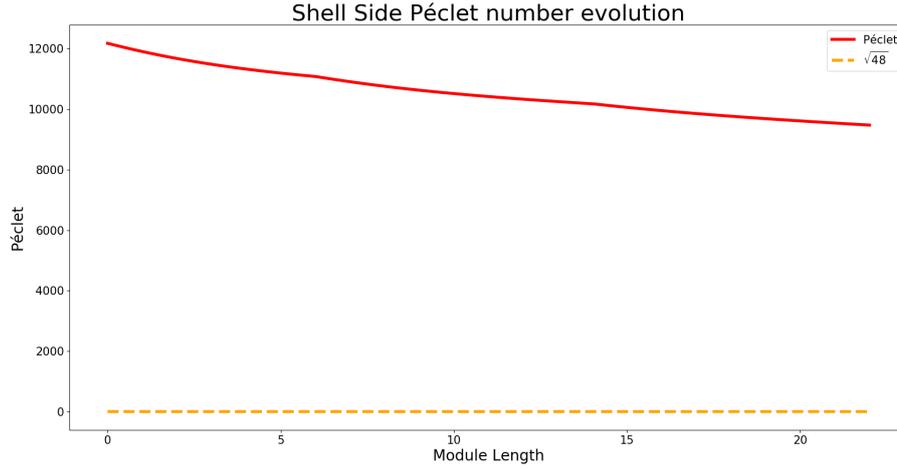


Figure 3.5: Average Péclet number over operation conditions for model shell side.

the mass balance equations. However, this significant participation fastly switches as the evaluation advances through each module. It is only meaningful for less than 5% of each set of modules.

Furthermore, the shell side Péclet behavior is proportionally inverse of the tube side. The values start at the higher Péclet number possible, dropping as mass permeates to the tube side. However, it never reaches any closer to the value where the diffusion is significant in the mass balance equations.

From this study, it can be implied that the diffusion parcel is only significant for less than 5% from the tube side mass balance. That seems reasonable enough and proves that this coefficient does not significantly influence the results from mass balance, and it can be omitted for enhancing computational purposes without losing modeling accuracy.

To sum up this discussion so far, the set of equations used for model validation study and further chapters discussion will be as follows:

$$u_{ziew} = u_{ziw} - \pi D_o N P_p (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta z \quad \text{Shell side} \quad (3.41)$$

$$v_{ziw} = v_{zie} + \pi D_o N P_p (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta z \quad \text{Tube side} \quad (3.42)$$

$$P_{xe}^2 = P_{xw}^2 - \frac{384 N D_o (D + N D_o) R T \mu_m}{\pi (D^2 - N D_o^2)^3 Z_y} \sum_{i=1}^2 (v_{ziP}) \Delta z \quad \text{Shell side} \quad (3.43)$$

Table 3.6: CHU *et al.* (2019) Scenario 1 parameters and feed conditions.

| Parameter/ Condition | Value |
|---|------------------------|
| Temperature [K] | 308 |
| Feed Pressure [bar] | 1 |
| Feed flow [mol/s] | 0.35 |
| Retentate feed flow [mol/s] | 0 |
| CO ₂ feed composition [mol%] | 10 |
| CH ₄ feed composition [mol%] | 90 |
| CO ₂ Permeance [mol/m ² Pa s] | 3.207x10 ⁻⁹ |
| CH ₄ Permeance [mol/m ² Pa s] | 0.133x10 ⁻⁹ |
| Module Inner Diameter [m] | 0.1 |
| Fiber outer diameter [μ m] | 250 |
| Fiber inner diameter [μ m] | 200 |
| Fiber length [m] | 0.6 |
| Number of fibers per module | 60000 |

Source: adapted from CHU *et al.* (2019)

$$P_{ye}^2 = P_{yw}^2 - \frac{256RT\mu_m}{\pi D_i^4 Z_x N} \sum_{i=1}^2 (u_{ziP}) \Delta z \quad \text{Tube side} \quad (3.44)$$

3.3.2 Model Validation 1

For the model validation study, two acknowledged works were chosen. The first one is the work of CHU *et al.* (2019), which studied the influence of module design parameters over the performance of CO₂ separation process in a hydrocarbon mixture. For their study, isothermic mass and momentum modeling was proposed. They also did not consider the site competition, membrane plastification effects, or viscosity change.

CHU *et al.* (2019) studied the influence of design parameters over seven different mixture scenarios. The one used to validate this study was scenario one. Initial conditions and parameters are described in Table 3.6 not only solved by the authors' proposed method (Mollocator equation) but also solved by NTNU membrane separation software Chembrane (GRAINGER, 2007) which resulted in Table 3.7.

As shown in Table 3.7, this work proposed equation relative error against the other studies was lesser than 1%, which indicates that this work proposal is yielding results similar to the ones expected by the literature. Also, for a lower CO₂ fraction on the feed flow, plastification effects and site competition should not significantly affect the results. For that reason, having similar results to the Molloca-

Table 3.7: Model validation.

| Studies | Permeate Flow [mol/s] | CO ₂ fraction on permeate [%mol] | Retentate flow [mol/s] | CH ₄ fraction on permeate [%mol] |
|-------------------------|-----------------------|---|------------------------|---|
| Mollocator ¹ | 0.0298 | 59.54 | 0.3202 | 94.60 |
| ChemBrane ¹ | 0.03 | 59.12 | 0.32 | 94.77 |
| This work | 0.02983 | 59.62 | 0.3201 | 94.62 |
| RE1 [%] | 0.101 | 0.134 | 0.031 | 0.021 |
| RE2 [%] | 0.57 | 0.839 | 0.031 | 0.159 |

Source: (1) adapted from CHU *et al.* (2019)

Table 3.8: Parameters and feed conditions for model validation 2.

| Parameter/ Condition | Value |
|---|----------------------|
| Temperature [K] | 343 |
| Feed Pressure [Pa] | 59.6x10 ⁵ |
| Feed flow [mol/s] | 0.377 |
| Retentate Pressure [Pa] | 1.7x10 ⁵ |
| Retentate feed flow [mol/s] | 0 |
| CO ₂ feed composition [mol%] | 10 |
| CH ₄ feed composition [mol%] | 90 |
| Module Inner Diameter [m] | 0.3048 |
| Fiber outer diameter [μ m] | 300 |
| Fiber inner diameter [μ m] | 150 |
| Fiber length [m] | 1 |
| Number of fibers per module | 500000 |

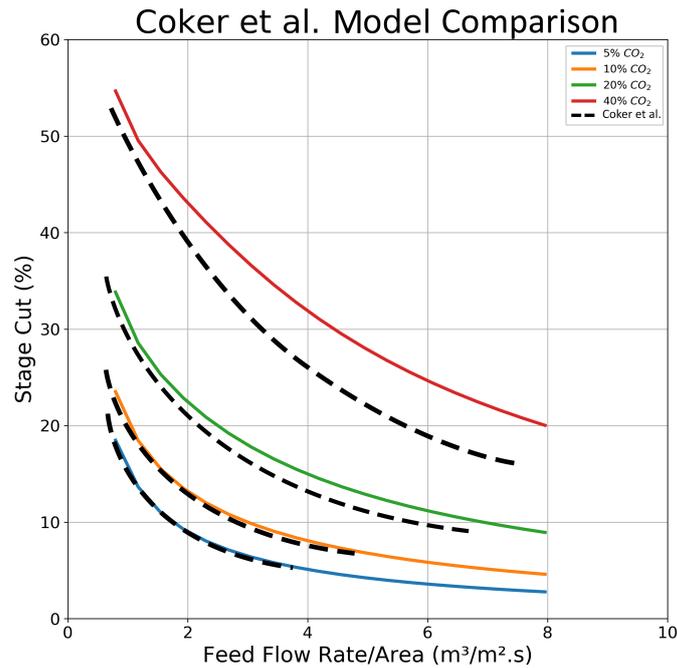
Source: COKER *et al.* (1999)

tor and ChemBrane is expected.

3.3.3 Model Validation 2

A second model validation was proposed. The chosen one was the work from COKER *et al.* (1999), which evaluated the effects of feed flow and membrane area over the stage cut from cellulose acetate hollow fiber membranes for a temperature-independent permeability coefficient. In this part of the study, the authors used a unidimensional equation. However, likewise CHU *et al.* (2019) work, COKER *et al.* (1999) also did not consider site competition or membrane plastification effects for high CO₂ concentration or the effects of temperature. Table 3.8 describes a summary list of parameters and conditions.

Differently from comparing with the last work, similar results are expected only for lower stage cut, where the effects of classification are not significant to promote a gap between the proposed model results and the one presented in the



Source: adapted from COKER *et al.* (1999)

Figure 3.6: Model validation.

literature. Following this expected mismatch for higher values of CO_2 , Figure 3.6 shows the comparison as expected. The dashed line, representing the results of COKER *et al.* (1999), tends to create a gap from this study proposal for higher CO_2 fraction, while it tends to overlay each other for lower CO_2 fraction.

To sum up, considering both model validation accuracy, the model proposed and developed in this chapter is proved to be a plausible and effective alternative and is under what is already established in the literature.

3.4 Conclusion

The development of a phenomenological model is imperative to be used as a benchmark for the next chapters development. The proposed model is a more detailed alternative to a membrane-based natural gas deacidification process. It balances computational costs with accuracy. It benefits from knowing the expected process parameters such as permeability, gas fugacity, membrane plastification rate, and viscosity.

Chapter 4

Multi-layer Perceptron Network development

4.1 Introduction

To achieve the second objective of this study, the permeability factor estimation will be estimated by a neural network model to skip the process of thermodynamic parameters calculation for each finite volume. The core idea is to study and propose a model with reduced computational demand to be implemented online for further process monitoring and optimization. For that process, an MLP neural network structure will be developed in this section and trained with real data process range.

All 4 input variables available, mentioned and treated in Section 4.2.1 (Temperature, Pressure, volumetric flow, and feed composition) will be used to predict the permeability factor for CO_2 and CH_4 compounds. Different structures of an MLP neural networks will be tested in search of the best configuration that gives smaller errors in validation and prediction. Also, a search on the literature was made to verify if similar works on permeability prediction and hybrid models were already developed. As a result, similar research was found, but none had used a hybrid model, which makes the present work the first of it until the date of its publication.

A neural network can be used to estimate critical parameters and predict the output variables even if they are bond to each other by a non-linear relation. Because of that, as a matter of performance comparison, another study regarding the assertiveness of an MLP network regarding the variable's boundary conditions at the process exit is done. The objective of this side study does not contribute directly to any of this study's objectives. However, it enriches this work with another alternative for comparison and discussion if it can be used for real-time evalua-

tion.

4.2 Methodology

In this chapter, the process described by Figure 3.1 and the SABERI *et al.* (2016) equation presented in Section 3.2.6 are considered. The thermodynamic parameters are given by the Peng-Robinson equation of state exposed in Section 3.2.3, and mixture rules detailed in the Appendix D. Those generated data will be used to adjust the weights and biases of a Neural Network to predict the permeability factor and achieve part of this work's second objective.

In Chapter 2, the structure of an MLP is presented. That structure will be optimized and oriented to model assertiveness and trained accordingly to the back-propagation algorithm described in Section 2.4.4.

All numerical algorithms were developed using an 8th generation i7 Intel processor PC with 16Gb of RAM and Python 2.0 programming language at Jupyter compiler using Keras Optuna test generator package for Neural Networks structuring and performance evaluation.

4.2.1 Data evaluation and treatment

Before any neural network development, an in-depth evaluation of the available data quality is required to avoid corrupted data or outliers that may interfere with the parameter estimation algorithm causing even worse neural network weights and biases adjustments. An analysis of the normalized raw feed variables database shown in Figure 4.1 indicates that some data might be corrupted. Please refer to Appendix F for the visualization from the other feature's raw data. Not only negative values for temperature, pressure, and molar percentage were found, but sensor errors with constant values for extended periods were also frequent. Besides, it also shows how this process is well-behaved and fast enough that the sampling time step almost does not catch a maintenance stop and restart dynamics.

All the data was gathered from May 2019 until the end of October of the same year with a sampling time of 1 hour, yielding an amount of 4425 points from each variable (temperature, CO_2 composition, pressure, and volumetric flow). Unfortunately, this FPSO does not have enough instrumentation to analyze CH_4 , SO_2 , or any long-chain hydrocarbons, which would be of great value to this work. And for that reason, some additional assumptions to the ones presented in Section 3.2.1 regarding permeate and retentate streams were made to deal with this missing data, as follows.

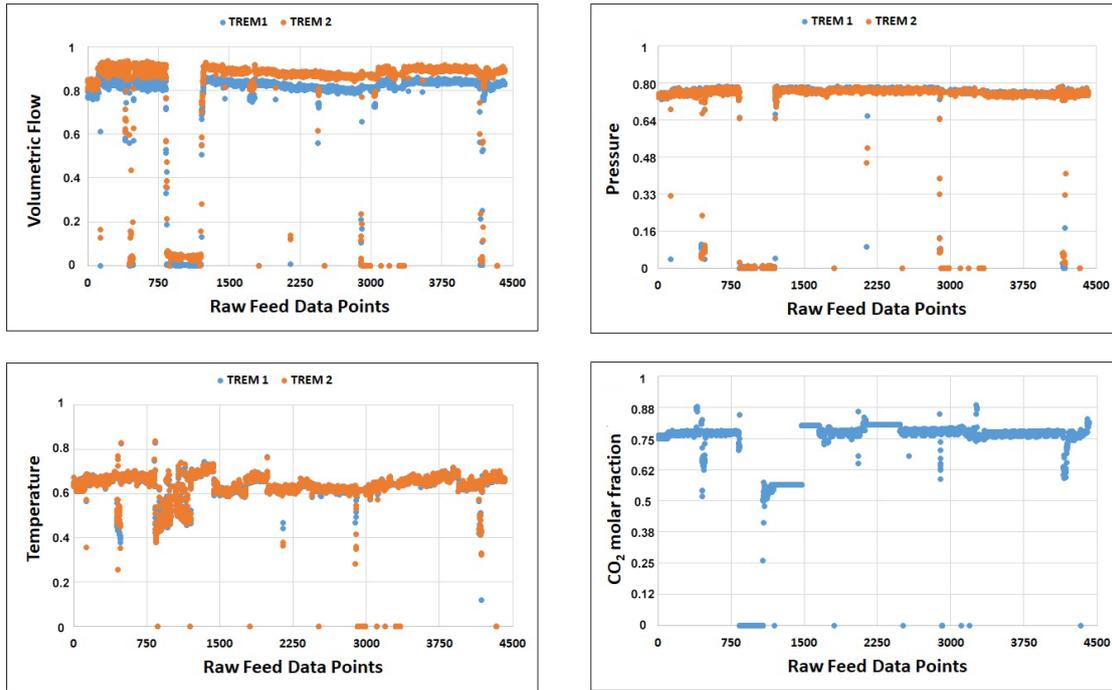


Figure 4.1: Normalized raw data sheet from feed variables. Sampling time of 1 hour.

Hypothesis 11 (H11). *Only the binary mixture CO_2 and CH_4 will be considered in this study. So the molar fraction of CH_4 will be considered as the complementary from CO_2 to reach 100%.*

Hypothesis 12 (H12). *The SO_2 is considered to be in a concentration low enough to meet the regulation standards at the end of the process.*

Hypothesis 13 (H13). *Since the analyzer's minimum sampling time is one hour, the output variables from a given timestamp can be approximated to be the output from the input from the same timestamp, i.e., disregarding the dynamic effects.*

To begin the data treatment procedure, measurement errors and missing data were removed from the series. That included empty values, not-a-number values, and data inconsistent values mentioned in the first paragraph of this section. Those points were also considered measurement errors and therefore disregarded.

After the treatment process, feed parameters were used to calculate each initial permeability factor for both compounds according to the SABERI *et al.* (2016) equation. That yielded around 2900 permeability values used as targets for the neural network. During the period when the data set was gathered, only TREM 1 and 2 were operating. For that reason, only these TREMs were considered.

4.2.2 Neural Network modeling

The strategy to build a neural network was based on the structure of the perceptron as illustrated in Figure 2.14 developed by MCCULLOCH and PITTS (1943) improved by ROSENBLATT (1958) where the network first layer is composed by the exact number of input variables. The last layer will also have as much as the number of output variables, likewise, as HAYKIN (2001) suggests.

All four feed variables were considered input for the first layer (Temperature, Flow, CO₂ composition, and Pressure). For the hidden layers, the same perceptron model will be used. However, Section 4.3 will discuss the best structure of how many hidden layers and how many neurons on each layer. KAVLAKOGLU (2020) published an article considering more than just one hidden layer as a deep learning algorithm. Although this work does not intend to create a deep learning model, it will also be considered if proven useful and more accurate in achieving its objectives. For the output, a 2-node layer will be used to simultaneously predict both compounds' permeability factors.

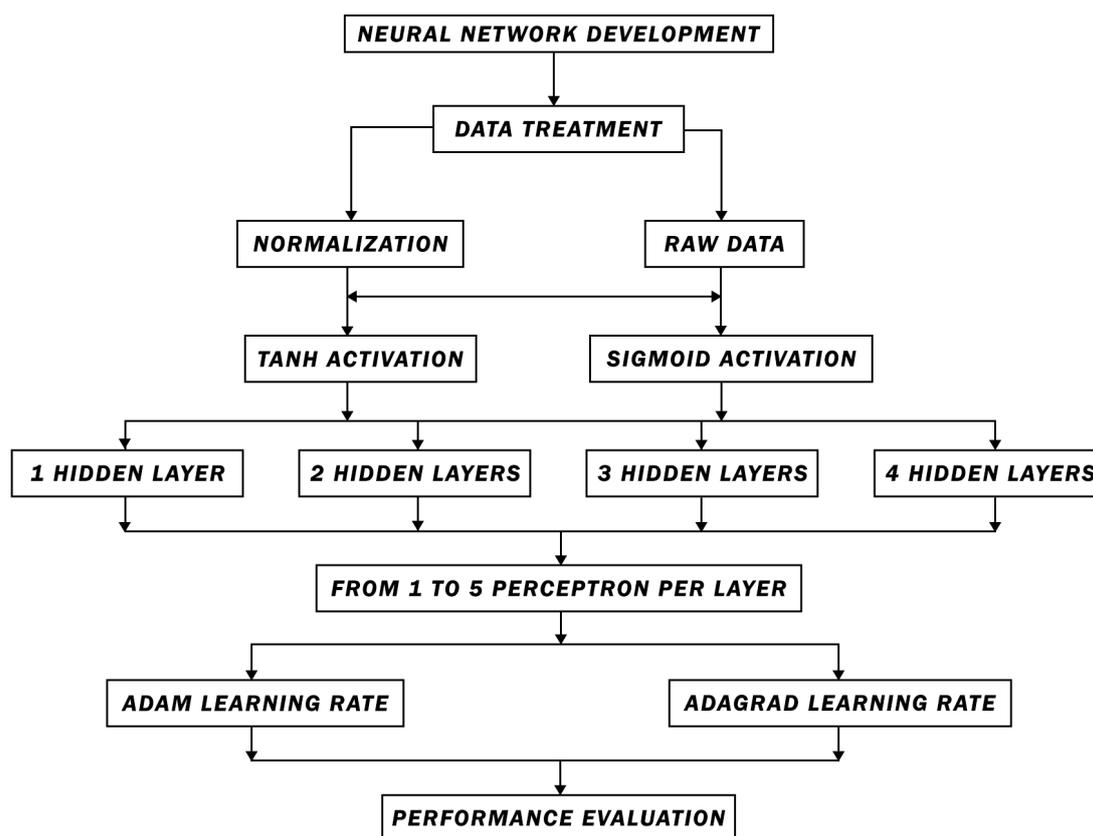
The backpropagation algorithm will be used to update each perceptron's weights, and biases as described by RUMELHART and MCCLELLAND (1986) and by Equations 4.1 and 4.2. However, the learning rate parameter will not use the fixed step to minimize overshooting and oscillation possibilities. The algorithms developed by KINGMA and BA (2015) (ADAM) and DUCHI *et al.* (2011) (AdaGrad) will be used instead and compared since one is a stochastic gradient-based algorithm, and the other is an adaptive gradient method that takes into account the function profile from previous interactions to adapt the learning rate. Different activation functions will also be tested.

$$\Delta w_{j-1}(t) = -\eta \nabla Obj(t)_j \quad (4.1)$$

$$Obj(t) = \frac{\|r(t) - y(t)\|^2}{2} \quad (4.2)$$

This search for the best neural network configuration will be done in two steps. The first one is a manual search to understand how the MLP Network will perform for different non-structural parameters, such as:

- Normalized features x Raw features;
- While normalizing, which normalization strategy is better;
- What is the maximum training epochs without producing over-fitting;



Source: developed by the author.

Figure 4.2: Roadmap for the search of key parameters non-structural from the manually-set Neural Network.

- Which optimizer algorithm performs better in terms of accuracy and evaluation speed.

For the manual search of those hyperparameters, Figure 4.2 visually represents a general strategic workflow for developing the neural network structure.

4.2.3 Architecture search

A full network architecture search using the test-trial package Optuna[®] was used to optimize the best set of hidden layers and nodes as well as corresponding weights and biases to achieve the second objective of this study which is to find a neural network model to predict the permeability parameters for both compounds studied and serve as an alternative to avoid computational costs with state equations evaluation every finite volume.

For this part of the study, more data was necessary. This architecture search aims to create a model capable of predicting the permeability factors for both compounds inside the range of operation the dataset indicated. Because of that, syn-

thetic data was necessary to simulate an actual process range of parameters and the evolution of the permeability factors among a wider range of operations.

