Universidade Federal do Rio de Janeiro
Escola de Química
Tecnologia dos Processos Químicos e Bioquímicos

Evaluation of the gaseous mixture N_2/CH_4 separation by packed bed adsorption using PSA technology

Hermes Ribeiro Sant' Anna

Rio de Janeiro
2017

Hermes Ribeiro Sant' Anna

Evaluation of the gaseous mixture N$_2$/CH$_4$ separation by packed bed adsorption using PSA technology

Dissertation submitted to the Faculty of the Chemical and Biochemical Process Technologies of Federal University of Rio de Janeiro in partial fulfillment of the requirements for the Degree of Master of Science (MSc).

Advisers:

Amaro Gomes Barreto Jr., D.Sc.

Frederico Wanderley Tavares, D.Sc.

Rio de Janeiro
2017

Avaliação da separação da mistura gasosa N2/CH4 via adsorção em coluna de leito fixo usando a tecnologia PSA.

Hermes Ribeiro Sant' Anna

Dissertação submetida ao Programa de Engenharia de Processos Qu[imicos e Bioquímicos da Universidade federal do Rio de Janeiro como parte dos requisitos necessários para a obtenção de grau de mestre em ciência.

Orientadores:

_____

Prof. Amaro Gomes Barreto Jr., D.Sc.

_____

Prof. Frederico Wanderley Tavares, D.Sc.

Committee Members:

_____

Luis Fernando Lopes Rodrigues Silva, D.Sc.

_____

Márcio Luis Lyra Paredes, D.Sc.

_____

Eduardo Rocha de Almeida Lima, D.Sc.

Rio de Janeiro
2017

UFRJ

Evaluation of the gaseous mixture N_2/CH_4 separation by packed bed adsorption using PSA technology.

Hermes Ribeiro Sant' Anna

Advisers: Amaro Gomes Barreto Jr., D.Sc. and Frederico Wanderley Tavares, D.Sc.

Abstract: Natural gas plays a key role among the national and worldwide sources of energy. Wellhead natural gas may contain a number of different contaminants, which requires removal prior to sending the gas to the pipeline and the final consumer. N_2 is an inert contaminant and its presence may cause inefficiency and energy waste in the pipeline and compression units downstream of the Gas Processing Unit (GPU). Nitrogen rejection units (NRU) can employ three different technologies: cryogenic distillation, pressure swing adsorption (PSA) and membranes. Cryogenic distillation is the most mature technology. However, since it operates above atmospheric pressure (14 to 27 barg) and at low temperature (-180 to -150 °C) this process is intensive in energy and capital. In this work, we analyze the technical viability of a PSA bed packed with silicalite to treat the waste nitrogen stream arising from a cryogenic distillation plant in order to reduce the cryogenic unit's size and increase its temperature. The governing equations of packed bed adsorption forms a system of differential algebraic equations (DAE), and can be solved using the method of lines. The method of lines uses two numerical methods, one for the time domain and another for the spatial domain. This work compares four different finite volume method (FVM) schemes to solve the DAE in space (UDS, VanLeer, Superbee and WENO) while using the DASSL method to solve the DAE in the time domain. WENO is the most adequate method in terms of numerical stability, numerical dispersion and computational cost. This method also demonstrates a good agreement with experimental points. Optimizing purity and recovery of the PSA separation using the DAE model can take uo to 16 hours on a desktop PC. We therefore use artificial neural networks as surrogate models to reduce the optimization time to less than a second. Using this approach, we found that the proposed PSA cycle and material can deliver a product with 99.5% N_2 purity from a waste nitrogen stream containing 15% CH_4 and also presents a maximum nitrogen recovery of 90%.

Keywords: Adsorption, PSA, Neural Networks, Optimization.

Universidade Federal do Rio de Janeiro
Escola de Química
Tecnologia dos Processos Químicos e Bioquímicos

UFRJ

Avaliação da separação da mistura gasosa N2/CH4 via adsorção em coluna de leito fixo usando a tecnologia PSA.