In total, 10000 data were generated using a 10-value range between each input variable minimum value to the maximum value and divided randomly for training, validation, and testing in the proportion of 3:3:1. A set of trials was scheduled, limiting up to 5 hidden layers and up to 5 nodes per layer. Also, the number of training epochs was limited to a maximum of 4000, and the loss function was kept as the mean squared error. The stop criteria were set for a mean squared error resulting in 10^{-4} or 100 trials, which came first.

4.2.4 The Hybrid Model

The Hybridization proposed in this section consists of the replacement of the thermodynamic calculations used to determine the mixture key coefficients that are needed in the Permeability model developed by SABERI *et al.* (2016), by the Neural Network model developed in this section.

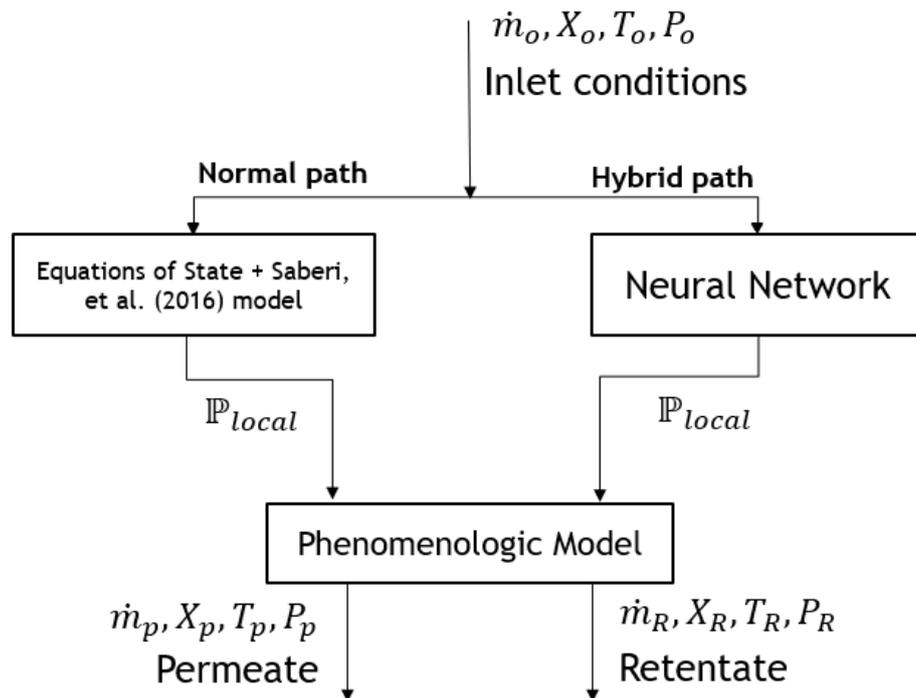
The core idea is to fasten the phenomenologic model evaluation. Furthermore, the hybridization process proposed does not make the model totally unaware of the process physics and restrictions. After the hybridization, a performance comparison is made to sum up and ponder the proposed model's benefits against the purely phenomenological model presented in the last chapter. Figure 4.3 illustrates how the hybridization described will be done.

In this section, the phenomenologic set of equations proposed and validated in Chapter 3 was used with the process conditions accordingly to the study of COKER *et al.* (1999) for the performance evaluation of the proposed models. For that, five key parameters were analyzed: permeate flow [mol/s], permeate CO_2 molar fraction [%], retentate flow [mol/s], retentate CH_4 molar fraction [%], and the time that the model needed to render the output variables. For this study, a 100-finite-volume mesh was used as the number of discrete volumes.

4.3 Results and Discussion

4.3.1 Data Treatment

As discussed in the methodology section, after the data treatment, almost 1500 points were disregarded from the 4400 total points initially available, which is more than one-third of invalid values. The first insight gathered from this treatment procedure is that if any of this work's findings come to be implemented in any membrane-based gas-sweetening process. Data filtering and cleaning must



Source: developed by the author.

Figure 4.3: Comparison between the proposed pure phenomenological modeling workflow and the Hybrid process.

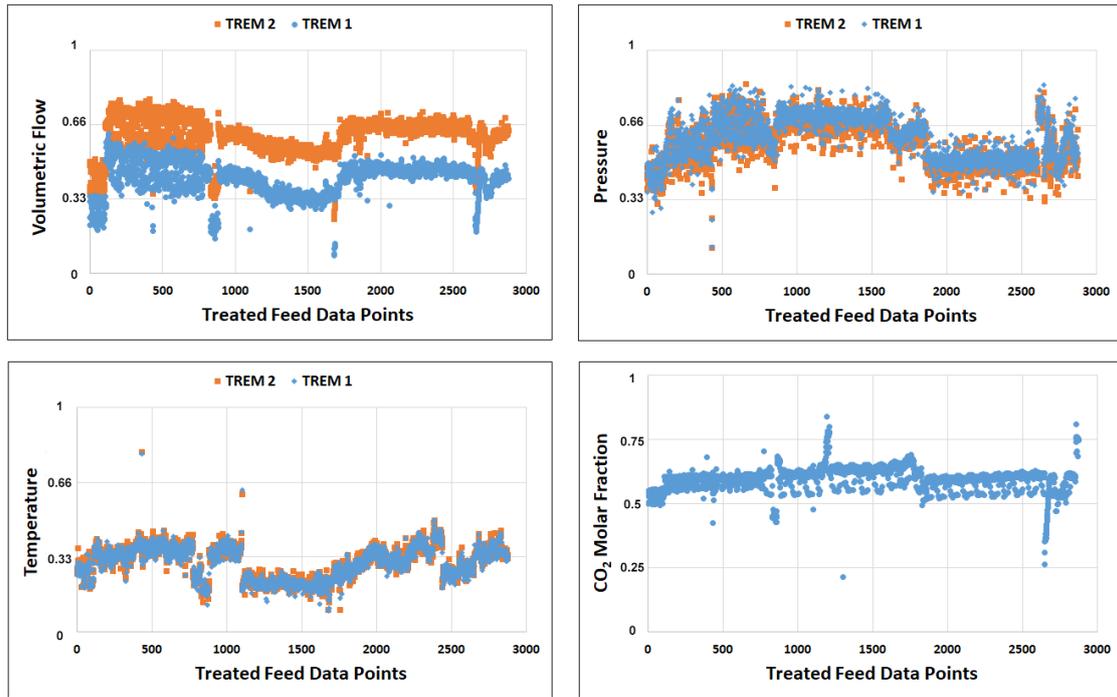


Figure 4.4: Normalized treated data dispersion from feed variables.

be done before model evaluation.

As a result, almost 1500 points were excluded, and the results from normalized feed variables are shown in Figure 4.4. Please refer to Appendix F for the results of the other features filtering results. For the retentate and permeate values, it is worth noting that the CO_2 analyzer is only available on the main lines of the process and corresponds to the combined TREM 1 and 2 flowrates. To promote better visualization of the process behavior, maximum and minimum values for each normalization procedure were fixed by variable type.

4.3.2 Neural network Parameter Study

Table 4.1 shows the differences between the network accuracy (represented by the average correlation factor from validation data) and the average total time required for the training step from using raw data without normalization and normalized training data using a min-max scaler. In this study, activation functions were fixed at the logistics equation (sigmoid), and the number of epochs (interactions over the entire data set) was fixed at 1000. Optimizers were also fixed to the ADAM algorithm in both training sessions.

Normalized input parameters proved more efficient and faster, with an average time difference of 3 s per training, just as mentioned by SOLA and SEVILLA (1997). Besides, the most relevant result was that normalized data yielded a far more accurate model adjustment than raw data and looked much more promising

Table 4.1: Model adjustment comparison between raw data and normalized data training sets.

| Compounds | Raw Data | | Normalized Data | |
|--|-----------------|---------|------------------------|---------|
| | CO_2 | CH_4 | CO_2 | CH_4 |
| Method | - | - | Min-Max | Min-Max |
| Activation | Sigmoid | Sigmoid | Sigmoid | Sigmoid |
| Epochs | 1000 | 1000 | 1000 | 1000 |
| Optimizer | ADAM | ADAM | ADAM | ADAM |
| Av. R^2 Validation | 0.2498 | 0.2368 | 0.9510 | 0.9262 |
| R^2 Maximum | 0.7833 | 0.7961 | 0.9992 | 0.9991 |
| Training time (s) | 17.87 | 17.87 | 14.10 | 14.10 |

Table 4.2: Model adjustment comparison between ADAM and ADAGRAD optimizer algorithms.

| Compounds | CO_2 | CH_4 | CO_2 | CH_4 |
|--|---------|---------|---------|---------|
| Activation | Sigmoid | Sigmoid | Sigmoid | Sigmoid |
| Epochs | 800 | 800 | 4000 | 4000 |
| Optimizer | ADAM | ADAM | ADAGRAD | ADAGRAD |
| Av. R^2 Validation | 0.9510 | 0.9262 | -0.0060 | 0.0033 |
| R^2 Maximum | 0.9992 | 0.9991 | 0.1468 | 0.0645 |
| Training time (s) | 14.10 | 14.10 | 58.97 | 58.97 |

as shown on Table 4.1.

In addition, an additional study on the optimizer was also made to identify the one with the best performance. Table 4.2 summarizes the milestones that indicated the ADAM algorithm yielded the best model adjustment. Only with 800 Epochs could the ADAM structure update the weights with better adjustment than the ADAGRAD algorithm, which had issues optimizing the mean-squared objective function using the sigmoid function. Each ADAGRAD model adjustment was evaluated from 800 to 4000 epochs, and only at 4000 was the loss between the predicted and real values presented no longer significant variation (order of 10^{-6}). Because of that, epochs evaluation was stopped at 4000.

However, as one can also see in Table 4.3, ADAGRAD algorithm, which shows the adjustment correlation from different network structures, showed a good performance adjusting the model with a hyperbolic tangent activation function only for a few network structures.

In addition, the study presented in Table 4.4 aimed to identify the activation function which yielded the best model adjustment. As a result, the logistic function yielded a better maximum R^2 . However, the hyperbolic tangent activation function resulted in faster training time (almost half) and better average R^2 . Besides the sigmoid function achieving higher maximum accuracy, the hyperbolic tangent study characteristics suggested that different network structures could be

Table 4.3: ADAGRAD performance with a hyperbolic tangent activation function.

| R^2 Validation | | | | | | |
|------------------|---------------|---------------------|---------------------|--------|--------------------|--------|
| Compound | Hidden layers | Nodes per layer | | | | |
| | | 1 | 2 | 3 | 4 | 5 |
| CO ₂ | 1 | -0.0092 | 0.3054 | 0.3370 | 0.3691 | 0.3407 |
| | 2 | -0.0612 | 0.0038 | 0.2639 | 0.8923 | 0.8470 |
| | 3 | -0.1451 | -3x10 ⁻⁴ | 0.0049 | 0.2718 | 0.9185 |
| | 4 | -0.0656 | -5x10 ⁻⁴ | 0.9050 | 1x10 ⁻⁴ | 0.8844 |
| CH ₄ | 1 | 0,0051 | 0.2116 | 0.269 | 0.2872 | 0.3181 |
| | 2 | 4x10 ⁻⁴ | 0.0028 | 0.2571 | 0.8796 | 0.8387 |
| | 3 | -6x10 ⁻⁴ | -3x10 ⁻⁴ | 0.0031 | 0.3047 | 0.9147 |
| | 4 | -6x10 ⁻⁴ | -4x10 ⁻⁴ | 0.5283 | -0.4819 | 0.8799 |

Table 4.4: Model adjustment comparison between Tanh and Sigmoid activation functions.

| Compounds | CO ₂ | CH ₄ | CO ₂ | CH ₄ |
|----------------------|-----------------|-----------------|-----------------|-----------------|
| Activation | Tanh | Tanh | Sigmoid | Sigmoid |
| Epochs | 800 | 800 | 800 | 800 |
| Optimizer | ADAM | ADAM | ADAM | ADAM |
| Av. R^2 Validation | 0.9951 | 0.9863 | 0.9510 | 0.9262 |
| R^2 Maximum | 0.9991 | 0.9985 | 0.9992 | 0.9991 |
| Training time (s) | 7.93 | 7.93 | 14.10 | 14.10 |

used without relying much on accuracy and be twice as fast as the logistic activation function network. For these reasons, the hyperbolic tangent equation seemed to be the best option.

To summarize, Table 4.5 shows a case study result on NN structure performance using ADAM optimizer and logistics activation function. As one can see, many different structures performed with at least a 0.999 correlation coefficient.

Figure 4.5 shows the loss behavior as the network is trained. As one can see, there is a sub-decay in the first stages of training, expected as the learning rate is adaptive. So, it is customary to be faster at the beginning and slower as the algorithm approaches the objective function minimum. For this step, 90% of the 2960 data were used for training and 10% for validation.

As expected, the training yielded a strong adjustment of 0.9962 for the CO₂ and 0.9973 for CH₄. Next, to evaluate the network's performance, brand new data set was gathered from Nov-2019 up to Feb-2020 with 3 thousand more points. The same data treatment described in Section 4.2.1 was made, and the trained network was evaluated.

At first glance, it is clear that the network performed well, with predicted values reasonably close to the data. The correlation coefficient of 0.9883 for CO₂ values and 0.9879 for CH₄.

Table 4.5: ADAM optimizer performance with a sigmoid activation function.

| R^2 Validation | | | | | | |
|------------------|---------------|-----------------|---------|---------|---------|---------|
| Compound | Hidden layers | Nodes per layer | | | | |
| | | 1 | 2 | 3 | 4 | 5 |
| CO ₂ | 1 | 0.99953 | 0.99472 | 0.99905 | 0.99925 | 0.99947 |
| | 2 | 0.99705 | 0.99566 | 0.95260 | 0.99678 | 0.99929 |
| | 3 | 0.78250 | 0.93294 | 0.98023 | 0.99778 | 0.99083 |
| | 4 | 0.64949 | 0.82321 | 0.97191 | 0.97437 | 0.98355 |
| CH ₄ | 1 | 0.99854 | 0.99289 | 0.99734 | 0.99827 | 0.99855 |
| | 2 | 0.99542 | 0.98993 | 0.95226 | 0.99692 | 0.99916 |
| | 3 | 0.64675 | 0.94122 | 0.98171 | 0.99767 | 0.99075 |
| | 4 | 0.30916 | 0.80408 | 0.97383 | 0.97728 | 0.98223 |

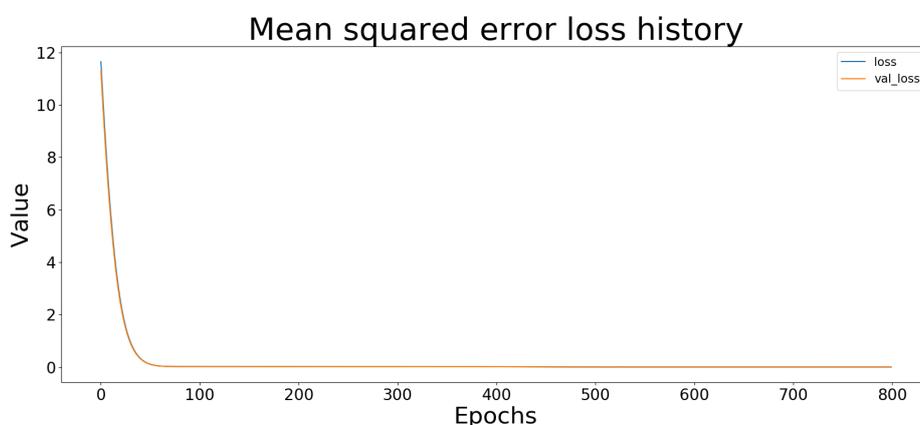


Figure 4.5: Loss behavior as the network is trained.

To sum up, this study confirmed some key parameters that served as simplifications for the following study of architecture search. For instance, the manual roadmap has shown that the ADAM optimizer and gradient search for the optimal weights and biases are the best approaches for this type of data and non-linearities prediction. Also, the number of training epochs can not be more than 4000, and 800 epochs are enough to reach the desired accuracy. The activation function can be both Tanh or Sigmoid approach. However, the Tanh model was twice as fast as the sigmoid model without relying too much on accuracy, as the maximum correlation coefficient is slightly lower than the sigmoid case study. Therefore, Tanh is a better model for reaching the hybridization objectives.

4.3.3 Neural network architecture search

For the best architecture, two approaches were used: a model that can predict both permeances at once or two models that can predict one compound's permeability. This will make sense in the next section, where different models and ways of calling them are studied to produce a faster result.

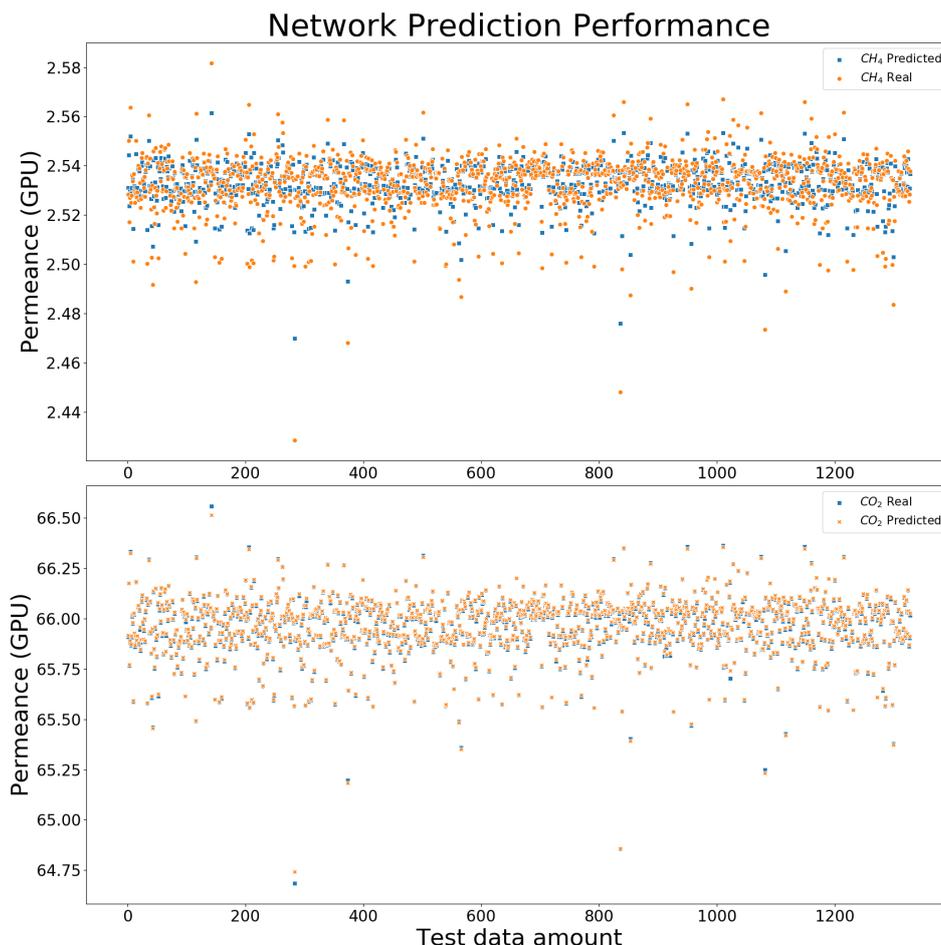


Figure 4.6: Permeance prediction result for the model with two output layers.

The first approach of a single model to predict both permeances resulted in the following architecture: 4 hidden layers using the configuration of 2x2x5x5 nodes on each layer, respectively. With 2000 epochs, it yielded a validation mean squared error of 0.0108. The model's performance against the test dataset shown by Figure 4.6 produced an R^2 coefficient of 0.9979 for the CO_2 and 0.9658 for CH_4 . It took 107 s for this model to be trained and less than 2 s to predict all the 1400 test values, which is a desired result for a neural network on both accuracy and evaluation speed.

The second approach was to develop two models to predict each compound's permeance separately. For comparison purposes, two single-output networks were proposed, and their architecture was searched using the same method as the two-output network search, with more restrictions.

As the single-output model is supposed to be simpler than the two-output model, and the best structure of the two-output model was with four hidden layers, the maximum number of hidden layers tried was 2 with a limit of 5 nodes per layer, with the same optimizer, activation function, loss function and maximum

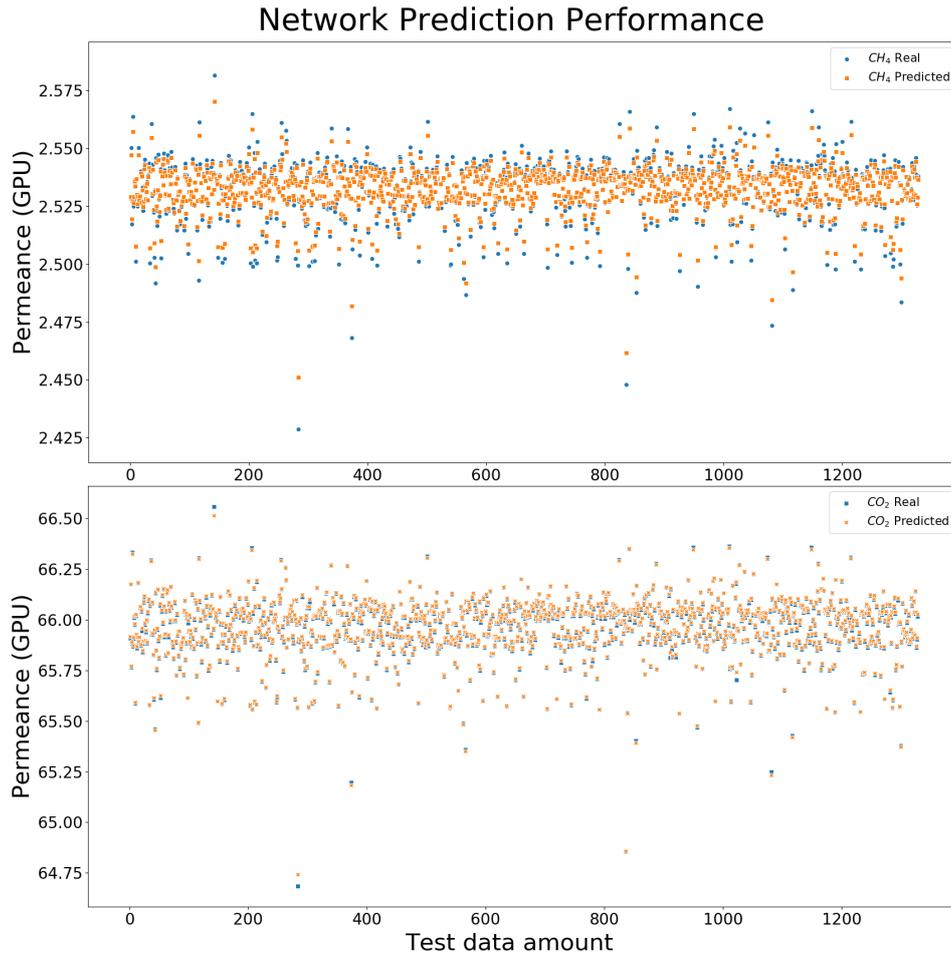


Figure 4.7: Permeance prediction result for the two models with one output layer each.

training epochs of 4000.