Hermes Ribeiro Sant Anna

Resumo: O gás natural desempenha um papel fundamental dentre as fontes de energia nacionais e mundiais. O gás natural na cabeça de poço pode conter vários contaminantes diferentes, as quais requerem remoção antes do envio do gás para o gasoduto e para o consumidor final. N_2 é um contaminante inerte e sua presença pode causar ineficiência e desperdício de energia nos dutos e unidades de compressão a jusante da Unidade de Processamento de Gás Natural (UPGN). As unidades de remoção de nitrogênio (NRU) podem empregar três tecnologias diferentes: destilação criogênica, adsorção por variação de pressão (PSA) e membranas. A destilação criogênica é a tecnologia mais madura. No entanto, uma vez que opera em pressões acima da atmosférica (14 a 27 barg) e em baixas temperaturas (-180 a -150 ° C), este processo é intensivo em energia e capital. Neste trabalho, analisamos a viabilidade técnica de um leito de PSA empacotado com silicalita para tratar a corrente de nitrogênio residual decorrente de uma planta de destilação criogênica, a fim de reduzir o tamanho da unidade criogênica e aumentar sua temperatura. As equações governantes da adsorção em leito empacotado formam um sistema de equações algébrico diferenciais (DAE) e podem ser resolvidas usando o método de linhas. O método das linhas usa dois métodos numéricos, um para o domínio do tempo e outro para o domínio espacial. Este trabalho compara quatro esquemas diferentes do método dos volumes finitos (FVM) para resolver o DAE no espaço (UDS, VanLeer, Superbee e WENO) enquanto utiliza o método DASSL para resolver o DAE no domínio do tempo. WENO é o método mais adequado em termos de estabilidade numérica, dispersão numérica e custo computacional. Este método também demonstra uma boa concordância com os pontos experimentais. Otimizar a pureza e a recuperação da separação de PSA usando o modelo DAE pode levar até 16 horas em um PC desktop. Usamos, portanto, redes neurais artificiais como modelos substitutos para reduzir o tempo de otimização para menos de um segundo. Usando essa abordagem, descobrimos que o ciclo de PSA e o material proposto podem fornecer um produto com 99,5% de pureza de N_2 a partir de uma corrente de nitrogênio residual contendo 15% de CH_4 e também apresenta uma recuperação máxima de nitrogênio de 90%.

Palavras chave: Adsorção, PSA, Redes Neurais, Otimização.

Table of Contents

Chapter I - Introduction

Natural gas is the third most used energy source worldwide. Behind oil and coal, its consumption amounted to 701.1 billion m^3 in 2015 (631.0 million tons oil equivalent) representing 23.8% of the total energy consumed worldwide (*BP statistical review of world energy*, 2016). The heating value of natural gas usually varies from 26.1 MJ/Sm^3 to 59.6 MJ/Sm^3, where S stands for standard (0 °C and 10^5 Pa) (Guo and Ghalambor, 2012). Typical natural gas reserves contains mostly small hydrocarbons and small amounts of inorganic gas (Häring, 2008). Table I.1 shows usual natural gas components, as well as their typical and extreme concentrations. It is possible to see that although predominantly composed of methane, some natural gases may be extremely rich in nitrogen and CO_2.

Table I.1 – Typical and extreme concentration of natural gas components in % mole fraction (Häring, 2008).