Figure 4.7 shows the prediction of the test dataset for the one-output network model. The result of this study was a model with one hidden layer, four nodes each, and 2000 epochs of training. Both models resulted in the same structure of one hidden layer with four nodes. They have had the training session mean squared error of 10^{-5} magnitude and an R^2 adjustment of 0.9968 for the CO_2 and 0.9752 for CH_4 . It took around 60 s for these models to be trained and less than 2 s to predict all the 1400 test values.

The results for the single-output network model are similar to the ones produced by the two-output network, especially on accuracy and speed with a more straightforward structure. For that reason, different types of algorithm hybridization using the neural networks found in this section are tested in the next section.

4.3.4 Hybridization performance evaluation

The hybridization process only replaced the thermodynamic equations algorithm with the neural networks callback on the finite volume solver for the phenomenological model. Table 4.6 shows the performance of each hybridization method in the face of the purely phenomenological model performance.

The models were abbreviated for Table 4.6 size. Their meaning and characteristics are as follows.

- **Phenom.** - The purely phenomenological model used as the benchmark for the hybrid models
- **Hybrid 1A** - The hybridization consisted of loading the two-output network in place of the thermodynamic equations for permeance prediction for each discretized finite volume.
- **Hybrid 2A** - The hybridization consisted of loading one-output networks in place of the thermodynamic equations for permeance prediction for each discretized finite volume.
- **Hybrid 1B** - The hybridization consisted of remounting the two-output network in place of the thermodynamic equations for permeance prediction for each discretized finite volume.
- **Hybrid 2B** - The hybridization consisted of remounting one-output networks in place of the thermodynamic equations for permeance prediction for each discretized finite volume.

It is also important to note that the terms "*loading*" and "*remounting*" have distinct meanings. To load a model means to import a model package straight to the compiler. The weights, biases and structure information are loaded at the same time using the TensorFlow backend interpreter. Conversely, to remount a model means to pass to the TensorFlow interpreter an array of weights and biases, then the neural network structure is remounted based on the array passed in the TensorFlow backend.

From the results, two major conclusions can be made out. The first is that the neural network can be used as an alternative to certain parts of the phenomenological models. In this case, 3 out of 5 resulted in average errors of less than 2% for both approaches, one and two output models. The second finding is the lack of speed of the hybrid model against the purely phenomenological model. However, models 1B and 2B indicate that the most time loss is related to the need for network callbacks for every discretized finite volume.

Table 4.6: Hybrid models and their comparison with the purely phenomenological model proposed on Chapter 3

| Model | Perm. Flow [mol/s] | CO₂ Perm. Frac. [%] | Ret. Flow [mol/s] | CH₄ Ret. Frac. [%] | Eval. Time [s] | Avg. Rel. Error [%] |
|------------------|-------------------------------|---|------------------------------|--|---------------------------|--------------------------------|
| Phenom. | 8.662 | 75.11 | 8.168 | 97.24 | 0.16 | - |
| Hybrid 1A | 8.474 | 76.62 | 8.356 | 97.13 | 13.85 | 1.65 |
| Hybrid 2A | 8.498 | 76.41 | 8.333 | 97.12 | 27.53 | 1.44 |
| Hybrid 1B | 8.327 | 77.84 | 8.502 | 97.06 | 10.57 | 2.94 |
| Hybrid 2B | 8.502 | 76.17 | 8.310 | 97.08 | 10.50 | 1.23 |

Besides, in the previous section, this study showed that the network models could predict more or less 1400 points in less than 2 s, while when hybridized, it could not predict 100 of them in less than 10 s. This also confirms the conclusion that the amount of network model callbacks is the main reason for the evaluation delay. To solve this issue, the number of callbacks should be reduced to the minimum possible (1 time only) to extract the maximum potential of the hybrid model.

Because of that, another hybridization approach is proposed as an alternative to decreasing the number of network callbacks. It is called Physics Informed Neural Networks (PINNs) and is further developed in the next chapter.

4.4 Conclusions

This chapter discussed the effective use of neural networks to predict the permeability factors to replace the thermodynamic evaluation of the phenomenological model proposed in Chapter 3 to achieve better model performance. It showed that a simple multi-layer perceptron network with one hidden layer can predict the permeability factor with a correlation coefficient of more than 0.96. However, data-cleaning treatments and normalization are crucial to achieving these results.

Furthermore, it was shown that the hybridization problem consists of the number of network callbacks made. The more network callbacks are made, the more the loading process happens. This hinders the model performance from becoming 100 times slower than the purely phenomenological models. For real-time monitoring, it would depend on the required sensors' sampling time. If the sensors are to sample in less than 15 s, this hybridization is not the best approach.

Chapter 5

Physics-Informed Neural Network

5.1 Introduction

To complement this study's third objective and complete its final one, data-driven and physics-informed neural network models are proposed to replace the phenomenological model. This replacement intends to use real-world data and to keep physical constraints without compromising accuracy and efficiency. Data-driven models learn patterns in data and make predictions based on that information. Physics-informed models, on the other hand, use physical laws to constrain the solution space and ensure that the model outputs are physically meaningful.

Physics-Informed Neural Networks (PINNs) is a type of neural network that incorporate physical laws and constraints into the model. These networks can be used to solve problems where traditional machine learning techniques may not be effective, such as problems in which the underlying physical laws are known but challenging to formulate mathematically. The approach consists of defining a loss function based on the physical laws and constraints, in addition to the data, and then training the neural network to minimize this loss function (RAISSI *et al.*, 2019).

One advantage of PINNs is that they can handle both forward and inverse problems. For example, they can be used to predict the future state of a system given its initial state or to determine the initial state of a system given its final form. Additionally, they can handle nonlinear and partial differential equations, making them well-suited for modeling complex physical phenomena.

In summary, PINNs provide a way to leverage the power of neural networks to solve problems guided by physical laws, making them valuable tools for various applications in fields such as engineering, physics, and materials science.

5.2 Methodology

Gas separation data will be generated through the phenomenological modeling proposed in Chapter 3. The measurements gathered will be used as a source of information from the operating range that the actual process undertook over the year the data was collected. The generated data will then train a physics-informed neural network and data-driven models.

A gaussian noise of a maximum of 1% will be added to the synthetically generated data to account for measurement uncertainties and other sources of error. This will help to improve the robustness and generalizability of the trained model. The end goal is to produce a model that can accurately predict gas separation outcomes based on the input data and physical constraints.

The PINN network will be designed to incorporate physical constraints and equations that govern the behavior of the gas mixtures. PINN and data-driven networks will then learn to predict the gas separation outcome based on input data and physical constraints during training. All coding algorithms were developed using an 8th generation i7 Intel processor PC with 16Gb of RAM and Python 3.0 programming language at Jupyter compiler using Keras packages Neural Networks structuring and HAGHIGHAT and JUANES (2021) SCIANN computing package for PINN development and performance evaluation.

For both approaches, a Multi-Layer Perceptron network structure will be used.

5.2.1 The loss function

Data-driven and physics-informed neural network models are distinct approaches to modeling complex physical systems. Both approaches have unique loss functions to evaluate the model's performance and guide the optimization process during training. Understanding the differences between these loss functions is crucial for choosing the appropriate model for a given problem and interpreting the model's results.

A data-driven model is trained by optimizing a loss function that measures the difference between the model's predictions and the observed data. The loss function commonly used in data-driven models is the mean squared error (MSE) sum function, which measures the average squared difference between each model's prediction variables and their true values represented by Equation 5.1. The optimization process tries to minimize the MSE sum by adjusting the model's parameters until the difference between the predictions and the true values is as small as possible, as presented in section 2.4.4 objective function. The data-driven model aims to fit the data as closely as possible without considering any underly-

ing physical constraints or equations.

$$Loss = MSE = \frac{1}{N} \sum_{i=0}^N |Target_i - Prediction_i|^2 \quad (5.1)$$

On the other hand, a physics-informed neural network model is trained by optimizing a loss function that incorporates the observed data and the underlying physical constraints or equations. The loss function measures the difference between the model's predictions and the observed data and the deviation from the physical constraints. The optimization process tries to minimize this combined loss by adjusting the model's weights and biases. The goal of the physics-informed neural network model is to produce a model that not only fits the data well but also respects the underlying physical principles.

The Equation 5.2 represents how the loss function for the PINN will be used. It is a weighted sum of mean squared errors from the following key parameters: data (dat), residuals (Res), outlet conditions (OC) and inlet conditions (IC).

$$Loss = \sum_{i=0}^N [w_1 MSE(Res)_i + w_2 MSE(OC)_i + w_3 MSE(IC)_i + w_4 MSE(dat)_i] \quad (5.2)$$

The data targets will be the normalized features given by simulated data added by 1% random noise generator. The residuals target will be given the phenomenological model from Chapter 3 without noise. However, the boundary conditions MSE is calculated only at the beginning and end of each dataset using the first and last positions from the data dataset, and set to zero for the other positions. The loss equation weights were set to 1 at first and then study their influence during the studies that followed this chapter.

The PINNs and data-driven models will be set to predict eight features (each gas separation unit side molar flowrate for CO_2 and CH_4 , each side pressure buildup and drop end each compound permeability factors) while using three input features (inlet molar flowrate for both compounds and the inlet shell-side pressure), apart from the module position argument, according to the roadmap of evaluation described in Figure 5.1.

5.2.2 Data generation

For the data generation process, the relative position of each discretized finite volume of the membrane modules was considered, the operating ranges of the CO_2 and CH_4 molar flowrate variables and the operating range of the shell pressure.

First, the performance of the proposed models considering only one mem-

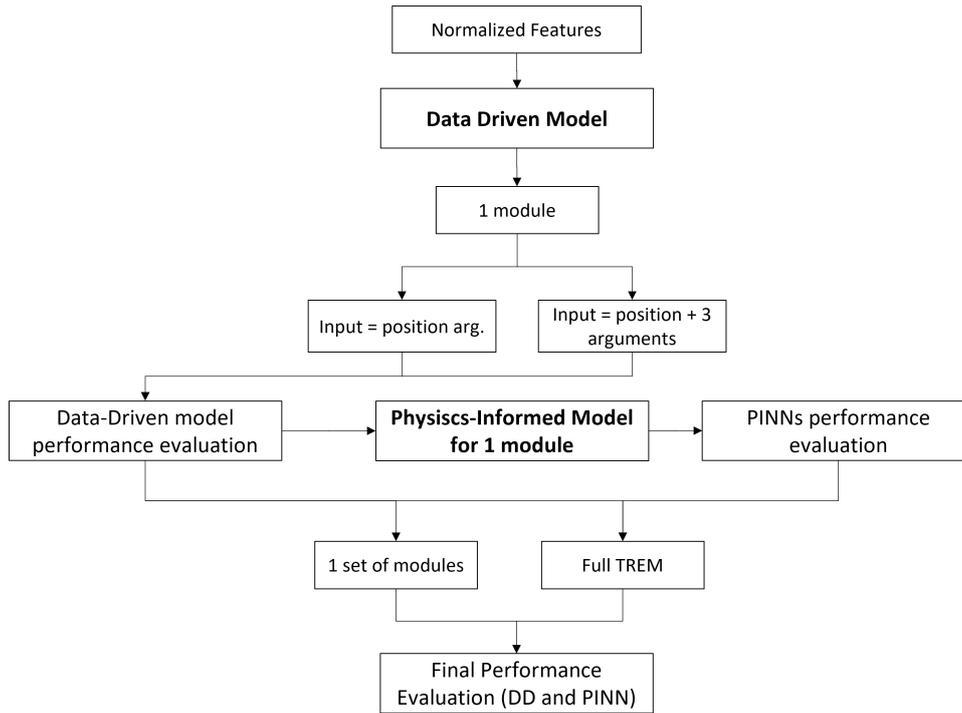


Figure 5.1: Roadmap for the PINN and data-driven model development and performance evaluation.

brane module was studied. For this first study, a 400 finite volume discretization was performed, and a range of values equally spaced over a specific range for each key variable, according to the following: nineteen values for the CO_2 molar flowrate, ten for the CH_4 molar flowrate, and thirteen for the shell pressure. The total number of points generated for training and validation of the models was 1,482,000. Since the operating range gave the values chosen from the dataset gathered from the actual process, those values are also covered in the non-disclosure agreement.

For the second part of this chapter, the performance of the models in the digital twin of the industrial process used as a basis for this dissertation was studied. In this phase, the same ranges of operation of the dataset of one year of operation of the industrial process were used as the benchmark. In this second step, 400 finite volumes for each TREM (please refer to the Figure C.1 in Appendix C) were used as discretization, totaling 1200 finite volumes in each operating condition. Additionally, a range with nineteen values for the CO_2 molar flowrate, ten for the CH_4 molar flowrate, and nineteen for the shell pressure were used. In total, the number of points generated for training and validation of the models was 4,332,000.

In addition, a noise signal was added, which was performed on the entire dataset, excluding the points of the boundary conditions. The added noise value was -1% to 1% of the value of each point randomly and with normal distribution.

5.2.3 PINN and Data-Driven models structure and parameters

For the structure of the neural networks for each proposed model, some parameters of the optimization algorithms and updating of the network's weights and biases were fixed according to knowledge acquired in the previous chapters. Thus, the following parameters were set:

- The activation function was set to Hyperbolic Tangent;
- The optimizer was set on the ADAM algorithm;
- The number of training batches was set to 500;
- The number of nodes per layer was fixed at 10, only varying the number of layers per model created;
- The objective function was set to MSE;
- Training and validation data were divided as 80% and 20% from the total dataset respectively;
- Test dataset inlet conditions will be given randomly and bounded to each feature range after the training process;
- For the data-driven and PINN models, a single MLP network will be given to predict each feature.

The hyperparameters available to be modified were the amounts of internal layers for both models, the number of data per batch, the initial learning rate and rate of decay for each training performed, the weights of the loss function, the normalization method and the criteria for abrupt stopping of the training process.

5.3 Results and Discussion

5.3.1 One module data-driven model

- **1 input parameter: Module discrete positioning**

As part of obtaining knowledge for better modeling and training strategy, first, an MLP network containing two internal layers was trained using only the position parameter of the module as input for each neural network. In this first step, the min-max normalization strategy was also used, where the values of each variable for each finite volume are normalized between the minimum and maximum values of the entire dataset for that variable. This first study aimed to understand

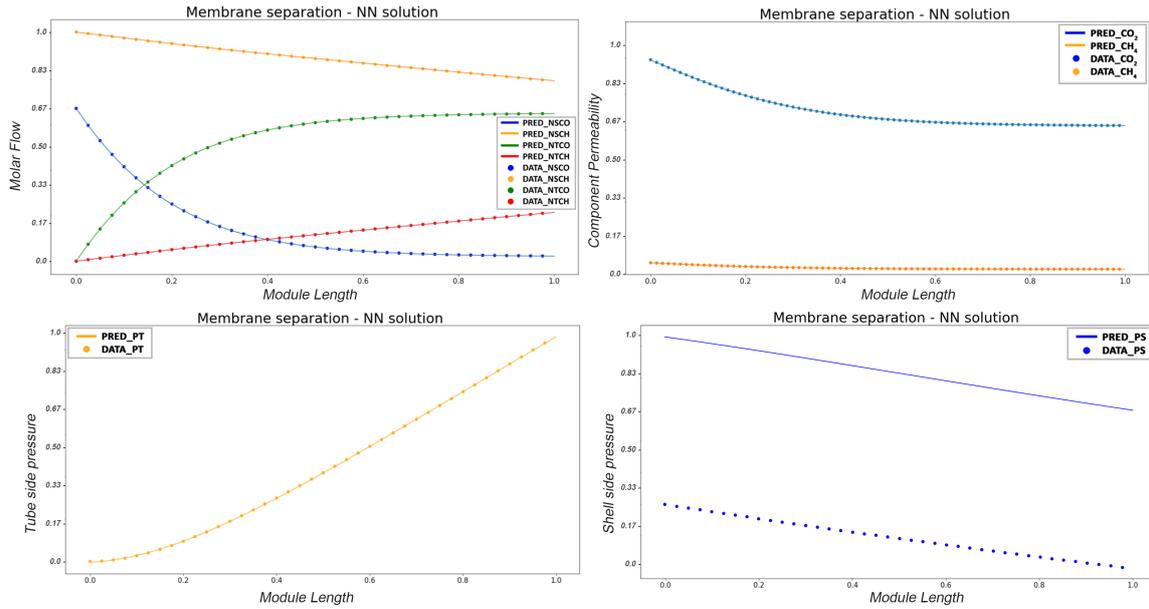


Figure 5.2: First results for molar flowrates, permeabilities, and side pressure evolution for the data-driven model using one input and one output for each network.

if only the module position parameter and a low number of internal layers were enough to accurately predict the evolution of the variation of the features along the module.

As a result of this first study, the data-driven net behaved very well when predicting the parameters of molar concentration and permeability of CO_2 and CH_4 as well as tube side pressure buildup. However, for the prediction of shell pressure drop, there was almost no fitting, as presented in Figure 5.2, for an inlet condition never before seen by neural nets, in comparison to its theoretical result using the phenomenological equation that gave origin to the training data.

The training procedure for this data-driven model configuration took, on average, 1 h to 1.5 h to complete; and the time this model took to evaluate the test dataset was 1.06 s.

A mismatch in pressure prediction results within the module is perceived. Consequently, it was decided to change the number of internal layers in the pressure model structures from 2 to 4, increase the learning rate from 0.01 to 0.05, and increase the amount of training data provided in each training batch from 512 to 1024.

These modifications were not carried out all at once, they were done in isolated steps and different combinations and comparing the results. However, the configuration that resulted in the lowest MSE value for the data-driven network was the one with the configurations mentioned above, and presented in the results below in Figure 5.3. However, it is still possible to see a slight offset in the shell pressure

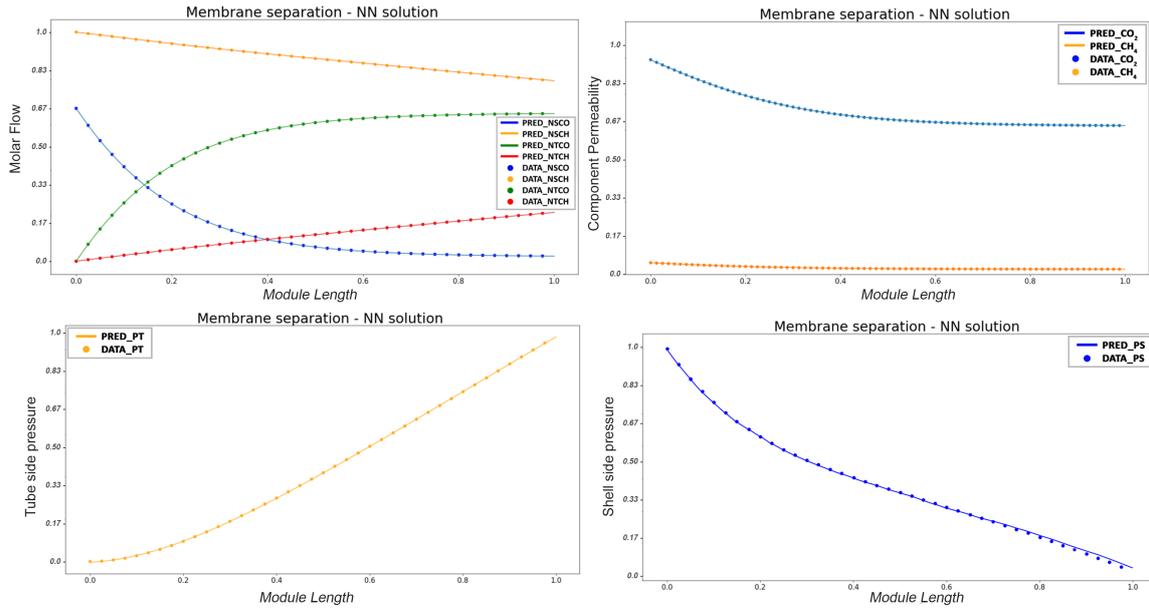


Figure 5.3: Best results for molar flowrates, permeabilities, and pressure evolution for the data-driven model using one input and one output for each network.

variable, although, in proportional terms, this offset represents less than 1% in the final deviation, it is still a more significant misfit than the other variables.

From this study, it can be understood that it is possible to model the proposed system using models purely based on data in a simple way, using only the discretization parameter of the modules as input to the network. However, something worth pointing out is that, when the inlet conditions of the provided test dataset were farther from the points used for training, the greater the MSE of the test data. Unfortunately, no set of charts was generated for those poor results, however, it can be noted on the shell pressure discrepancy in Figure 5.2. According to the test data randomly generated, the shell pressure condition given to the data-driven model that generated the results from Figure 5.2 was far from the conditions used in training. This indicates that the discretization of the range of network operations would need to be large enough to reduce these gaps from inlet conditions away from the training data.

An alternate possibility to minimize the effects of this data gap is to include the inlet conditions as inputs to the neuronal model. This way, the weights and biases can be updated to make the model better understand how each variable behaves throughout the process, considering these inlet parameters. And so, it was understood that a study with these conditions would be appropriate to verify the best fits of the model and reduce the amount of data needed for training and validation.

Another benefit of this approach would be the reduction of overfitting since

the network would update its parameters based on the weighting of the inlet conditions and not dependent only on the applied normalization and the number of spatial discretizations of the gas separation module.

Although seven of the eight variables yielded satisfactory results, the utilization of spatial discretization alone in the development of a data-driven network was an additional attempt to evaluate the flexibility of data-driven models in the absence of essential information. Moreover, it aimed to simplify the model by reducing the number of weights and biases. However, this approach fails to consider the diversity of boundary conditions, as the models are incapable of learning different boundary conditions, which are essentially distinct boundary value problems. Consequently, the models can only predict specific operating conditions with minimal variation.

Furthermore, the solution generated by the network represents an average of all solutions contained in each boundary condition in the dataset. While normalization of training data may mitigate this effect, the network's assertiveness becomes unreliable when input conditions differ from the training dataset. To enable the network to learn the process behavior for various initial conditions, it is crucial to incorporate changes in boundary conditions as input to the network. Nonetheless, the resulting solution, while promising, cannot be applied to the construction of a Digital Twin as it is not generalizable.

- **4 input parameters: Module position + inlet conditions**

In this study, only the input variables of the data-driven model were changed. At this time, not only the position in the module but also the inlet molar flowrates of CO_2 and CH_4 , and also the value of the inlet shell pressure were considered as input variables. The same network structure, training, and optimization conditions of the previous study that yielded better results were maintained.

The results of this study are shown in Figure 5.4. It can be seen in the figure a fit of the final molar flowrates and the prediction of permeability are very similar to those found by the previous model with one input variable. However, when comparing the results for both tube and shell pressure settings, the new model with four input variables performed much better than the best model proposed in the last section.

The sum of the loss function fits of all normalized variables was 2.3×10^{-4} with an average relative error of less than 0.1% and an average R^2 of 0.9843. Significantly better result than that obtained with the previous model. Despite these promising results, the model took about 60% longer to predict the values of each module finite volume. While in the previously proposed model, the evaluation time was only one second, the model with four input variables took, on average,

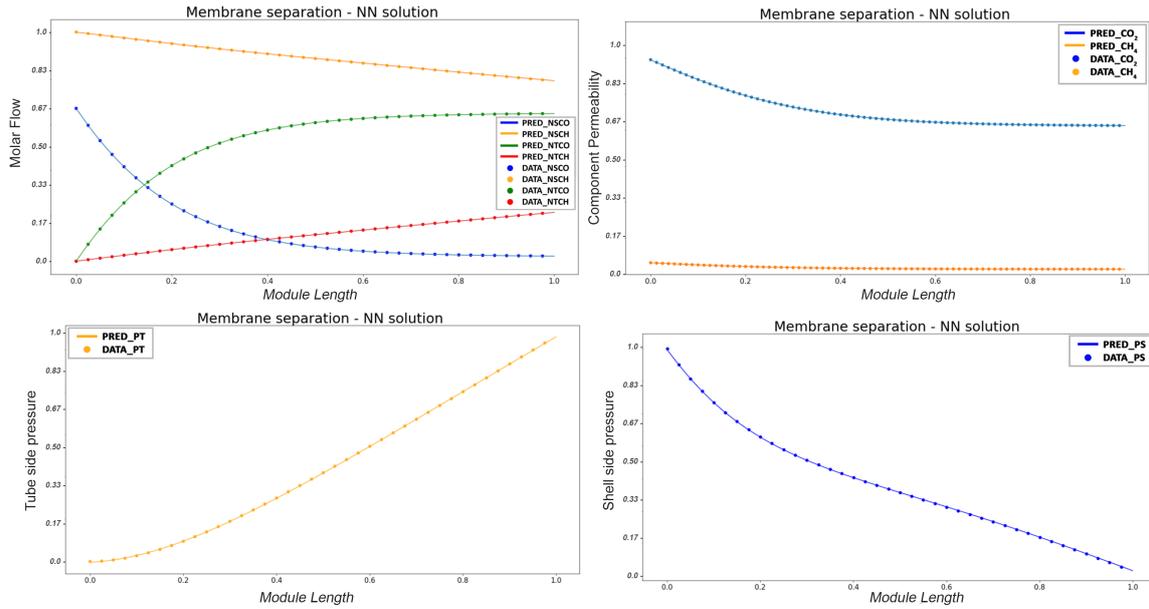


Figure 5.4: Best results for molar flowrates, permeabilities, and side pressure evolution for the data-driven model using four inputs and one output for each network.

1.6 s to predict the values and profiles of the eight variables within the membrane module.

Another essential point evaluated was the network performance for inlet conditions far from the discretization points provided in the training dataset. Unlike the network with only one variable, the model with four input variables showed no apparent difference in the adjustment to the test data of the phenomenological model, adding another positive point for the preference for using this new structure found.

For all these reasons, the modeling of the PINNs presented in the following sections was all carried out considering four input variables.

5.3.2 One module PINN model

With the promising results of the last section, the first version of PINN was modeled following the same conditions as the neural network found in the previous section (four inputs). The loss equation weights were all kept equal to one in this first study. However, the first results were not presented in such a way as to be considered valid. See an example of the result obtained in Figure 5.5, which shows the comparison between the values predicted by the PINN and those whose development comes from the phenomenological equation for a given inlet condition.

From this result, some actions were taken, and changes in the structural parameters of the network were evaluated. Although each training session took almost

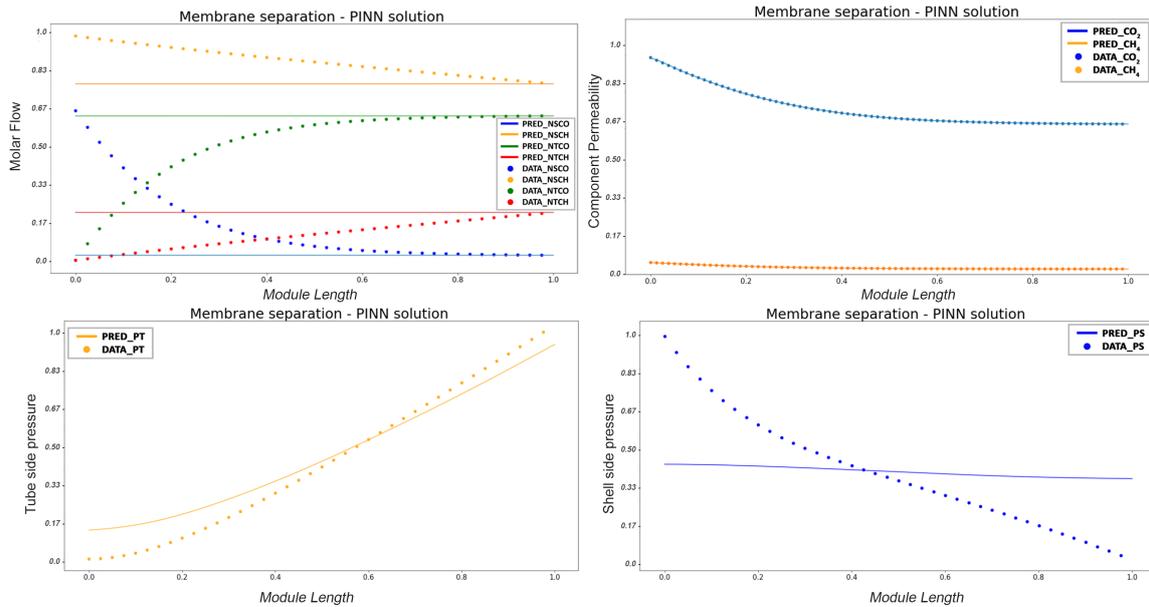


Figure 5.5: First results for molar flowrates, permeabilities, and side pressure evolution for the PINN model using four inputs and one output for each network.

4 h to be done, the following changes were proposed and evaluated:

- Changing each neural network hidden layers number;
- Varying the number of training epochs;
- Changed the learning rate;
- Changed the weights of each feature of the loss function.

At the end of the search roadmap, some conclusions could be drawn. The first is that a minimum of 4 hidden layers were needed to produce good results for all output variables. Adjustment deviations in the MSE of the validation dataset are found from 1000 epochs of training, characterizing possible overfitting of the PINN, which is why the 500 epochs initially proposed were maintained. The initial learning rate of the PINN could not be greater than 0.01, with a decay rate of ten times less every 100 training epochs. Otherwise, it would show signs of losing track of the lowest MSE point.

In addition to this acquired knowledge, the most significant improvements occurred when changing the weights of the variables that make up the PINN loss function. By decreasing the weights of the residual equations parcel, a better fit was repaired in all data. They presented results very similar to those shown in the previous section for the data-driven model.

Based on this knowledge, a training and retraining path was expected to gradually update the weights of the loss equation at each new PINN training session.

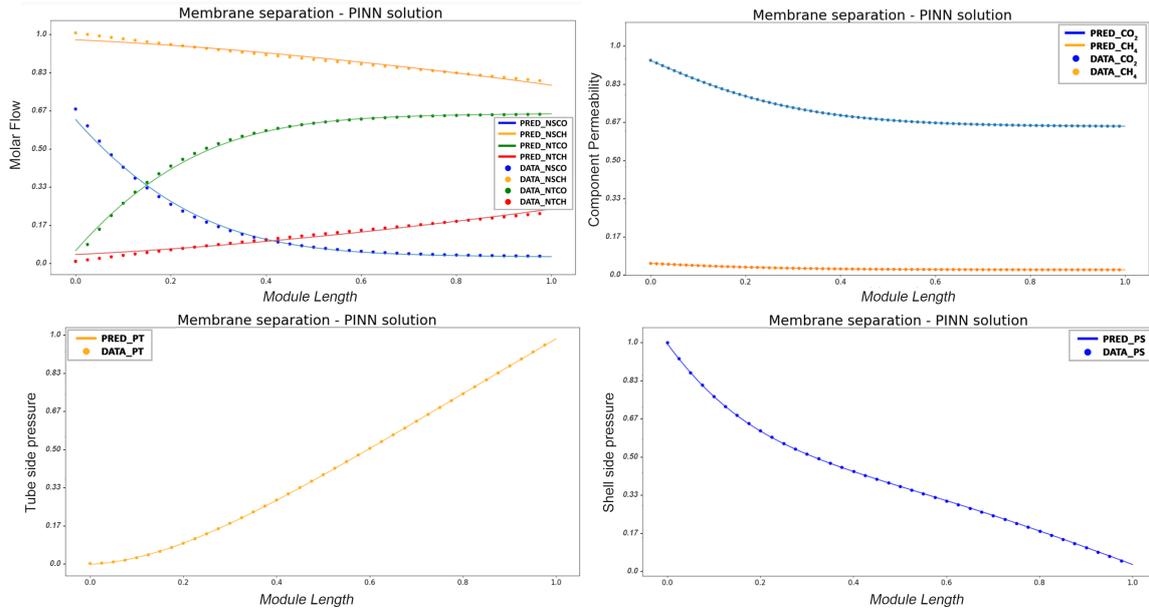


Figure 5.6: Best results for molar flowrates, permeabilities, and cavities pressure evolution for the PINN model using four inputs and one output for each network.

Initially, weights equal to 1 were considered for adjusting the data and boundary conditions, and 0.1 for all residual equations. Once trained, the states of the network were saved, and from this state, the weight of the portion of the residual equations was increased by 0.2, and the initial learning rate was reduced by 50%. This was done successively until it reached equal weight distribution for all variables.

The best results of this study are shown in Figure 5.6. Adjustments to the test data provided showed good results, fitting the molar flowrate, permeability, and pressures both in the tube and the shell. Despite the loss function having given almost 25 times more than the results of the data-driven model, the adjustments of each plot individually were in the order of 10^{-3} for the boundary conditions and residuals and the order of 10^{-4} for the PINN fitting of the data.

Table 5.1 compares the results obtained by the best four-input data-driven and the PINN models. In this table, it is possible to compare the performance of both proposed models numerically. Both models were shown to be reasonably accurate. However, the time both models took to evaluate the test given is still higher than the theoretical model used as a benchmark. Besides, when considering the MSE sum and evaluation time, the proposed models from this chapter proved to be more accurate and faster than the hybridization method proposed in Chapter 4. Evaluation time dropped 90% for the data-driven model and 30% for the PINN model.

Table 5.1: Curve adjustment comparison between the data-driven and PINN models' KPIs for normalized features.

| Model | IC | OC | Res | Data | $\sum MSE$ | Eval. Time |
|-------------|----------------------|----------------------|----------------------|----------------------|----------------------|------------|
| Theoretical | - | - | - | - | - | 0.58s |
| Data-driven | 9×10^{-5} | 1.2×10^{-4} | 7×10^{-6} | 3.6×10^{-6} | 2.3×10^{-4} | 1.6s |
| PINN | 2.9×10^{-4} | 1.4×10^{-3} | 1.9×10^{-3} | 5.2×10^{-3} | 5.6×10^{-3} | 6.9s |

5.3.3 Digital Twin data-driven model

In this part of the study, the process known and available in Appendix C for consultation was considered to model the digital twin and thus conclude this dissertation's third and final objectives.

By evaluating the flowchart of the actual process, it is possible to identify that the entire permeate flow is removed with each set of membranes. Therefore, the inlet conditions change as the retained flowrate changes sets, in such a way that it was necessary to create another dataset for training and validation that considered this change of inlet conditions every time the flowrates changed from one TRAIN to the other. This new dataset was synthetically generated, as presented in the methodology section of this chapter.

The modeling proposal for the digital twin using the data-driven model is a little different from the one studied in Section 4. As there are new inlet conditions in each set, a model was proposed that considers the 3 or 4 modules coupled as if they were a single gas separation module. Instead of considering the length of the module according to each one individually, it was considered as if the end of each module was directly connected to the next one.

In this way, the behavior of the variables in the first set of membranes would be equivalent to that of a module with the same configurations as the standardized module used so far in the course of this work but with a length three times greater due to the number of 3 modules in set 1. The same was given to sets 2 and 3 of each TREM, whose number of membrane modules is 4, with a length four times greater than the standardized module used for the study until then.

First, a set of 4 coupled membranes was trained (1 data-driven model to predict the evolution of the variables by the size 4x more significant than the one studied in Section 4). The training process took about 6 h using the 4.3 MM dataset. Despite the time and the expectation of providing good results, the result obtained for the pressure loss in the shell drew attention. The data-driven model pursued the other variables with low errors and MSE (on the order of 10^{-2}). However, it is clear that as the Model evaluation progresses along the longer length of the coupled modules, there is a significant offset close to the output conditions.

As this behavior occurred only in one of the output variables, it was consid-

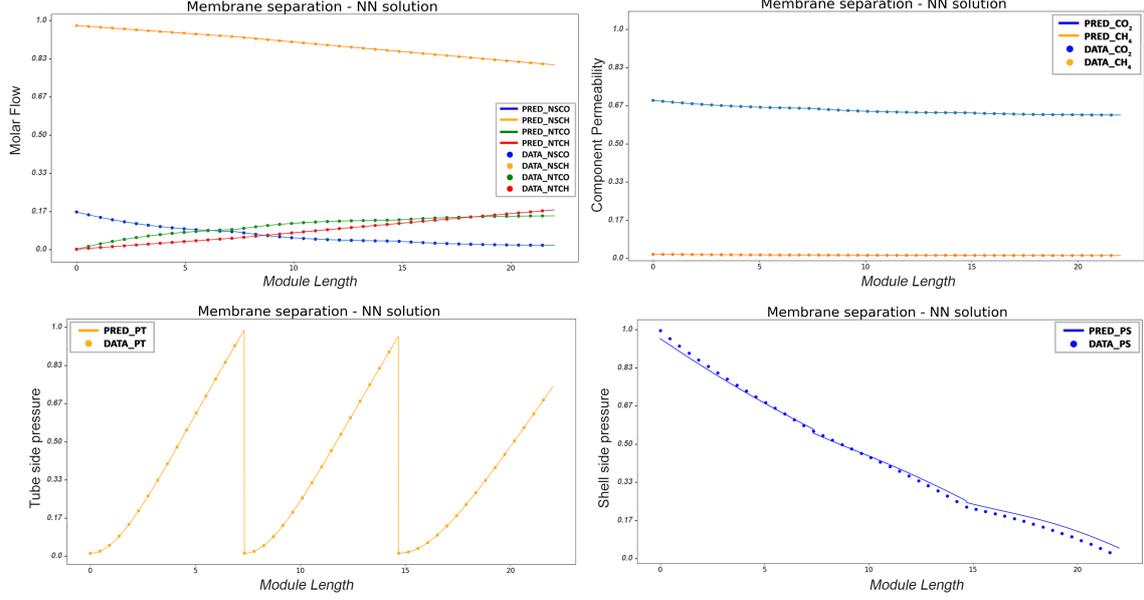


Figure 5.7: Best results for molar flowrates, permeabilities, and side pressure evolution for the Digital Twin data-driven-modeled.

ered to study another method of normalization of the shell pressure to overcome this problem. Among the proposals, the most effective one was the min-max normalization of the pressure difference between the value of each point (*value*) and the dataset maximum value ($MAX(value)$), as represented by Equation 5.3.

$$\Delta V = MAX(value) - value \quad (5.3)$$

where ΔV is the difference between the maximum value from the dataset and each point value. The normalization method is made over those ΔV values given by the min-max method presented in Equation 5.4. Where $MAX(\Delta V)$ and $MIN(\Delta V)$ are the ΔV maximum and minimum value from the dataset, respectively.

$$Normalized\ value = 1 - \frac{MAX(\Delta V) - \Delta V}{MAX(\Delta V) - MIN(\Delta V)} \quad (5.4)$$

This change resulted in significantly fewer errors than those of the first study for the evolution of the shell pressure, keeping the other variables in the same prediction performance. Figure 5.7 shows the final result for a given inlet condition and its comparison with its theoretical development.

At the end of the study, the sum of the MSE of the purely data-based model was of the order of 10^{-2} considering normalized data. An expressive result with minor errors. Despite the strictly data-based model having no contribution from the phenomenological equations that govern the process's nature, the model's ac-

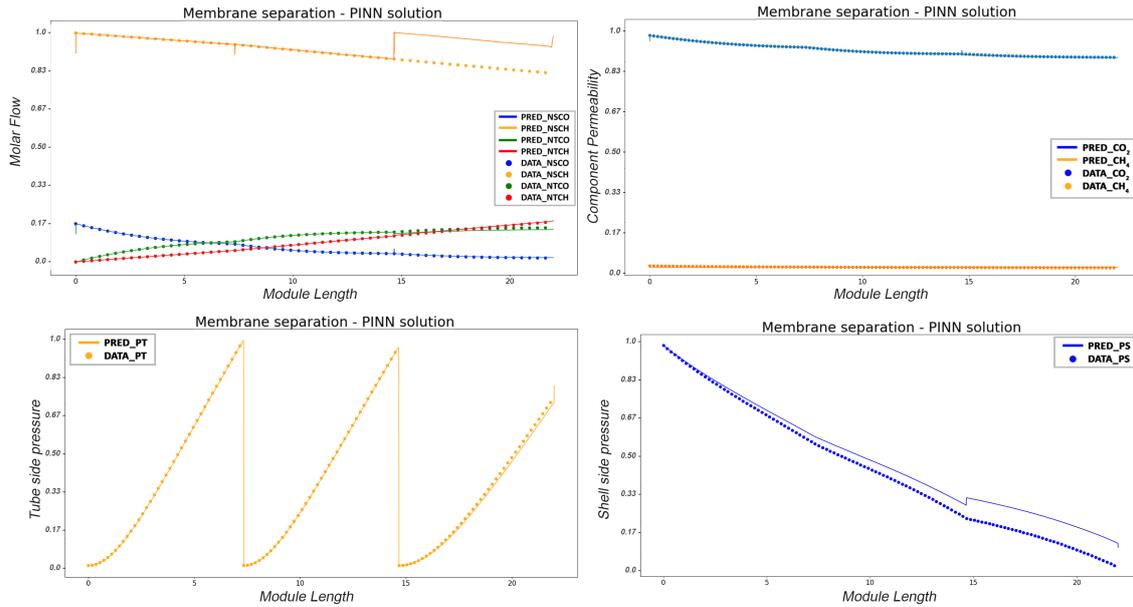


Figure 5.8: First results of molar flowrates, permeabilities, and side pressure evolution for the PINN model using four inputs and one output for each network.

curacy was much improved.

Like most metrics of this model, the evaluation of the given inlet test conditions took an average of 3.2 s. The sum of the MSE of the adjusted variables of the model's predicted variables concerning the data was 1.2×10^{-3} .

5.3.4 Digital Twin PINN model

In this study, the same PINN structure developed in Section 5.3.2 was used with the same training strategy from the beginning. First, the proposed PINN was trained using only data from the boundary conditions and the dataset and with the loss equation weights set to 0 for the residual equations, in order to generate a starting point which was given by the data-driven model.

Then, the weights of the residual portion were gradually increased at the same time as the learning rate was reduced, as was done in Section 5.3.2. The results obtained with this strategy for the digital twin clearly showed similar behavior to those obtained for the first study carried out for the digital twin considering the data-driven model with the original pressure normalization conditions. Figure 5.8 shows the mentioned results, however, alongside the deviation some continuation boundary conditions errors were found alongside the shell pressure offset.