Components	Typical	Extreme
Methane	80-95	50-95
Ethane	2-5	2-20
Propane	1-3	1-12
Butane	0-1	0-4
C5+	0-1	0-1
Carbon dioxide	1-5	0-99
Nitrogen	1-5	0-70
Hydrogen sulfide	0-2	0-6
Oxygen	0	0-0.2
Helium	0-0.1	0-1
Other inert gases	traces	

In this work, we deal with nitrogen-contaminated natural gases. Pipeline content of nitrogen has to meet national and international specifications for transport and commercialization. Most regulatory organs caps the inert content of fuel gases at 4% (Scholes et al., 2012). On Brazilian soil, the limit is 6% for all regions except for the northeast and north regions where the limits are respectively 8% and 18% (ANP, 2008). Such high limits are due to the Urucu field on Amazonas, and Manati field on Bahia (Madeira, 2008). Regarding worldwide occurrence, natural gas wells containing above specification nitrogen may be found in (but not limited to) Algeria, Estonia, Germany, Netherlands, Russia, Sweden, UK and USA with usual compositions ranging between 4% - 14%, although fields containing between 85%-100% exist (Jackson et al., 2005; Krooss et al., 1995; Madeira, 2008). There

are various paths to the formation of N_2 in the subsurface of sedimentary basins. It can derive immediately from the deep crust and mantle (primordial) or through reincorporation from the atmosphere into the lithosphere (recycled) (Krooss et al., 1995).

Nitrogen separation from natural gas is possible through the usage of Nitrogen Rejection Units (NRU). The natural gas industry employs NRU for two main purposes: gas upgrading and enhanced oil recovery (EOR). While the first aims to specify the natural gas to pipeline quality, the second aims to increase oil recovery on production wells through N_2 reinjection on stimulation wells (Kidnay and Parrish, 2006). The main difference of both applications is that EOR should be planned to handle with an increase in nitrogen concentration throughout the time (MacKenzie et al., 2002). There are various types of NRU such as cryogenic distillation, pressure swing adsorption (PSA), membrane permeation, solvent absorption, lean oil absorption and chelating chemicals (Kuo et al., 2012). These technologies may differentiate in terms of stages of research and commercialization. The main processes considered in the literature are cryogenic distillation, PSA and membranes where the first is the most mature (Kidnay and Parrish, 2006). The literature cites a cryogenic plant startup dating back to 1969, in the Parisian region of Alfortville, to treat natural gas coming from the Netherlands with 14% N_2 bringing it down to 2.5% (in the present work "%" stands for mole percentage, except when otherwise is specified). The major disadvantages of such technology are its energy and capital demands, which makes it only feasible to facilities processing high amounts of natural gas (> 400 MSm^3/d) for an extended amount of time (10-20 years) (Kidnay and Parrish, 2006; Lokhandwala et al., 2010). For small and medium sized wells, the only feasible technologies are membrane permeation and PSA. Table I.2 further compares these units.

Table I.2 – Comparison of cryogenic distillation, PSA and membrane NRU (Kidnay and Parrish, 2006)

Process	Feed (10^3 Nm^3/d)	Complexity	Hydrocarbon recovery	Stage
Cryogenic Distillation	>400	Complex, Continuous	Product	Mature
PSA	60-400	Simple, Semi-continuous, Multiple Beds	Reject	Early commercialization
Membrane	15-700	Simple, Continuous, Multiple Beds	Product	Early commercialization

The literature cites the study of many adsorbents to perform methane-nitrogen separation, such as activated carbon made of various substrates, zeolites with various crystalline configurations, carbon molecular sieves and metal-organic frameworks. However, almost all adsorbents' selectivity favor the adsorption of methane. Since methane is the predominant component in the natural gas, the ideal adsorbent should favor the adsorption of nitrogen over methane and present high adsorption capacity in order to yield a high purity product for large flows. To demonstrate the state of the art in PSA process technology, carbon molecular sieves with commercial applications for flow rates ranging from 65×10^3 to 425×10^3 Sm^3/d. On the other hand, one of the few known adsorbents with equilibrium selectivity favoring N_2 adsorption, CTS-1 (Molecular Gate™) is only capable of processing natural gas at flow rates up to 141×10^3 Sm^3/d. Although there are adsorbents with kinetic selectivity favoring the adsorption of N_2, no commercial applications using such materials are known so far (Tagliabue et al., 2009). A similar problem occurs with membrane separation. Membranes can be manufactured to permeate selectively N_2 ($\alpha_{N_2,CH_4} \approx 2.5$) or CH_4 ($\alpha_{CH_4,N_2} \approx 4$). However, an ideal membrane to perform commercially suitable single stage gas separation should present a nitrogen selectivity of 6 or a methane selectivity of 17 (Baker and Lokhandwala, 2008). Therefore, commercial nitrogen separation plants are composed of bed arrangements ranging from two to four beds (NitroSep™), depending on the feed composition. Although the company states that this technology can be scaled up to 2800×10^3 Sm^3/d (MTRI, 2017), the literature reports commercial applications with inlet flow rates only up to 340×10^3 Sm^3/d (Lokhandwala et al., 2010).