Although the continuation problem, the adjustments for the variables of molar flowrate and tube pressure permeability were well-adjusted. Only a coarse offset in the prediction of the shell pressure was noticed. The total training time of the

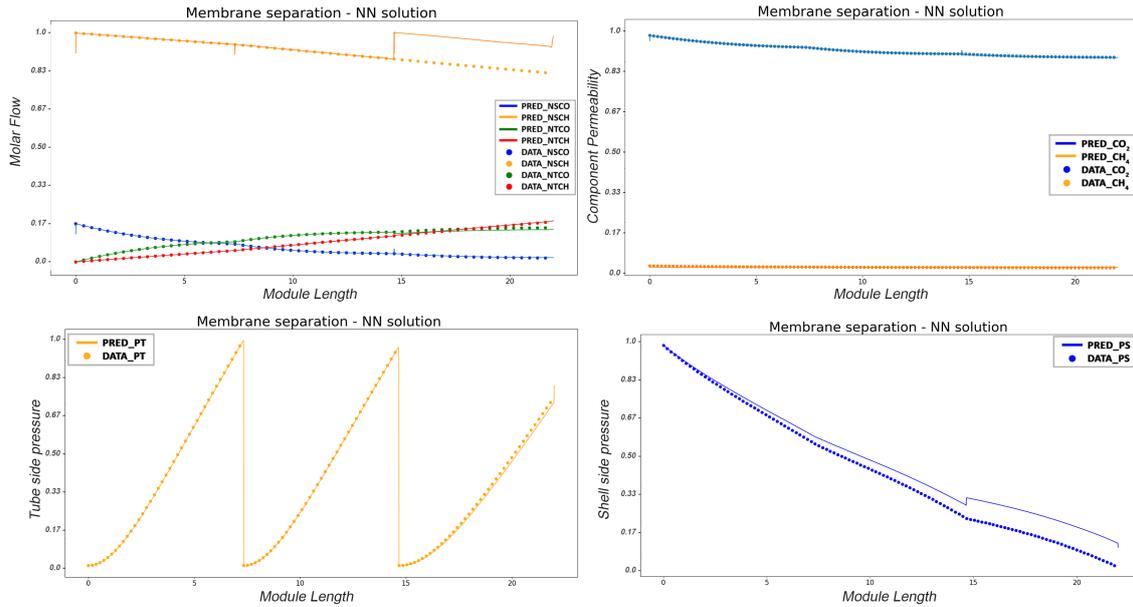


Figure 5.9: Best results for molar flowrates, permeabilities, and side pressure evolution for the digital twin PINN-modeled.

PINN, considering the dataset of the digital twin, was approximately 8 h for each parameter update, totaling a training run of roughly 32 h.

With this result, the need to consider a new normalization strategy to improve the adjustment of the pressure variable was evident. Thus, the same normalization strategy used for the data-driven model of the digital twin was used. The new results are shown in Figure 5.9. In this figure, one can see an improvement in the pressure adjustment. However, the offset did not decrease as much as in the data-driven model. Despite this, in absolute values, the deviation from the boundary condition was close to 1% as the pressure unit of measurement was Pascal. The normalized chart presented does not make that clear. The sum of the MSE of the adjustment was 9.2×10^{-2} considering the normalized values.

The PINN model took approximately 8.2 s to resolve the entire digital twin for the inlet test conditions. This result is slightly lower than expected but satisfactory regarding neural network prediction.

Table 5.2 shows a summary of the results obtained by data-driven and the PINN models for the digital twin. In this table, it is possible to compare the performance of both proposed models numerically. Both models were shown to be reasonably accurate. The boundary conditions and data higher values of MSE for the PINN model were majorly due to the shell pressure carried-over errors. The MSE was of the 10^{-5} order of magnitude for the other variables. Besides, the time both models took to evaluate the test given is still higher than the theoretical model used as a benchmark. However, when considering the MSE sum and evaluation time, the proposed models from this chapter proved to be accurate

Table 5.2: Digital-Twin curve adjustment comparison between the data-driven and PINN models' KPIs for normalized features.

| Model | IC | OC | Res | Data | $\sum MSE$ | Eval. Time |
|--------------------|----------------------|----------------------|----------------------|----------------------|----------------------|------------|
| Theoretical | - | - | - | - | - | 0.82s |
| Data-driven | 2×10^{-5} | 5×10^{-3} | 3×10^{-6} | 1.2×10^{-3} | 7.2×10^{-3} | 3.2s |
| PINN | 1.2×10^{-3} | 5.2×10^{-2} | 5.0×10^{-3} | 3.3×10^{-2} | 9.2×10^{-2} | 8.2s |

Table 5.3: Digital-Twin curve adjustment comparison between the data-driven and adjusted PINN (aPINN) models' KPIs for normalized features.

| Model | IC | OC | Res | Data | $\sum MSE$ | Eval. Time |
|--------------------|----------------------|----------------------|----------------------|----------------------|----------------------|------------|
| Theoretical | - | - | - | - | - | 0.82s |
| Data-driven | 2×10^{-5} | 5×10^{-3} | 3×10^{-6} | 1.2×10^{-3} | 7.2×10^{-3} | 3.2s |
| PINN | 1.2×10^{-3} | 5.2×10^{-2} | 5.0×10^{-3} | 3.3×10^{-2} | 9.2×10^{-2} | 8.2s |
| aPINN | 1.2×10^{-3} | 3.1×10^{-4} | 1.0×10^{-6} | 1.4×10^{-4} | 1.7×10^{-3} | 8.2s |

and faster enough to be implemented for a 10-second sampling time monitoring system.

As one can see in Figure 5.9, there is still a considerable offset as the prediction advances through the digital twin length. In order to improve the prediction accuracy of the shell pressure, an adjustment function is proposed to correlate the offset errors between the predicted and actual values. To achieve this, a 3rd order polynomial approach was considered. For that study, 12,000 offset data points from 10 different initial conditions were used to estimate the polynomial fit. The results obtained were good, with the mean square error (MSE) for boundary conditions reducing from 10^{-2} to 10^{-4} .

The resulting polynomial equation is shown in Equation 5.5 which correlation coefficient was 0.9990.

$$\xi = 7.47 \times 10^{-3} z^3 - 0.631 z^2 + 30.09 z - 4.19 \quad (5.5)$$

As a matter of comparison, the same case study in Figure 5.9 is shown in Figure 5.10 with the polynomial error adjustment fitted. This approach outperformed the data-driven model, indicating that the proposed polynomial fit was effective in reducing the offset errors and improving the accuracy of the PINN model. Overall, the results of this study highlight the potential of polynomial fitting in enhancing the accuracy of physics-informed neural networks, which is an important step toward improving the reliability of digital twin models.

Table 5.3 summarizes the findings and compares the adjustment curves considering the error-adjusted PINN.

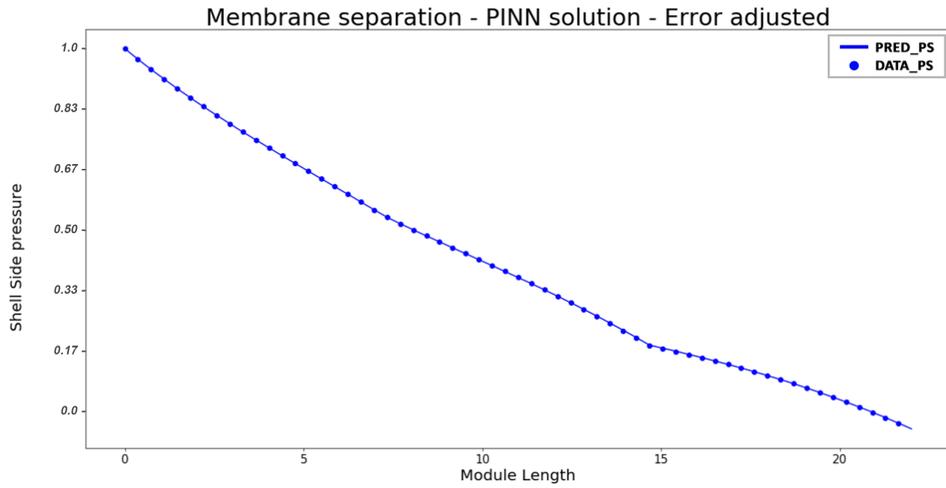


Figure 5.10: Best results for shell pressure after the application of the error adjustment function.

5.4 Conclusions

This chapter presents strategies for modeling the gas separation process using pure neural networks. However, both proposed models move away from the idea of a black-box model since the data-driven model was trained to understand how the variables behave from data generated by the phenomenological model. In the same way, the PINN, in its conception, aims to pursue not only the data but the restrictions of the process and the trends of variations described by the residual equations that govern the physical process.

The results presented in this chapter show the ability of neural networks to solve coupled differential equations and deal with non-linearities in the same way that phenomenological models are capable. They presented much more accurate and faster results than the hybrid models proposed in the previous chapter. For this reason, they were chosen for the digital twin modeling and thus fulfilled the 3rd objective proposed in this work.

However, despite the efficiency of the models proposed in this chapter, it is worth mentioning the importance of a detailed study of the structure of the proposed networks, as well as a detailed study of the best configuration of parameters and conditions for optimization and training of the neural network models. Each model is unique and has specific needs and conditions of use and setup. More time was taken to find the optimal settings than testing the model against test data.

It is noteworthy that the data-driven models and PINNs do not need all the parameters of the membrane modules to understand the data relation. Once trained with theoretical equations, it is possible to extrapolate without knowing all the

details of each membrane module, which makes these models highly flexible and scalable, as long as the operating conditions are among the states used in training the models.

Chapter 6

Conclusion

6.1 Developments Overview

In the introduction of this document, we highlighted the main motivations for developing a hybrid approach to develop a digital twin for a gas separation unit. Among them, one can highlight:

1. *Develop a phenomenological model to describe each membrane module's mass, momentum, and heat transfer.*
2. *Develop, train, and evaluate a Multi-Layer Perceptron network to predict compound permeability.*
3. *Determine the best hybridization structure that can predict the output with significant accuracy and efficiency to be implemented in real-time.*
4. *Create and evaluate the performance of a digital twin with the best alternative model developed for the third objective.*

Until recently, gas separation unit models were purely theoretical or phenomenological and were based on thermodynamic, physical, and chemical properties. These models respect physical laws and result in the process of detailed comprehension of how each variable interacts. However, it has some restraints when facing non-linearity problems, mainly predicting and analyzing mass transfer through different membrane fibers in a gas separation process (ASGHARI *et al.*, 2018).

Machine learning-based models, though, could deal with this non-linearity more efficiently. However, being an empirical model, it also has uncertainty in describing a complex manufacturing process. Besides, it does not take into account any thermodynamic or physical law. Because of that, combining both models into a hybrid one can result in a powerful approach to respect the physical laws and

deal with non-linearity faster than purely phenomenological, which also enables more accurate real-time optimization algorithms (ASGHARI *et al.*, 2018; QUIZA *et al.*, 2012). Also, the use of machine learning is an attribute of the new digital era in industry, one that promises more profitability and reliability (KANE *et al.*, 2015).

Furthermore, as presented during the theory and bibliography review chapter, there were no studies on natural gas sweetening considering a hybrid ANN and Phenomenological approach. On the contrary, many phenomenological approaches were found, and only a few studies used a purely black box model, trained using experimental data. None was found using gas separation unit data up to the time the bibliography review was finished (February 2023).

In Chapter 3, the hollow-fiber models were presented according to the continuity equations from BIRD *et al.* (2004); WHITAKER (1977) and MULDER (1996). Also, a one-dimensional mesh and finite volume discretization were presented for mass, momentum, and energy conservation. The present study has successfully proposed a phenomenological model that balances robustness and computational efficiency. The model considers essential aspects of the studied phenomena and provides a reliable and accurate representation. Furthermore, its computationally friendly nature makes it easily implementable and computationally efficient, making it an ideal solution for practical applications. Also, the proposed model confirms that a gas separation unit using membranes on the oil & gas field using high pressures and high feed flows is pointless considering a diffusive parcel. The results of this study showed that it is possible to achieve a good balance between robustness and computational efficiency and provide valuable insights for future model development studies in a similar field.

The data treatment and manual hybridization method were investigated in Chapter 4. It was shown that normalizing all data proved essential for not only speeding up training but also in terms of accuracy. Furthermore, the gradient-based network parameter actualization method (ADAM) proved more accurate than the adaptive gradient (ADAGRAD). Also, the hyperbolic tangent activation function was slightly off-target compared to the sigmoid function. Its maximum correlation coefficient to the test datasets was at least 0.99 for each compound. However, the training time was cut in half using the hyperbolic tangent activation function. For that, the hyperbolic tangent was the most suitable parameter to address the problem better. Still, in Chapter 4, a study has compared the performance of a hybrid network-phenomenological model with a purely phenomenological model. The results showed that the hybrid model achieved an average relative error of less than 2%, showing its accuracy compared to the phenomenological model. However, the evaluation time of the hybrid model was at least 100

times higher than that of the phenomenological model, making it infeasible for real-time implementation in the proposed digital twin. This highlights the trade-off between accuracy and computational efficiency in selecting models for practical applications. Although the hybrid model has similar results, the hybridization method proposed in this chapter resulted in high computational costs. It may limit its use in specific applications where real-time performance is crucial. It is also worth noting that a trade-off from the hybridization method proposed in this chapter still requires a phenomenological model with all equipment parameters and configurations known. That level of detail may not be available to all FPSOs or units, just as it was not available for this work.

Chapter 5 presented the development route of two different models from those previously proposed. One 100% data-driven and another whose hybridization did not take place to incorporate a neural network into a phenomenological model but instead incorporated the results of the phenomenological model as a target for estimating the data-driven network, a concept known since 2019 in literature such as Physics-Informed Neural Networks (PINNs). During the development of these models, the high relevance of a good strategy for normalizing the variables, the amount of information available for estimation, and the need for slow and gradual training in the case of PINNs became evident.

In addition, both models proposed in this chapter met the efficiency, agility, and accuracy goals that had not been found together during the studies of the previous chapter. Instead of taking about 10 s to evaluate the profiles of each variable within 1 module, the data-driven, and PINN models were able to evaluate the entire digital twin in 3 s and 7 s, respectively. Hence, these strategies can be implemented in real-time if the frequency of sending sensors is 10s or more. Furthermore, even with the shell pressure prediction needing a post adjustment, by fitting a polynomial function after the PINN, it does not correspond to deviations more significant than 1% in absolute terms, so it is possible to conclude that both the data-driven model and the PINNs fulfilled the main objective of this study.

6.2 Work Contributions

At the time of completion of this work, no other work was proposed to study alternatives to existing models for gas separation processes on offshore platforms. Furthermore, no digital twin modeling was found in the literature using physical-informed neural networks or comparing a pure data-driven model with PINNs. The results and conclusions presented in this work are expected to be significantly valuable and a time saver for oil & gas industry managers in research and development. They are proposed to help streamline the digitalization and expansion

of the oil & gas segment for industry 4.0.

6.3 Proposal for Continuation

This work showed and proved models that work both in the phenomenological and data-driven parts. However, some parallel paths can be better explored to improve the scope of the proposed models.

In terms of application, the present work can serve as a basis for further developments, such as:

- Use the proposed phenomenological model with actual gas separation unit parameters and compare the variables against what is read on the sensors and validate the model also with real data;
- Use another hybridization method rather than just estimating the permeability factors;
- With enough data, consider expanding the evaluation for more than just 2 compounds;
- Extrapolate the proposed models and try to predict the parameters for a pre-salt gas composition;
- Create an application capable of taking in the models proposed in this work and testing it in real-time simulations to verify its theoretical capabilities found in this work;
- Use the proposed models as the basis of an AI-based control system.

In terms of methodology, some open matters to the development of the proposed algorithm can be cited:

- Evaluate and compare the proposed phenomenological model against other pressure estimation equations different from the PAN's study;
- Evaluate if a deep learning algorithm would have performed faster and more accurately for the hybridization method using the neural network to predict the permeability;
- Use the PINN normalization and structure as a benchmark and propose modifying its structural parameters and try to achieve better results for the shell pressure;
- If available, use real data to train and test data-driven and PINN models;

- Extrapolate the inputs used on the test dataset to verify how all models deal with offset variables.

In conclusion, the oil and gas industry is at the threshold of a digital revolution, and there is a pressing need to integrate Industry 4.0 technologies. The presented work has proposed several models that could accelerate the digital maturity of the oil and gas industry, providing a head start for the national industry. With the low digital maturity of the oil and gas segment, the proposed models offer a promising solution to enhance oil and gas operations' efficiency, safety, and sustainability. The oil and gas industry can unlock new operational excellence and competitiveness levels by leveraging cutting-edge technologies such as machine learning, artificial intelligence, and the Internet of Things. This dissertation highlighted the importance of digital transformation in the national oil and gas industry and demonstrated that with the right approach, it can achieve its digital maturity faster and become a head start in the digital era.

Bibliography

- 2H OFFSHORE, 2020. "Riser System Digital Twin". Availability: <<https://2hoffshore.com/services/integrity-life-extension/riser-system-digital-twin/>>.
- AARON, D., TSOURIS, C., 2005, "Separation of CO₂ from flue gas: A review", *Separation Science and Technology*, v. 40, n. 40:1-3, pp. 321–348. doi: <10.1081/SS-200042244>.
- AHARI, J. S., SADEGHI, M. T., PASHNE, Z. S., 2011, "Optimization of OCM reaction conditions over Na – W – Mn / SiO₂ catalyst at elevated pressure", *Journal of the Taiwan Institute of Chemical Engineers*, v. 42, n. 5, pp. 751–759. doi: <10.1016/j.jtice.2011.02.005>.
- AHMAD, F., LAU, K. K., SHARIFF, A. M., et al., 2012, "Process simulation and optimal design of membrane separation system for CO₂ capture from natural gas", *Computers and Chemical Engineering*, v. 36, n. 1, pp. 119–128. doi: <10.1016/j.compchemeng.2011.08.002>.
- AL-BAGHLI, N. A., PRUESS, S. A., YESAVAGE, V. F., et al., 2001, "A rate-based model for the design of gas absorbers for the removal of CO₂ and H₂S using aqueous solutions of MEA and DEA", *Fluid Phase Equilibria*, v. 185, n. 1-2, pp. 31–43. doi: <10.1016/S0378-3812(01)00454-X>.
- AL-MUTAZ, I. S., SOLIMAN, M. A., AL-ZAHRANI, A. E., 1997, "Modeling and simulation of hollow fine fiber modules with radial dispersion - A parametric sensitivity study", *Desalination*, v. 110, n. 3, pp. 239–250. doi: <10.1016/S0011-9164(97)00102-1>.
- ALEKSANDER, I., MORTON, H., 1990, *An introduction to neural computing*. 1 ed. The University of Michigan, Michigan/USA, Chapman and Hall. ISBN: 9780412377808. Availability: <<https://books.google.com.br/books?id=b4ZQAAAAMAAJ>>.
- ANP, 2008, *AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, RESOLUÇÃO ANP Nº 16, DE 17.6.2008.*

Relatório técnico, AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS, Brasília. Availability: <<http://legislacao.anp.gov.br/?path=legislacao-anp/resol-anp/2008/junho&item=ranp-16-2008>>.

ANP, 2019a, *Gás Natural e Biocombustíveis: Anuário Estatístico*. Relatório técnico, Agência Nacional do Petróleo, Brasília, a. Availability: <<http://www.anp.gov.br/publicacoes/anuario-estatistico/5237-anuario-estatistico-2019#Se{ç}{~}{a}}o1>>.

ANP, 2019b, *Painel Dinâmico de Produção de Petróleo e Gás Natural*. Relatório técnico, Agência Nacional do Petróleo, Brasília, b. Availability: <<http://www.anp.gov.br/exploracao-e-producao-de-oleo-e-gas/painel-dinamico-de-producao-de-petroleo-e-gas-natural>>.

ANSARINASAB, H., MEHRPOOYA, M., PARIVAZH, M. M., 2017, "Evaluation of the cryogenic helium recovery process from natural gas based on flash separation by advanced exergy cost method – Linde modified process", *Cryogenics*, v. 87, pp. 1–11. doi: <10.1016/j.cryogenics.2017.08.005>. Availability: <<https://doi.org/10.1016/j.cryogenics.2017.08.005>>.

ARINELLI, L. O., DE MEDEIROS, J. L., ARAJO, O. Q., 2015, "Performance analysis and comparison of membrane permeation versus supersonic separators for CO₂ removal from a plausible natural gas of libra field, Brazil", *Offshore Technology Conference Brasil*, pp. 996–1007. doi: <10.4043/26164-ms>.

ASGHARI, M., DASHTI, A., REZAKAZEMI, M., et al., 2018, "Application of neural networks in membrane separation", *Reviews in Chemical Engineering*. doi: <10.1515/revce-2018-0011>.

ATCHARIYAWUT, S., JIRARATANANON, R., WANG, R., 2007, "Separation of CO₂ from CH₄ by using gas-liquid membrane contacting process", *Journal of Membrane Science*, v. 304, n. 1-2, pp. 163–172. doi: <10.1016/j.memsci.2007.07.030>.

BAKER, R. W., 2004, *Membrane Technology and Applications*. Second ed. Chichester, England, John Wiley Sons, Ltd. ISBN: 0471440264.

BALABIN, R. M., LOMAKINA, E. I., SAFIEVA, R. Z., 2015, "Neural network (ANN) approach to biodiesel analysis : Analysis of biodiesel density , kinematic viscosity , methanol and water contents using near infrared

- (NIR) spectroscopy”, *Fuel*, v. 90, n. 5, pp. 2007–2015. doi: <10.1016/j.fuel.2010.11.038>.
- BELTRAO, R. L., SOMBRA, C., LAGE, A., et al., 2009, “Challenges and New Technologies for the Development of the Pre-Salt Cluster, Santos Basin, Brazil”, *Offshore Technology Conference Brasil*. doi: <10.4043/otc-19880-ms>.
- BHIDE, B. D., VOSKERICYAN, A., STERN, S. A., 1998, “Hybrid processes for the removal of acid gases from natural gas”, *Journal of Membrane Science*, v. 140, n. 1, pp. 27–49. doi: <10.1016/S0376-7388(97)00257-3>.
- BIRD, R. B., STWEWART, W. E., LIGHTFOOT, E. N., 2004, *Transport Phenomena*. John Wiley Sons, Inc. ISBN: 0471410772.
- BOTTOMS, R. R., 1930, “Process for separating acidic gases, United States Patent office”, .
- BP, 2019, *BP Statistical Review of World Energy*. Relatório técnico, BP. Availability: <<https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2019-full-report.pdf>>.
- BP, 2022, *BP Statistical Review of World Energy*. Relatório técnico, BP. Availability: <<https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2022-full-report.pdf>>.
- CAPLOW, M., 1968, “Kinetics of Carbamate Formation and Breakdown”, *Journal of the American Chemical Society*, v. 90, n. 24, pp. 6795–6803. doi: <10.1021/ja01026a041>.
- CARVALHO, F. S., BRAGA, J. P., 2022, “Physics Informed Neural Networks applied to liquid state theory”, *Journal of Molecular Liquids*, v. 367, pp. 120504. doi: <10.1016/j.molliq.2022.120504>.
- CASKEY, T. L., 1993. “METHOD OF FORMING TUBESHEET FOR HOLLOW FIBERS”. .
- CHIOU, J. S., PAUL, D. R., 1987, “Effects of CO₂ exposure on gas transport properties of glassy polymers”, *Journal of Membrane Science*, v. 32, n. 2-3, pp. 195–205. doi: <10.1016/S0376-7388(00)85006-1>.

- CHIOU, J. S., BARLOW, J. W., PAUL, D. R., 1985, "Polymer Crystallization Induced by Sorption of CO₂ Gas", *Journal of Applied Polymer Science*, v. 30, n. 9, pp. 3911–3924. doi: <doi:10.1002/app.1985.070300929>.
- CHOWDHURY, M. H. M., FENG, X., DOUGLAS, P., et al., 2005, "A new numerical approach for a detailed multicomponent gas separation membrane model and AspenPlus simulation", *Chemical Engineering and Technology*, v. 28, n. 7, pp. 773–782. doi: <10.1002/ceat.200500077>.
- CHU, Y., LINDBRÅTHEN, A., LEI, L., et al., 2019, "Mathematical modeling and process parametric study of CO₂ removal from natural gas by hollow fiber membranes", *Chemical Engineering Research and Design*, v. 148, pp. 45–55. doi: <10.1016/j.cherd.2019.05.054>.
- CNOP, T., DORTMUNDT, D., SCHOTT, M., 2007, *Continued Development of Gas Separation Membranes for Highly Sour Service*. Relatório técnico, UOP LLC, Illinois - USA. Availability: <<https://www.uop.com/?document=uop-continued-development-of-gas-separation-membranes-technical-paper&download=1>>.
- COKER, D. T., FREEMAN, B. D., FLEMING, G. K., 1998, "Modeling multicomponent gas separation using hollow-fiber membrane contactors", *AIChE Journal*, v. 44, n. 6, pp. 1289–1302. doi: <10.1002/aic.690440607>.
- COKER, D. T., ALLEN, T., FREEMAN, B. D., et al., 1999, "Nonisothermal model for gas separation hollow-fiber membranes", *AIChE Journal*, v. 45, n. 7, pp. 1451–1468. doi: <10.1002/aic.690450709>.
- COLOMER, M., ALMEIDA, E., 2015, "Desafios da Produção de Gás Não Convencional no Brasil", *5th Latin American Energy Economics Meeting*, p. 10. Availability: <http://www.ie.ufrj.br/images/pesquisa/pesquisa/textos_sem_peq/texto0704.pdf>.
- DALANE, K., DAI, Z., MOGSETH, G., et al., 2017, "Potential applications of membrane separation for subsea natural gas processing: A review", *Journal of Natural Gas Science and Engineering*, v. 39, pp. 101–117. doi: <10.1016/j.jngse.2017.01.023>.
- DANCKWERTS, P. V., 1979, "The reaction of CO₂ with ethanalamines", *Chemical Engineering Science*, v. 34, n. 4, pp. 443–446. doi: <10.1016/0009-2509(79)85087-3>.

- DE SOUZA JÚNIOR, M. B., 1993, *Redes neuronais multicamadas aplicadas a modelagem e controle de processos químicos*. Doctorate thesis, Universidade Federal do Rio de Janeiro, COPPE.
- DEBNATH, B., MUKHERJEE, A., MULLICK, A., et al., 2019, “Desublimation based separation of CO₂ inside a cryogenic packed bed: Performance mapping with the spatiotemporal evolution of CO₂ frost”, *Chemical Engineering Research and Design*, v. 146, n. 1962, pp. 166–181. doi: <10.1016/j.cherd.2019.03.024>.
- DEON, B., COTTA, K. P., SILVA, R. F., et al., 2022, “Digital twin and machine learning for decision support in thermal power plant with combustion engines”, *Knowledge-Based Systems*, v. 253, pp. 109578. doi: <10.1016/j.knosys.2022.109578>.
- DING, Y., LIU, C., ZHU, H., et al., 2023, “A supervised data augmentation strategy based on random combinations of key features”, *Information Sciences*, v. 632, n. January, pp. 678–697. doi: <10.1016/j.ins.2023.03.038>.
- DONOHUE, M. D., MINHAS, B. S., LEE, S. Y., 1989, “Permeation behavior of carbon dioxide-methane mixtures in cellulose acetate membranes”, *Journal of Membrane Science*, v. 42, n. 3, pp. 197–214. doi: <10.1016/S0376-7388(00)82376-5>.
- DUCHI, J. C., BARTLETT, P. L., WAINWRIGHT, M. J., 2011, “Adaptative Sub-gradient Methods for Online Learning and Stochastic Optimization”, *Journal of Machine Learning Research*, v. 12, pp. 2121–2159. doi: <10.1109/CDC.2012.6426698>.
- EBRAHIMI, A., ZIABASHARHAGH, M., 2017, “Optimal design and integration of a cryogenic Air Separation Unit (ASU) with Liquefied Natural Gas (LNG) as heat sink, thermodynamic and economic analyses”, *Energy*, v. 126, pp. 868–885. doi: <10.1016/j.energy.2017.02.145>.
- ECHT, W. I., SINGH, M., 2008, *Integration of membranes into natural gas process schemes*. Relatório técnico, UOP LLC, A Honeywell Company, Illinois - USA.
- EPE, 2018, *Plano Nacional de Energia 2050: Cenários de Demanda para o PNE 2050*. Relatório técnico, Empresa de Pesquisa Energética - Ministério de Minas e Energia, Brasília. Availability: <<http://www.epe.gov.br/sites-pt/>>

publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/
publicacao-227/topico-202/Cen{â}riosdeDemanda.pdf>.

- FERREIRA, R. S., 2006, *Desenvolvimento de materiais poliméricos uretânicos para a purificação de gás natural: remoção de mercúrio e compostos à base de enxofre*. Masters dissertation, Universidade Federal do Paraná, Curitiba.
- FRANKLIN, T. S., SOUZA, L. S., FONTES, R. M., et al., 2022, "A Physics-Informed Neural Networks (PINN) oriented approach to flow metering in oil wells: an ESP lifted oil well system as a case study", *Digital Chemical Engineering*, v. 5, n. August, pp. 100056. doi: <10.1016/j.dche.2022.100056>.
- FREEMAN, S. A., DAVIS, J., ROCHELLE, G. T., 2010, "Degradation of aqueous piperazine in carbon dioxide capture", *International Journal of Greenhouse Gas Control*, v. 4, n. 5, pp. 756–761. doi: <10.1016/j.ijggc.2010.03.009>.
- FUTUREON, 2020. "FieldTwin - Digital Twion Solution". Availability: <<https://www.futureon.com/digital-field-twin/>>.
- GABELMAN, A., HWANG, S. T., 1999, "Hollow fiber membrane contactors", *Journal of Membrane Science*, v. 159, n. 1-2, pp. 61–106. doi: <10.1016/S0376-7388(99)00040-X>.
- GAO, T., SELINGER, J. L., ROCHELLE, G. T., 2019, "Demonstration of 99scrubbing", *International Journal of Greenhouse Gas Control*, v. 83, n. October 2018, pp. 236–244. doi: <10.1016/j.ijggc.2019.02.013>.
- GOKHALE, G., CLAESSENS, B., DEVELDER, C., 2022, "Physics informed neural networks for control oriented thermal modeling of buildings", *Applied Energy*, v. 314, n. March, pp. 118852. doi: <10.1016/j.apenergy.2022.118852>. Availability: <<https://doi.org/10.1016/j.apenergy.2022.118852>>.
- GRAHAM, T., 1861, "Liquid Diffusion Applied to Analysis", *Philosophical Transactions of the Royal Society of London*, pp. 183–224. doi: <10.1098/rstl.1861.0011>.
- GRAHAM, T., 1866, "On the Absorption and Dialytic Separation of Gases by Colloid Septa", *Philosophical Transactions of the Royal Society of London*, v. 156, n. 1866, pp. 399–439.

- GRAINGER, D. R., 2007, *A simple method for solving mixed-permeate and counter-current membrane models with multi-component feeds*. Ph.D. Thesis, Doctoral thesis appendix.
- GRIEVES, M., 2014, *Digital Twin : Manufacturing Excellence through Virtual Factory Replication*. Relatório técnico. Availability: <http://www.apriso.com/library/Whitepaper_Dr_Grieves_DigitalTwin_ManufacturingExcellence.php>.
- HAGHIGHAT, E., JUANES, R., 2021, "SciANN: A Keras/TensorFlow wrapper for scientific computations and physics-informed deep learning using artificial neural networks", *Computer Methods in Applied Mechanics and Engineering*, v. 373. doi: <10.1016/j.cma.2020.113552>.
- HAMMERSCHMIDT, E. G., 1934, "Formation of Gas Hydrates in Natural Gas Transmission Lines", *Industrial and Engineering Chemistry*, v. 26, n. 8, pp. 851–855. doi: <10.1021/ie50296a010>.
- HARAYA, K., OBATA, K., HAKUTA, T., et al., 1986, "The permeation of gases through a new type polyimide membrane." *Membrane*, v. 11, n. 1, pp. 48–52. doi: <10.5360/membrane.11.48>.
- HASEGAWA, H., NISHIDA, K., OGURO, S., et al., 2017, "Gas Separation Process for CO₂ Removal from Natural Gas with DDR-type Zeolite Membrane", *Energy Procedia*, v. 114, n. November 2016, pp. 32–36. doi: <10.1016/j.egypro.2017.03.1143>.
- HAYKIN, S., 2001, *Redes neurais: princípios e prática*. 2nd ed. Porto Alegre, Bookman. ISBN: 9788573077186.
- HE, X., HÄGG, M. B., 2011, "Hollow fiber carbon membranes: Investigations for CO₂ capture", *Journal of Membrane Science*, v. 378, n. 1-2, pp. 1–9. doi: <10.1016/j.memsci.2010.10.070>.
- HE, X., LINDBRÅTHEN, A., KIM, T. J., et al., 2017, "Pilot testing on fixed-site-carrier membranes for CO₂ capture from flue gas", *International Journal of Greenhouse Gas Control*, v. 64, n. August, pp. 323–332. doi: <10.1016/j.ijggc.2017.08.007>. Availability: <<http://dx.doi.org/10.1016/j.ijggc.2017.08.007>>.
- HEBB, D. O., 1949, *The Organization of Behavior; A Neuropsychological Theory*. New York, John Wiley Sons, Inc.

- HOLMES, A. S., RYAN, J. M., PRICE, B. C., et al., 1983, "Pilot Tests Prove Out Cryogenic Acid-Gas/Hydrocarbon Separation Processes." *Oil and Gas Journal*, v. 81, n. 26.
- HUERTAS, J. I., GOMEZ, M. D., GIRALDO, N., et al., 2015, "CO₂ absorbing capacity of MEA", *Journal of Chemistry*, v. 2015, n. 2. doi: <10.1155/2015/965015>.
- IPCC, 2014, *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. N. 5. 5 ed. Geneva, Switzerland, IPCC. ISBN: 978-92-9169-143-2. Availability: <https://archive.ipcc.ch/pdf/assessment-report/ar5/syr/SYR_AR5_FINAL_full_wcover.pdf>.
- IPCC, 2021, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press. Availability: <<https://www.ipcc.ch/report/ar6/wg1/>>.
- JUNG, H. J., HAN, S. H., LEE, Y. M., et al., 2010, "Modeling and simulation of hollow fiber CO₂ separation modules", *Korean Journal of Chemical Engineering*, v. 28, n. 7, pp. 1497–1504. doi: <10.1007/s11814-010-0530-y>.
- KAGERMANN, H., WAHLSTER, W., HELBIG, J., 2013, *Securing the future of German manufacturing industry: Recommendations for implementing the strategic initiative INDUSTRIE 4.0*. Relatório Técnico April, National Academy of Science and Engineering (ACATECH), Frankfurt, Germany.
- KANE, G. C., PALMER, D., PHILLIPS, A. N., et al., 2015, *Strategy, Not Technology, Drives Digital Transformation*. Relatório Técnico July, MIT Sloan Management Review and Deloitte University Press.
- KAVLAKOGLU, E., 2020. "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?" Availability: <<https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>>.
- KENDER, R., KAUFMANN, F., RÖSSLER, F., et al., 2021, "Development of a digital twin for a flexible air separation unit using a pressure-driven simulation approach", *Computers and Chemical Engineering*, v. 151, pp. 107349. doi: <10.1016/j.compchemeng.2021.107349>.

- KIDNAY, A. J., PARRISH, W. R., 2006, *Fundamentals of Natural Gas Processing*. First ed. Boca Raton Fl, CRC Press, Taylor and Francis Group LLC. ISBN: 0-8493-3406-3.
- KINGMA, D. P., BA, J. L., 2015, "Adam: A method for stochastic optimization", *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15.
- KNAPIK, E., KOSOWSKI, P., STOPA, J., 2018, "Cryogenic liquefaction and separation of CO₂ using nitrogen removal unit cold energy", *Chemical Engineering Research and Design*, v. 131, pp. 66–79. doi: <10.1016/j.cherd.2017.12.027>.
- KOHL, A. L., NIELSEN, R. B., 1997, *Gas Purification*. 5th ed. Houston, Texas, Gulf Publishing Company. ISBN: 0-88415-220-0.
- KOVÁCS, Z. L., 1996, *Redes neurais artificiais: fundamentos e aplicações: um texto básico*. 2nd ed. São Paulo. ISBN: 85-86396-02-8.
- KOVVALI, A. S., VEMURY, S., ADMASSU, W., 1994, "Modeling of Multicomponent Countercurrent Gas Permeators", *Industrial and Engineering Chemistry Research*, v. 33, n. 4, pp. 896–903. doi: <10.1021/ie00028a016>.
- LEONARD, J., KRAMER, M. A., 1990, "Improvement of the Backpropagation Algorithm for Training Neural Networks", *Computers Chemical Engineering*, v. 14, n. 3, pp. 337–341. doi: <10.1016/0098-1354(90)87070-6>.
- LI, H., 2008, *Thermodynamic properties of CO₂ mixtures and their applications in advanced power cycles with CO₂ capture processes*. 1st ed. Department of Chemical Engineering and Technology, Stockholm, Sweden, KTH, Royal Institute of Technology. ISBN: 978-91-7415-091-9.
- LIMA, L. C. A., GONÇALVES, P. R. C., 2016, *Desafios do Desenvolvimento Sustentável: Estudo Sobre as Mudanças na Matriz Energética Brasileira*. Graduation monograph, Universidade Federal Fluminense, Rio de Janeiro.
- LIN, F., CHEN, P. C., HUANG, T. C., et al., 2021, "3D in-situ simulation and particle tracing of gas filtration process for ultrafine particles removal using a hollow fiber membrane", *Journal of Membrane Science*, v. 632, n. January, pp. 119380. doi: <10.1016/j.memsci.2021.119380>.
- LOCK, S. S., LAU, K. K., AHMAD, F., et al., 2015, "Modeling, simulation and economic analysis of CO₂ capture from natural gas using cocurrent,

- countercurrent and radial crossflow hollow fiber membrane”, *International Journal of Greenhouse Gas Control*, v. 36, pp. 114–134. doi: <10.1016/j.ijggc.2015.02.014>.
- LU, H., GUO, L., AZIMI, M., et al., 2019, “Oil and Gas 4.0 era: A systematic review and outlook”, *Computers in Industry*, v. 111, n. July, pp. 68–90. doi: <10.1016/j.compind.2019.06.007>.
- MAKOGON, Y. F., 2010, “Natural gas hydrates - A promising source of energy”, *Journal of Natural Gas Science and Engineering*, v. 2, n. 1, pp. 49–59. doi: <10.1016/j.jngse.2009.12.004>.
- MALISKA, C. R., 1995, *Transferência de Calor e Mecânica dos Fluidos Computacional: Fundamentos e Coordenadas Generalizadas*. 1st ed. Rio de Janeiro, LTC - Livros Técnicos e Científicos. ISBN: 9788521633358.
- MAQSOOD, K., PAL, J., TURUNAWARASU, D., et al., 2014, “Performance enhancement and energy reduction using hybrid cryogenic distillation networks for purification of natural gas with high CO₂ content”, *Korean Journal of Chemical Engineering*, v. 31, n. 7, pp. 1120–1135. doi: <10.1007/s11814-014-0038-y>.
- MARRIOTT, J., SØRENSEN, E., 2003, “A general approach to modelling membrane modules”, *Chemical Engineering Science*, v. 58, n. 22, pp. 4975–4990. doi: <10.1016/j.ces.2003.07.005>.
- MARRIOTT, J., 2001, “Detailed modelling and optimal design of membrane separation systems”, *Ph.D. Thesis*, , n. February. Availability: <<http://discovery.ucl.ac.uk/1349890/1/368065.pdf>>.
- MARZOUK, S. A., AL-MARZOUQI, M. H., TERAMOTO, M., et al., 2012, “Simultaneous removal of CO₂ and H₂S from pressurized CO₂-H₂S-CH₄ gas mixture using hollow fiber membrane contactors”, *Separation and Purification Technology*, v. 86, pp. 88–97. doi: <10.1016/j.seppur.2011.10.024>.
- MCCLELLAND, J. L., RUMELHART, D. E., 1988, *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. 1st ed. Cambridge, Massachusetts, The MIT Press. ISBN: 0-262-63129-6.
- MCCULLOCH, W. S., PITTS, W., 1943, “A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY”, *Bulletin of Mathematical Biology*, v. 5, pp. 115–133. doi: <10.1007/BF02478259>.

- MEAD, C., 1989, *Analog VLSI and Neural Systems*. Addison-Wesley. ISBN: 9780201059922. Availability: <<https://books.google.com.br/books?id=-j8PAQAAMAAJ>>.
- MEHL, A., 2009, *Estudo do equilíbrio de fases de hidrocarbonetos e CO₂ supercrítico*. Doctorate thesis, Universidade Federal do Rio de Janeiro, UFRJ.
- MIKOŁAJCZYK, A., GROCHOWSKI, M., 2018, “Data augmentation for improving deep learning in image classification problem”, *2018 International Interdisciplinary PhD Workshop, (IIPHDW)*, pp. 117–122. doi: <10.1109/IIPHDW.2018.8388338>.
- MUDUNURU, M. K., O’MALLEY, D., SRINIVASAN, S., et al., 2020, “Physics-informed machine learning for real-time unconventional reservoir management”, *CEUR Workshop Proceedings*, v. 2587.
- MULDER, M., 1996, *Basic Principles of Membrane Technology*. 2nd ed. Dordrecht, The Netherlands, Kluwer Academic Publishers. ISBN: 978-0-7923-4248-9.
- NASIR, R., SULEMAN, H., MAQSOOD, K., 2022, “Multiparameter Neural Network Modeling of Facilitated Transport Mixed Matrix Membranes for Carbon Dioxide Removal”, *Membranes*, v. 12, n. 4. doi: <10.3390/membranes12040421>.
- NELLIKATH, R., CHATZIVASILEIADIS, S., 2022, “Physics-Informed Neural Networks for AC Optimal Power Flow”, *Electric Power Systems Research*, v. 212, n. 949899, pp. 108412. doi: <10.1016/j.epsr.2022.108412>. Availability: <<https://doi.org/10.1016/j.epsr.2022.108412>>.
- NIST, 2019. “National Institute of Standards and Technology ,Thermophysical properties of fluid systems, U.S. Department of Commerce”. Availability: <<https://webbook.nist.gov/chemistry/fluid/>>.
- OKUBO, Y., SHUDO, A., OTA, Y., et al., 2018. “SPIRAL - WOUND ACID GAS SEPARATION MEMBRANE ELEMENT , ACID GAS SEPARATION MEMBRANE MODULE , AND ACID GAS SEPARATION APPARATUS”. Availability: <<https://patents.google.com/patent/US10092880B2/en>>.
- PAN, C. Y., 1986, “Gas separation by high-flux, asymmetric hollow-fiber membrane”, *AIChE Journal*, v. 32, n. 12, pp. 2020–2027. doi: <10.1002/aic.690321212>.