Chapter I.1- Motivation

One of the main difficulties in separating the binary mixture of nitrogen and methane is the similarity in their molecular properties such as kinetic diameter, dipole moment and polarizability (Tagliabue et al., 2009). Cryogenic distillation is an efficient separation technique due to the difference in the molecules boiling points 75.36 K (-197.8 °C) and 111.5 K (-161.6 °C) respectively. However, another work in the literature points out that one of the main challenges of cryogenic distillation processes is the difficulty in obtaining a reject with low methane concentration, keeping them below 3% (MacKenzie et al., 2002). Minimizing this concentration incurs in the need to operate the process in extremely low temperatures with a great compression energy input (MacKenzie et al., 2002). However, because most adsorbents selectively retain methane over nitrogen on their surfaces, it is favorable to purify the aforementioned rejected stream using PSA. Such approach can lead to the relaxation on the working conditions of the cryogenic unit, allowing it to produce a reject stream

with higher methane concentration, consequently increasing the operating temperature and reducing the energy consumption.

Chapter I.2 - General Objectives

This work focus on analyzing numerically the separation of the binary mixture N_2/CH_4, using a fixed bed packed with the adsorbent silicalite. A similar unit has been analyzed in the literature for the separation of methane rich streams (Delgado et al., 2006, 2011). In this work, we investigate the PSA performance for nitrogen rich streams. The results are separated into two main chapters. Chapter II contains the process modelling and analysis correspond to an article published in the Chemical Engineering Process: Process Intensification Journal (Sant Anna et al., 2016). It covers the phenomenological modelling equations; numerical methods to solve the system in space and time (method of lines); comparison against experimental data for breakthrough curves; four step PSA sequence and its boundary conditions and response surface optimization of the operating pressures for an inlet gas containing 15% methane. Chapter III contains materials about Machine learning model and optimization of a PSA unit for methane-nitrogen separation, published in the Computers and Chemical Engineering (Sant Anna et al., 2017). It regards the usage of artificial neural networks (ANN) as surrogate models to optimize the PSA process. It covers the pattern generation using the Latin Hypercube Sampling technique (LHS); training of ANN surrogate models; validation of the ANN model against the phenomenological model; comparison against a linear model; single objective optimization of both the ANN and phenomenological models and multi-objective optimization of the ANN surrogate model. Chapter IV presents the general conclusions and suggestions for future work on N_2/CH_4 separation process. In the end of Chapters II and III, the reader can find an addendum expanding information not published for the sake of expanding the explanation in the published articles.

References

Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.-T., 2012. Learning from data. AMLBook New York, NY, USA:

Agarwal, A., Biegler, L.T., Zitney, S.E., 2009. Simulation and Optimization of Pressure Swing Adsorption Systems Using Reduced-Order Modeling. Ind. Eng. Chem. Res. 48, 2327–2343. doi:10.1021/ie071416p

ANP, 2008. Resolução n°16 de 2008.

Baker, R.W., Lokhandwala, K., 2008. Natural gas processing with membranes: an overview. Ind. Eng. Chem. Res. 47, 2109–2121.