- PANDEY, P., CHAUHAN, R. S., 2001, "Membranes for gas separation", *Progress in Polymer Science (Oxford)*, v. 26, n. 6, pp. 853–893. doi: <10.1016/S0079-6700(01)00009-0>.
- PARO, A. D. C., 2005, *Estudo da contribuição do gás natural no setor elétrico – uma análise de cenários de sua expansão nos setores de geração termelétrica, cogeração e residencial*. Masters dissertation, Universidade de São Paulo, São Paulo.
- PARROTT, A., WARSHAW, L., 2017, *Industry 4.0 and the digital twin*. Relatório técnico.
- PASQUALETTE, M. A., REMPTO, M. J., CARNEIRO, J. N., et al., 2017, "Parametric study of the influence of GOR and CO₂ content on the simulation of a pre-salt field configuration", *Offshore Technology Conference Brasil*, pp. 1599–1610. doi: <10.4043/28093-ms>.
- PEER, M., MAHDYARFAR, M., MOHAMMADI, T., 2008, "Evaluation of a mathematical model using experimental data and artificial neural network for prediction of gas separation", *Journal of Natural Gas Chemistry*, v. 17, n. 2, pp. 135–141. doi: <10.1016/S1003-9953(08)60040-7>.
- PENG, D. Y., ROBINSON, D. B., 1976, "A New Two-Constant Equation of State", *Industrial and Engineering Chemistry Fundamentals*, v. 15, n. 1, pp. 59–64. doi: <10.1021/i160057a011>.
- PITZER, K. S., LIPPMANN, D. Z., CURL, R. F., et al., 1955, "The Volumetric and Thermodynamic Properties of Fluids. II. Compressibility Factor, Vapor Pressure and Entropy of Vaporization", *Journal of the American Chemical Society*, v. 77, n. 13, pp. 3433–3440. doi: <10.1021/ja01618a002>.
- POBEREZHNY, L., HRYSANCHUK, A., GRYTSULIAK, H., 2019, "Influence of the gas hydrates on the corrosion rate of gas gathering pipelines", *Procedia Structural Integrity*, v. 16, pp. 141–147. doi: <10.1016/j.prostr.2019.07.033>.
- POBEREZHNY, L., MARUSCHAK, P., HRYSANCHUK, A., et al., 2017, "Impact of Gas Hydrates and Long-Term Operation on Fatigue Characteristics of Pipeline Steels", *Procedia Engineering*, v. 187, pp. 356–362. doi: <10.1016/j.proeng.2017.04.386>.
- QI, R., HENSON, M. A., 1998, "Optimization-based design of spiral-wound membrane systems for CO₂-CH₄ separations", *Separation and Purifica-*

tion Technology, v. 13, n. 3, pp. 209–225. doi: <10.1016/S1383-5866(98)00044-6>.

QUEIROZ, L. H., SANTOS, F. P., OLIVEIRA, J. P., et al., 2021, “Physics-Informed deep learning to predict flow fields in cyclone separators”, *Digital Chemical Engineering*, v. 1, n. September, pp. 100002. doi: <10.1016/j.dche.2021.100002>.

QUIZA, R., LÓPEZ-ARMAS, O., DAVIM, J. P., 2012, *Hybrid Modeling and Optimization of Manufacturing Combining Artificial Intelligence and Finite Element Method*. New York, Springer Science Business Media. ISBN: 9783642280849.

RAISSI, M., PERDIKARIS, P., KARNIADAKIS, G. E., 2019, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”, *Journal of Computational Physics*, v. 378, pp. 686–707. doi: <10.1016/j.jcp.2018.10.045>. Availability: <<https://doi.org/10.1016/j.jcp.2018.10.045>>.

ROCHELLE, G., CHEN, E., FREEMAN, S., et al., 2011, “Aqueous piperazine as the new standard for CO₂ capture technology”, *Chemical Engineering Journal*, v. 171, n. 3, pp. 725–733. doi: <10.1016/j.cej.2011.02.011>.

ROSENBLATT, F., 1958, “The perceptron: A probabilistic model for information storage and organization in the brain”, *Psychological Review*, v. 65, n. 6, pp. 386–408. doi: <10.1037/h0042519>.

RUMELHART, D. E., MCCLELLAND, J. L., 1986, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. 1 ed. Cambridge, MA, USA, MIT Press. ISBN: 9780262132183.

SABERI, M., HASHEMIFARD, S. A., DADKHAH, A. A., 2016, “Modeling of CO₂/CH₄ gas mixture permeation and CO₂ induced plasticization through an asymmetric cellulose acetate membrane”, *RSC Advances*, v. 6, n. 20, pp. 16561–16567. doi: <10.1039/c5ra23506e>.

SAHLI COSTABAL, F., YANG, Y., PERDIKARIS, P., et al., 2020, “Physics-Informed Neural Networks for Cardiac Activation Mapping”, *Frontiers in Physics*, v. 8, n. February, pp. 1–12. doi: <10.3389/fphy.2020.00042>.

SANDLER, S. I., 2017, *Chemical, Biochemical and Engineering Thermodynamics*. 5th ed. Delaware, John Wiley Sons, Ltd. ISBN: 9780470504796.

- SBM OFFSHORE, 2013, *Currents SBM Offshore*. Relatório Técnico 8, SBM Offshore. Availability: <<https://www.sbmoffshore.com/wp-content/uploads/2013/04/Currents-May-2013.pdf>>.
- SECCHI, A. R., WADA, K., TESSARO, I. C., 1999, "Simulation of an ultrafiltration process of bovine serum albumin in hollow-fiber membranes", *Journal of Membrane Science*, v. 160, n. 2, pp. 255–265. doi: <10.1016/S0376-7388(99)00091-5>.
- SHAHSAVAND, A., CHENAR, M. P., 2007, "Neural networks modeling of hollow fiber membrane processes", *Journal of Membrane Science*, v. 297, n. 1-2, pp. 59–73. doi: <10.1016/j.memsci.2007.03.011>.
- SIDHOUM, M., SENGUPTA, A., SIRKAR, K. K., 1988, "Asymmetric cellulose acetate hollow fibers: Studies in gas permeation", *AIChE Journal*, v. 34, n. 3, pp. 417–425. doi: <10.1002/aic.690340309>.
- SIPÖCZ, N., TOBIESEN, F. A., ASSADI, M., 2011, "The use of Artificial Neural Network models for CO₂ capture plants", *Applied Energy*, v. 88, n. 7, pp. 2368–2376. doi: <10.1016/j.apenergy.2011.01.013>.
- SMITH, A. R., KLOSEK, J., 2001, "A review of air separation technologies and their integration with energy conversion processes", *Fuel Processing Technology*, v. 70, n. 2, pp. 115–134. doi: <10.1016/S0378-3820(01)00131-X>.
- SOLA, J., SEVILLA, J., 1997, "Importance of input data normalization for the application of neural networks to complex industrial problems", *IEEE Transactions on Nuclear Science*, v. 44, n. 3, pp. 1464–1468. doi: <10.1109/23.589532>.
- TAHERI, M., MOHEBBI, A., HASHEMIPOUR, H., et al., 2016, "Simultaneous absorption of carbon dioxide (CO₂) and hydrogen sulfide (H₂S) from CO₂-H₂S-CH₄ gas mixture using amine-based nanofluids in a wetted wall column", *Journal of Natural Gas Science and Engineering*, v. 28, pp. 410–417. doi: <10.1016/j.jngse.2015.12.014>.
- TAO, F., ZHANG, M., NEE, A., 2019, "Background and Concept of Digital Twin", *Digital Twin Driven Smart Manufacturing*, pp. 3–28. doi: <10.1016/b978-0-12-817630-6.00001-1>.
- TERRON, L. R., 2009, *Termodinâmica química aplicada*. 1st ed. São Paulo, Manole. ISBN: 9788520420829.

- ULLAH, S., ASSIRI, M. A., AL-SEHEMI, A. G., et al., 2019, "High-temperature CO₂ removal from CH₄ using silica membrane: experimental and neural network modeling", *Greenhouse Gases: Science and Technology*, v. 9, n. 5, pp. 1010–1026. doi: <10.1002/ghg.1916>.
- VAN GAUWBERGEN, D., BAEYENS, J., 2000, "Macroscopic fluid flow conditions in spiral wound membrane elements: Packed bed approach", *Water Science and Technology*, v. 41, n. 10-11, pp. 85–91. doi: <10.2166/wst.2000.0615>.
- VIEIRA, P. L., GARCIA, C. B., TORRES, E. A., et al., 2005, *Gás Natural: Benefícios Ambientais do Gás Natural no Estado da Bahia*. Salvador, Salvador: Solisluna Design e Editora.
- VON BARTHELD, C., BAHNEY, J., HERCULANO-HOUZEL, S., 2016, "The Search for True Numbers of Neurons and Glial Cells in the Human Brain: A Review of 150 Years of Cell Counting: Quantification of Neurons and Glia in Human Brain", *Journal of Comparative Neurology*, v. 524, n. 18, pp. 3865–3895. doi: <10.1002/cne.24040>.
- WHITAKER, S., 1977, "Fundamental Principles of Heat Transfer". 1 ed., Pergamon Press, Inc., p. 556. ISBN: 0-08-017866-9.
- WIDROW, B., HOFF, M. E., 1960, "Adaptive switching circuits", *IRE Wescon Conference Record*, pp. 96–104.
- WILKE, C. R., 1950, "A viscosity equation for gas mixtures", *The Journal of Chemical Physics*, v. 18, n. 4, pp. 517–519. doi: <10.1063/1.1747673>.
- XUE, B., YU, Y., CHEN, J., et al., 2017, "A comparative study of MEA and DEA for post-combustion CO₂ capture with different process configurations", *International Journal of Coal Science and Technology*, v. 4, n. 1, pp. 15–24. doi: <10.1007/s40789-016-0149-7>.

Appendix A

Membrane diffusion equation

As mentioned in Section 2.1.3 Fick's law of material diffusion can be described by the following equation where D^{AB} is the diffusion coefficient and J is the material flux under a concentration gradient (dc/dh) through a membrane sheet of the thickness (h).

$$J = -D^{AB} \frac{dc}{dh} \quad (\text{A.1})$$

However when comes to considering gas separation, concentration can be rewritten as a function of partial pressure (p_i) according to Henry's law which:

$$c_i = S_{ii} \quad (\text{A.2})$$

where S_i is the i^{th} coefficient solubility inside membrane, which gives:

$$J_i = -D_i^{AB} S_i \frac{dp_i}{dh} \quad (\text{A.3})$$

According to MULDER (1996) the product $D^{AB}S$ can be joint in another parameter called permeability (k_i), which finally returns:

$$J_i = -k_i \frac{dp_i}{dh} \quad (\text{A.4})$$

Concluding that an element i flux through a membrane is proportional to its permeability factor and partial pressure gradient and inversely to its thickness h .

Appendix B

MLP network back-propagation algorithm

To deduct a back-propagation algorithm weight correction from an exit layer node, let's use the perceptron's equations from section 2.4.1 and 2.4.3. First, defining the error as:

$$e_m(t) = r_m(t) - y_m(t) \quad (\text{B.1})$$

Next, according to Widrow-Hoff law and its LMS algorithm, we have the objective function to be minimized as:

$$Obj_m(t) = \frac{\sum^m e_m^2(t)}{2} \quad (\text{B.2})$$

Besides, the linear combination from all inputs can be described accordingly to equation 2.32, which adjusted to a last node neural cell, gives:

$$u_m(t) = \sum^j w_{mj}(t)y_j(t) \quad (\text{B.3})$$

However, as it is an exit layer node, its response can be defined accordingly to equation 2.31, changing the process's input signal $x_i(t)$ to a previous layer response output $y_j(t)$, which gives

$$y_k = \varphi \left(\sum_{n=1}^N w_{mj}(t)y_j(t) \right) = \varphi_m(u_m(t)) \quad (\text{B.4})$$

According to LMS algorithm, the node weight correction is proportional to the partial differentiation $\partial Obj_m(t)/\partial w_{mj}(t)$ as it follows:

$$\Delta w_{mj}(t) = -\eta \frac{\partial Obj_m(t)}{\partial w_{mj}(t)} \quad (B.5)$$

Which, according to chain rule, $\partial Obj_m(t)/\partial w_{mj}(t)$ can be rewritten as:

$$\frac{\partial Obj_m(t)}{\partial w_{mj}(t)} = \frac{\partial Obj_m(t)}{\partial e_m(t)} \frac{\partial e_m(t)}{\partial y_m(t)} \frac{\partial y_m(t)}{\partial u_m(t)} \frac{\partial u_m(t)}{\partial w_{mj}(t)} \quad (B.6)$$

However if we differentiate equation equation B.1 ,B.2, B.3 and B.4 with respect of $y_m(t)$, $e_m(t)$, $w_{mj}(t)$ and $u_m(t)$ respectively, it gives the following equations B.7, B.8, B.9 and B.10 respectively:

$$\frac{\partial e_m(t)}{\partial y_m(t)} = -1 \quad (B.7)$$

$$\frac{\partial Obj_m(t)}{\partial e_m(t)} = e_m(t) \quad (B.8)$$

$$\frac{\partial u_m(t)}{\partial w_{mj}(t)} = y_j(t) \quad (B.9)$$

$$\frac{\partial y_m(t)}{\partial u_m(t)} = \varphi'_m(u_m(t)) \quad (B.10)$$

Replacing equations B.7, B.8, B.9 and B.10 into equation B.6, gives:

$$\frac{\partial Obj_m(t)}{\partial w_{mj}(t)} = -e_m(t) \varphi'_m(u_m(t)) y_j(t) \quad (B.11)$$

Which, replacing onto equation B.5 gives:

$$\Delta w_{mj}(t) = \eta e_m(t) \varphi'_m(u_m(t)) y_j(t) \quad (B.12)$$

Besides, a local gradient can be defined as:

$$\delta_m(t) = -\frac{\partial Obj_m(t)}{\partial u_m(t)} \quad (B.13)$$

Which, again, accordingly to chain law:

$$\delta_m(t) = -\frac{\partial Obj_m(t)}{\partial e_m(t)} \frac{\partial e_m(t)}{\partial y_m(t)} \frac{\partial y_m(t)}{\partial u_m(t)} \quad (B.14)$$

Replacing equations B.7, B.8 and B.9 into equation B.14, gives:

$$\delta_m(t) = e_m(t) \phi'_m(u_m(t)) \quad (B.15)$$

Then, replacing equation B.15 into equation B.12 gives finally that the last node weight correction depends on learning rate, local gradient, and previous node response only, which concludes that:

$$\Delta w_{mj}(t) = \eta \delta_m(t) y_j t \quad (B.16)$$

For a non-last-node layer, equations turn out to be a bit different from the example above. However, the deduction mechanism is quite the same. Let's then pretend that j is a neural cell from a hidden layer before exit node m . Since we don't have an error measure for an internal layer, the local gradient can be written as:

$$\delta_j(t) = -\frac{\partial Obj_j(t)}{\partial y_j(t)} \frac{\partial y_j(t)}{\partial u_j(t)} \quad (B.17)$$

Replacing equation B.10 we have:

$$\delta_j(t) = -\frac{\partial Obj_j(t)}{\partial y_j(t)} \phi'_j(u_j(t)) \quad (B.18)$$

However, we still want to minimize the exit layer error, so that differentiating equation B.2 to respect to $y_j(t)$ gives:

$$\frac{\partial Obj_j(t)}{\partial y_j(t)} = \sum_{m} e_m \frac{\partial e_m(t)}{\partial y_j(t)} \quad (B.19)$$

According to chain law:

$$\frac{\partial Obj_j(t)}{\partial y_j(t)} = \sum_m e_m \frac{\partial e_m(t)}{\partial u_m(t)} \frac{\partial u_m(t)}{\partial y_j(t)} \quad (B.20)$$

However, equation B.1 suggests that:

$$e_m(t) = r_m(t) - \phi_m(u_m(t)) \quad (B.21)$$

Which differentiating with respect to $u_L(t)$ gives:

$$\frac{\partial e_m(t)}{\partial u_m(t)} = -\phi'_m(u_m(t)) \quad (\text{B.22})$$

Besides, differentiating equation B.3 with respect to $y_j(t)$, gives:

$$\frac{\partial u_m(t)}{\partial y_j(t)} = w_{mj}(t) \quad (\text{B.23})$$

Then, replacing equations B.22 and B.23 into equation B.20, we have:

$$\frac{\partial Obj_j(t)}{\partial y_j(t)} = -\sum^m e_m \phi'_m(u_m(t)) w_{mj}(t) \quad (\text{B.24})$$

However, if noticed, equation B.24 contains equation B.15 as the two first terms inside summation. That way, equation B.24 can be rewritten as:

$$\frac{\partial Obj_j(t)}{\partial y_j(t)} = -\sum^m \delta_m(t) w_{mj}(t) \quad (\text{B.25})$$

Which, replacing on equation B.18, gives that a prior exit layer node local gradient can be described as:

$$\delta_j(t) = \phi'_j(u_j(t)) \sum^m \delta_m(t) w_{mj}(t) \quad (\text{B.26})$$

Then, replacing equation B.26 into equation B.16 gives finally that a prior last node weight correction depends on learning rate, local gradient (which depends on local response gradient, a sum of final node local gradients and weights) and previously node response, which concludes that:

$$\Delta w_{j(j-1)}(t) = \eta \phi'_j(u_j(t)) \left[\sum^m \delta_m(t) w_{mj}(t) \right] y_{(j-1)}(t) \quad (\text{B.27})$$

Appendix C

Larger real FPSO Gas separation process diagram

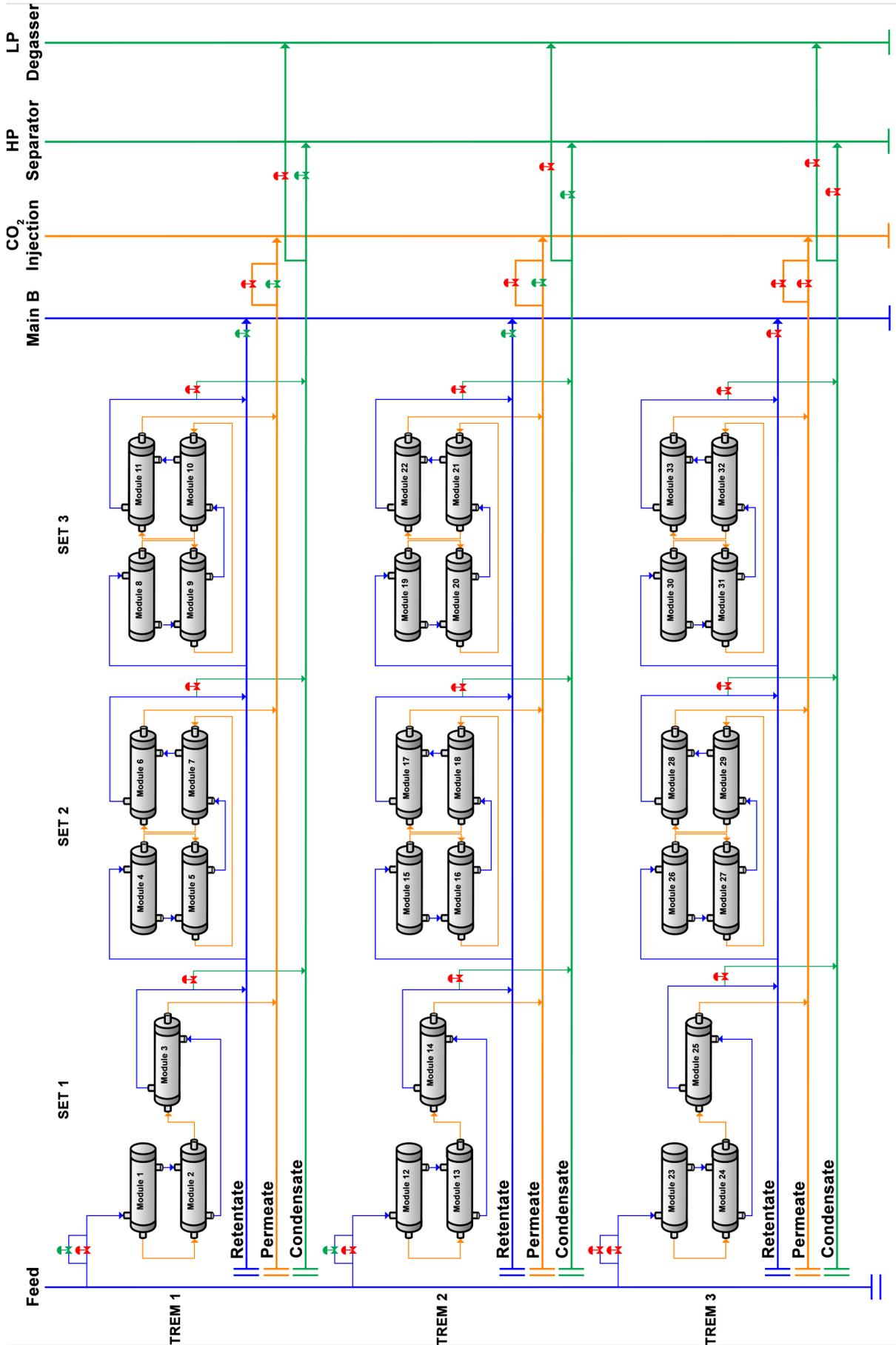


Figure C.1: Enlarged operational FPSO gas separation process diagram.

Appendix D

Fugacity calculation

Peng-Robinson Equation of State can be described by the following:

$$P = \frac{RT}{V-b} - \frac{a}{V(V+b) + b(V-b)} \quad (\text{D.1})$$

$$a = 0.45724 \frac{R^2(T^c)^2}{P^c} \left(1 + \left(0.37464 + 1.54226\omega - 0.26992\omega^2 \right) \left(1 - \sqrt{T^r} \right) \right)^2 \quad (\text{D.2})$$

$$\omega = -1 - \log_{10} \frac{P^{sat}}{P^c} \Big|_{T^r=0.7} \quad (\text{D.3})$$

$$b = 0.0778 \frac{RT^c}{P^c} \quad (\text{D.4})$$

$$T^r = \frac{T}{T^c} \quad (\text{D.5})$$

Where P are the pressure, T stands for temperature, R is the universal gas factor, V is the molar volume, T^c stands for critical temperature, P^c is the critical pressure, T^r stands for reduced temperature, ω is PITZER *et al.* (1955) acentric factor and P^{sat} is the saturation pressure. The compressibility factor can be determined by calculating the roots from the cubic equation as follows.

$$Z^3 - (1 - B)Z^2 + (A - 3B^2 - 2B)Z - (AB - B^2 - B^3) = 0 \quad (D.6)$$

Where:

$$A = \frac{a_{mix}P}{R^2T^2} \quad (D.7)$$

$$B = \frac{b_{mix}P}{RT} \quad (D.8)$$

Therefore:

$$V = \frac{ZP}{RT} \quad (D.9)$$

According to MEHL (2009) the a_{mix} and b_{mix} for a mixture of elements can be calculated using Van der Waals mixture rules where a is quadratic approximation whilst b is linear. The expression to calculate each parameter is as follows:

$$a_{mix} = \sum_{i=1}^n \sum_{j=1}^n y_i y_j a_{ij} \quad (D.10)$$

$$b_{mix} = \sum_{i=1}^n y_i b_i \quad (D.11)$$

Where:

$$a_{ij} = (1 - k_{CO_2-CH_4}) \sqrt{a_i a_j} \quad (D.12)$$

Where y_i is the molar fraction from i^{th} component in the mixture. Also, according to SANDLER (2017) the vapor phase fugacity coefficient can be given by a combination of the mixture rules of Van der Waals and Peng-Robinson equation of state. This expression is given by the following:

$$\ln \phi_i^v = \frac{b_i}{b_{mix}} (Z - 1) - \ln (Z - B_{mix}) - \frac{A_{mix}}{\sqrt{8}B_{mix}} \left(\frac{2 \sum_{j=1}^n y_j a_{ij}}{a_{mix}} - \frac{b_i}{b_{mix}} \right) \ln \left(\frac{Z + (1 + \sqrt{2})B_{mix}}{Z + (1 - \sqrt{2})B_{mix}} \right) \quad (D.13)$$

Table D.1: Peng-Robinson thermodynamic parameters for the fugacity calculation

| Parameter | CO_2 | CH_4 |
|-------------------------|--------|--------|
| $P^c(\text{bar})^{(1)}$ | 73.9 | 46.3 |
| $T^c(K)^{(1)}$ | 304.1 | 190.4 |
| $\omega^{(1)}$ | 0.239 | 0.011 |
| $k_{CO_2-CH_4}^{(2)}$ | 0.103 | |

Source: (1)TERRON (2009); (2)LI (2008)

Finally, the fugacity of each component is given by the equation:

$$f_i = y_i \phi_i^v P \quad (D.14)$$

Table D.1 are the parameters used to calculate each fugacity in this work.

Appendix E

Model discretization procedure

Following the idea integrating Equations 3.11 to 3.14 under the finite volume represented by Figure E.1, gives the following resolution for mass balance for the shell side:

$$\int_w^e \frac{\partial u_{zi}}{\partial z} dz - D_i^{ABs} \int_w^e \frac{\partial}{\partial z} \frac{\partial u_{zi}}{\partial z} dz = - \int_w^e \pi D_o N P (P_x X_i - P_y Y_i) dz \quad (E.1)$$

$$u_{zi}|_e - u_{zi}|_w - D_i^{ABs} \left[\frac{\partial u_{zi}}{\partial z} \Big|_e - \frac{\partial u_{zi}}{\partial z} \Big|_w \right] = - \pi D_o N P (P_x X_i - P_y Y_i) \Big|_p \Delta z \quad (E.2)$$

However:

$$\frac{\partial u_{zi}}{\partial z} \Big|_e \approx \frac{u_{zie} - u_{ziP}}{\Delta z} \quad (E.3)$$

$$\frac{\partial u_{zi}}{\partial z} \Big|_w \approx \frac{u_{ziP} - u_{ziw}}{\Delta z} \quad (E.4)$$

Replacing Equations E.3 and E.4 into Equation E.2, changing all the terms on

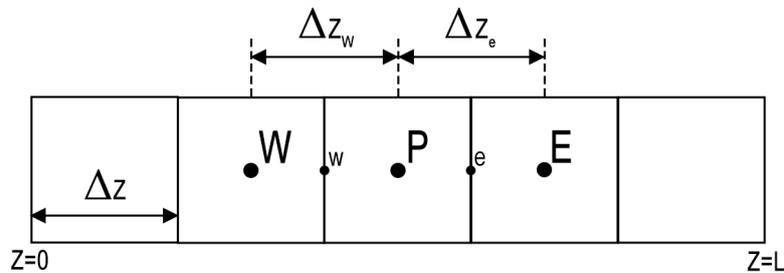


Figure E.1: General 1D mesh representation of finite-volume method, with the central volume P and the neighboring volumes W and E. The border midpoints are represented by w and e, respectively.

the right side of the equation by B_p , and manipulating some variables gives the following:

$$\left[\frac{2D_i^{ABs}}{\Delta Z} \right] u_{ziP} = \left[\frac{D_i^{ABs}}{\Delta Z} - 1 \right] u_{zie} + \left[\frac{D_i^{ABs}}{\Delta Z} + 1 \right] u_{ziw} = B_p \quad (E.5)$$

Rewriting the results in a linear system as shown in Equation E.6.

$$A_p u_{ziP} = A_e u_{zie} + A_w u_{ziw} + B_p \quad (E.6)$$

where:

$$A_p = \frac{2D_i^{ABs}}{\Delta Z} \quad (E.7)$$

$$A_e = \frac{D_i^{ABs}}{\Delta Z} - 1 \quad (E.8)$$

$$A_w = \frac{D_i^{ABs}}{\Delta Z} + 1 \quad (E.9)$$

$$B_p = -\pi D_o N P_p (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta Z \quad (E.10)$$

Likewise, this process, the same was made for the other two model equations, which resulted in the following:

$$\left[\frac{2D_i^{ABs}}{\Delta Z} \right] v_{ziP} = \left[\frac{D_i^{ABs}}{\Delta Z} - 1 \right] v_{ziw} + \left[\frac{D_i^{ABs}}{\Delta Z} + 1 \right] v_{zie} + \pi D_o N P_p (P_{xP} X_{iP} - P_{yP} Y_{iP}) \Delta Z \quad (E.11)$$

$$P_{xe}^2 = P_{xw}^2 - \frac{384ND_o(D + ND_o)RT\mu_m}{\pi(D^2 - ND_o^2)^3 Z_y} \sum_{i=1}^2 (v_{ziP}) \Delta Z \quad (E.12)$$

$$P_{ye}^2 - P_{yw}^2 = -\frac{256RT\mu_m}{\pi D_i^4 Z_x N} \sum_{i=1}^2 (u_{ziP}) \Delta Z \quad (E.13)$$

where the properties on each control volume center are calculated using the av-

erage of the properties on each control volume faces, Equation E.14 shows an example of center properties calculation for shell side pressure:

$$P_{xP} = \frac{P_{xe} + P_{xw}}{2} \quad (\text{E.14})$$

Appendix F

Raw dataset X Treated dataset chart comparison

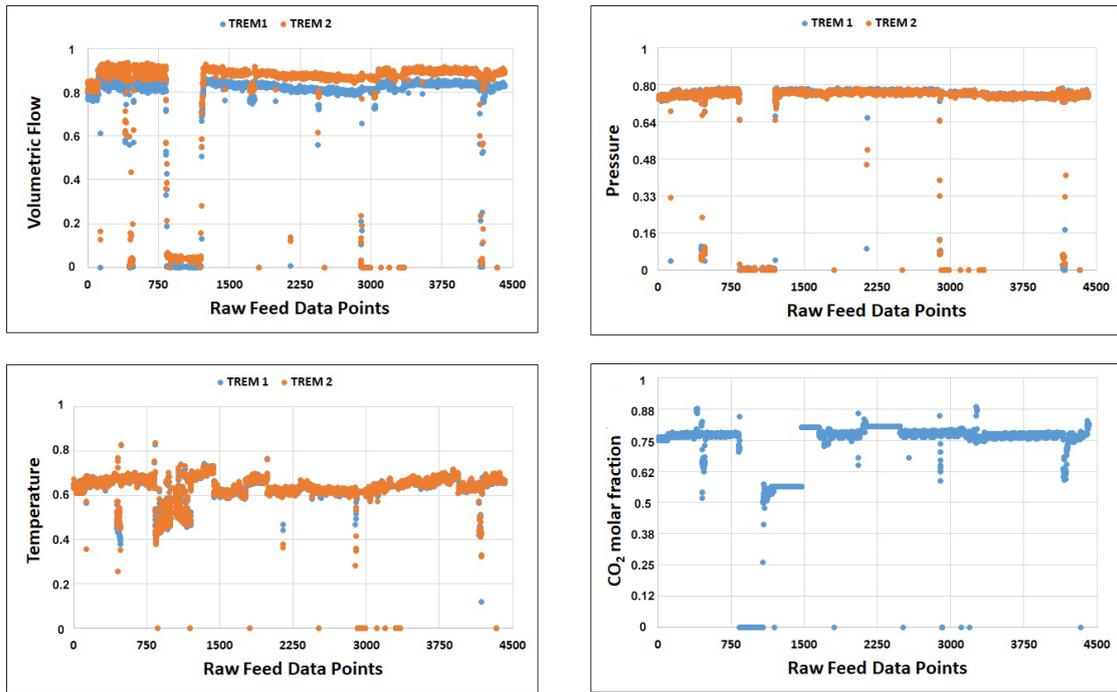


Figure F.1: Normalized raw data sheet from feed variables. Time step=1hour.

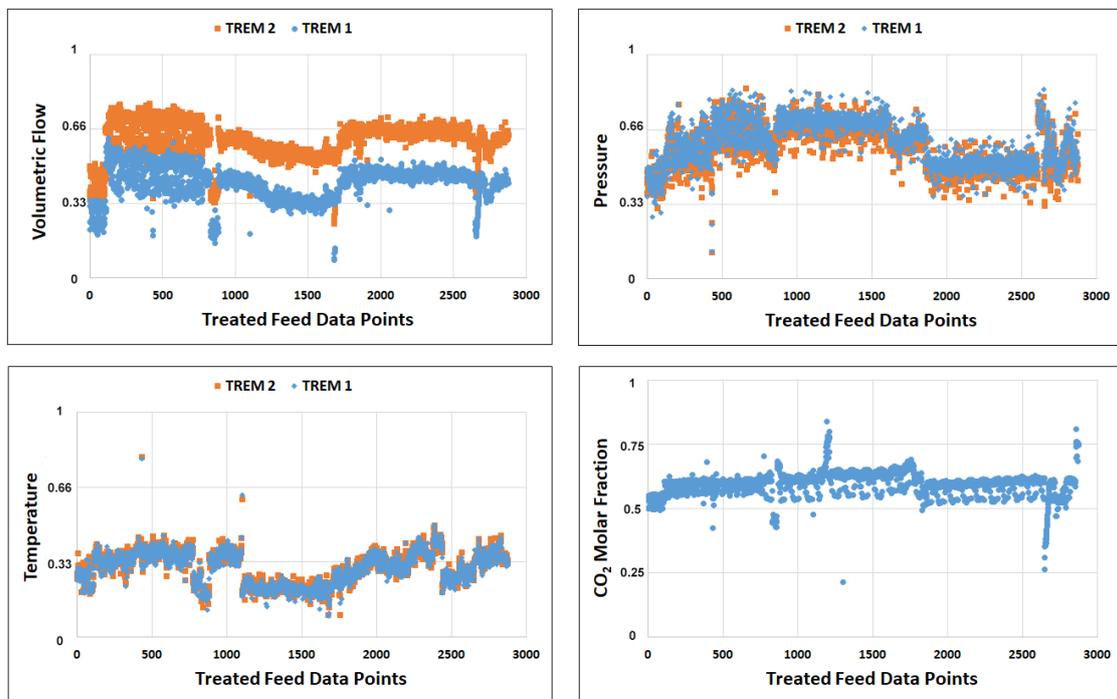


Figure F.2: Normalized treated data dispersion from feed variables.

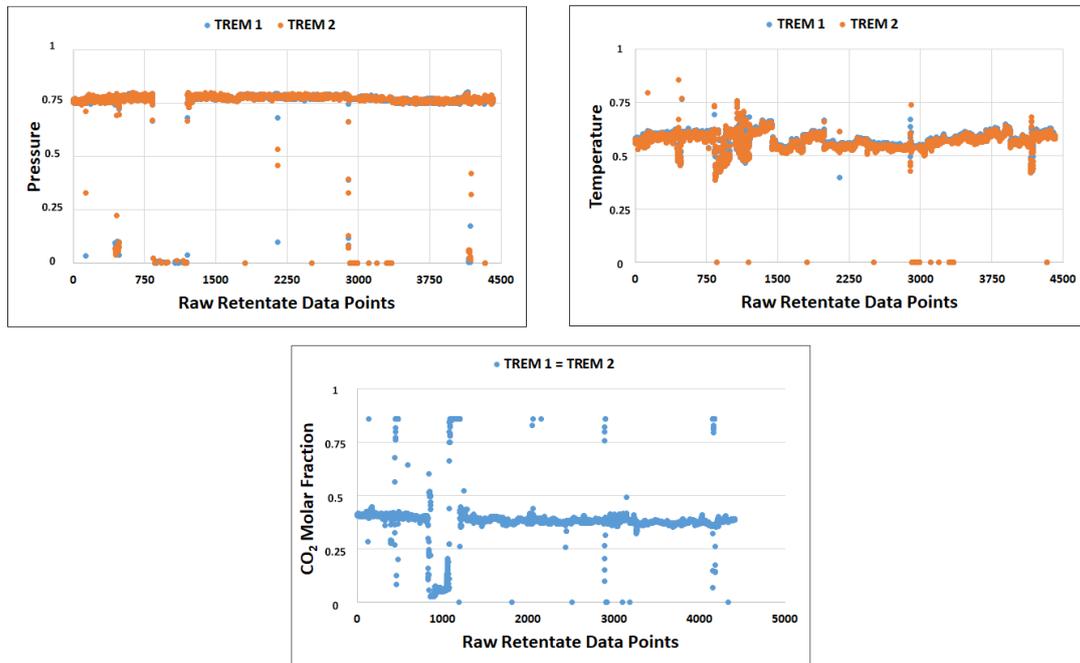


Figure F.3: Normalized raw data sheet from retentate variables. Time step=1hour.

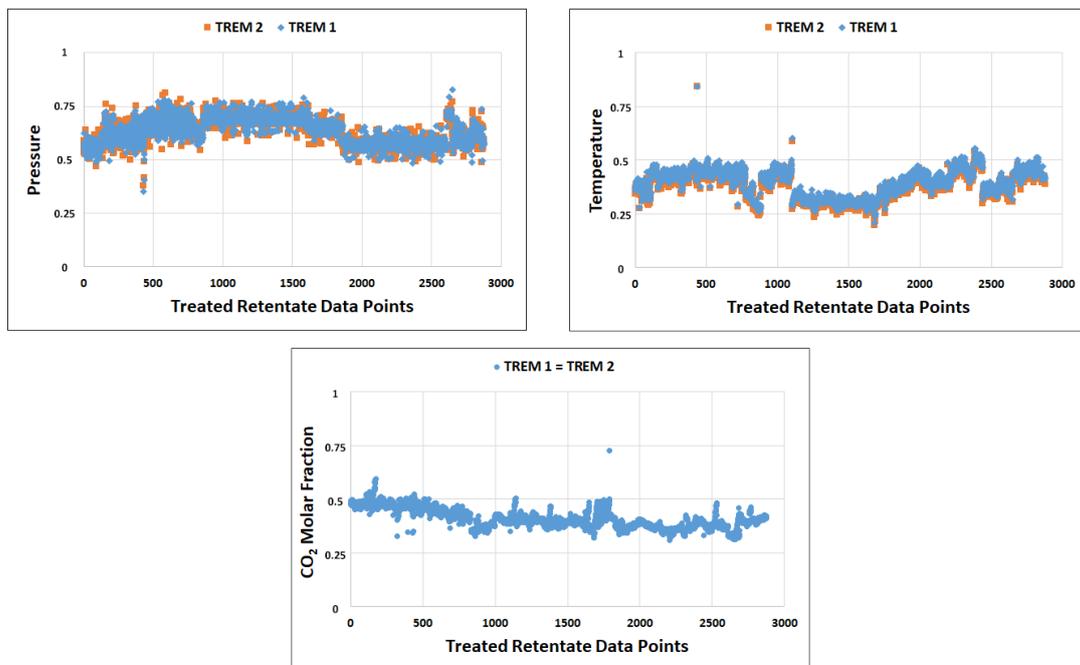


Figure F.4: Normalized treated data dispersion from retentate variables.

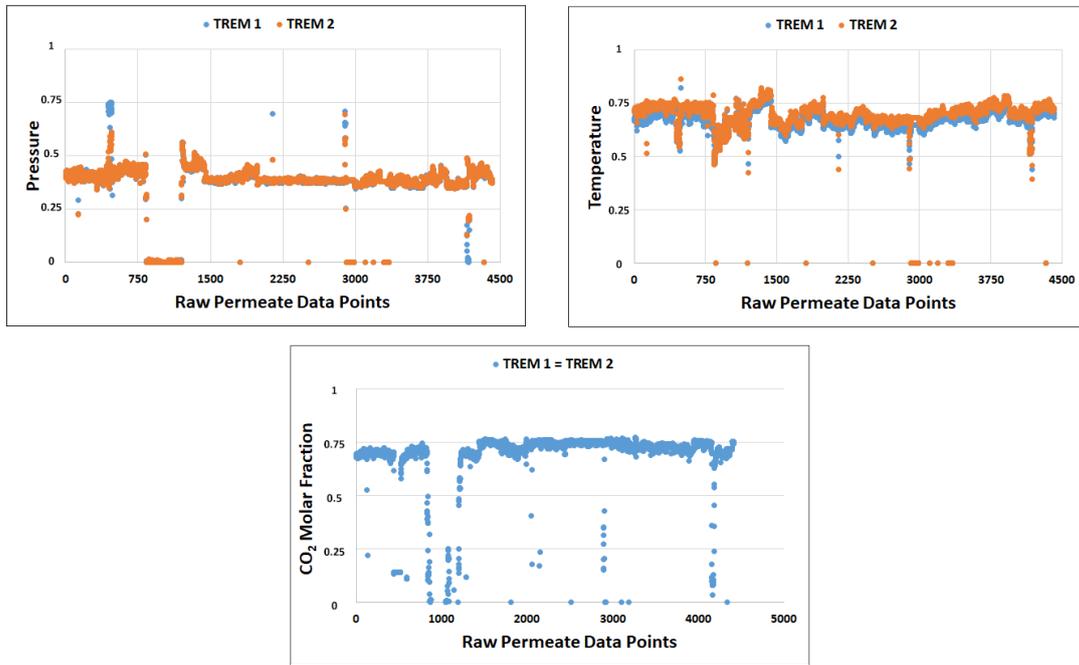


Figure F.5: Normalized raw data sheet from permeate variables. Time step=1hour.

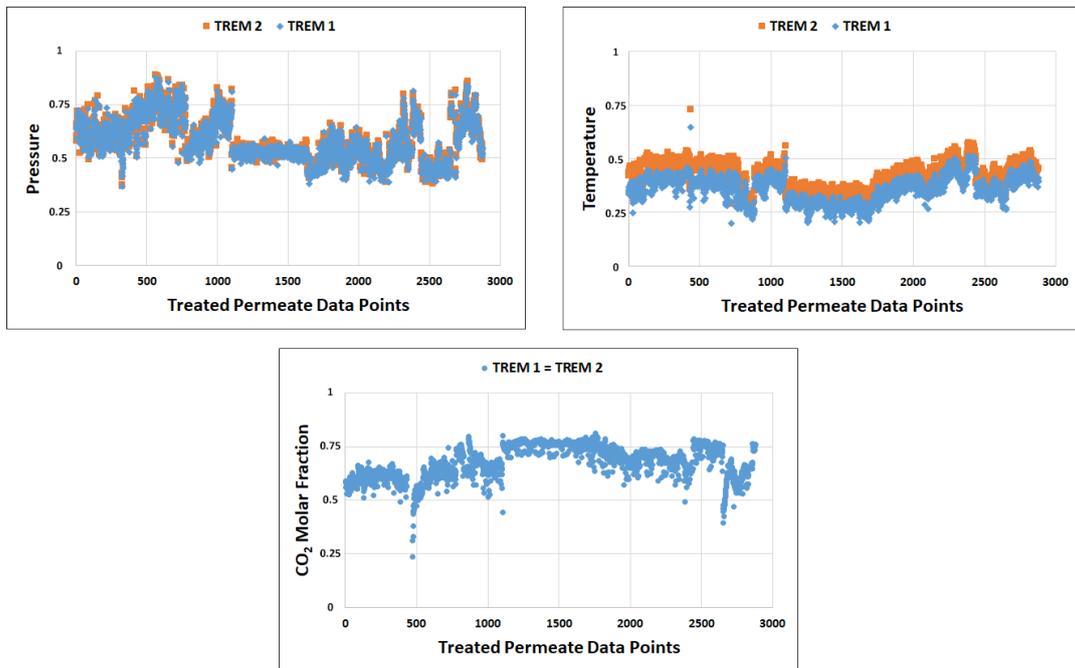


Figure F.6: Normalized treated data dispersion from permeate variables.