Baughman, D.R., Liu, Y.A., 1995. Neural networks in bioprocessing and chemical engineering. Academic Press, San Diego.

Beck, J., Friedrich, D., Brandani, S., Fraga, E.S., 2015. Multi-objective optimisation using surrogate models for the design of VPSA systems. Comput. Chem. Eng. 82, 318–329. doi:10.1016/j.compchemeng.2015.07.009

Beck, J., Friedrich, D., Brandani, S., Guillas, S., Fraga, E.S., 2012. Surrogate based optimisation for design of pressure swing adsorption systems, in: Proceedings of the 22nd European Symposium on Computer Aided Process Engineering. pp. 1217–1221.

Bergmeir, C.N., Benítez, J.M., 2012. Neural networks in R using the Stuttgart neural network simulator: RSNNS.

Biegler, L.T., Jiang, L., Fox, V.G., 2005. Recent Advances in Simulation and Optimal Design of Pressure Swing Adsorption Systems. Sep. Purif. Rev. 33, 1–39. doi:10.1081/SPM-120039562

Boukouvala, F., Hasan, M.M.F., Floudas, C.A., 2017. Global optimization of general constrained grey-box models: new method and its application to constrained PDEs for pressure swing adsorption. J. Glob. Optim. 67, 3–42. doi:10.1007/s10898-015-0376-2

BP statistical review of world energy, 2016. . British Petroleum.

Brenan, K.E., Campbell, S.L., Petzold, L.R., 1996. Numerical solution of initial-value problems in differential-algebraic equations, Classics in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia.

Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. Biometrika 503–514.

Capitanescu, F., Ahmadi, A., Benetto, E., Marvuglia, A., Tiruta-Barna, L., 2015. Some efficient approaches for multi-objective constrained optimization of computationally expensive black-box model problems. Comput. Chem. Eng. 82, 228–239. doi:10.1016/j.compchemeng.2015.07.013

Delgado, J.A., Uguina, M.A., Sotelo, J.L., Águeda, V.I., Gómez, P., 2011. Numerical simulation of a three-bed PSA cycle for the methane/nitrogen separation with silicalite. Sep. Purif. Technol. 77, 7–17. doi:10.1016/j.seppur.2010.11.004

Delgado, J.A., Uguina, M.A., Sotelo, J.L., Ruíz, B., 2006. Modelling of the fixed-bed adsorption of methane/nitrogen mixtures on silicalite pellets. Sep. Purif. Technol. 50, 192–203. doi:10.1016/j.seppur.2005.11.026

Eason, J., Cremaschi, S., 2014. Adaptive sequential sampling for surrogate model generation with artificial neural networks. Comput. Chem. Eng. 68, 220–232. doi:10.1016/j.compchemeng.2014.05.021

Fahmi, I., Cremaschi, S., 2012. Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. Comput. Chem. Eng. 46, 105–123. doi:10.1016/j.compchemeng.2012.06.006

Farooq, S., Ruthven, D.M., 1990. Heat effects in adsorption column dynamics. II: Experimental validation of the one-dimensional model. Ind. Eng. Chem. Res. 29, 1084–1090.

First, E.L., Hasan, M.M.F., Floudas, C.A., 2014. Discovery of novel zeolites for natural gas purification through combined material screening and process optimization. AIChE J. 60, 1767–1785. doi:10.1002/aic.14441

Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Prog. Aerosp. Sci. 45, 50–79. doi:10.1016/j.paerosci.2008.11.001

Forrester, A.I.J., Sóbester, A., Keane, A.J., 2008. Engineering design via surrogate modelling: a practical guide. J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ.

Galton, F., 1886. Regression towards mediocrity in hereditary stature. J. Anthropol. Inst. G. B. Irel. 15, 246–263.

Garud, S.S., Karimi, I.A., Kraft, M., 2017. Smart Sampling Algorithm for Surrogate Model Development. Comput. Chem. Eng. 96, 103–114. doi:10.1016/j.compchemeng.2016.10.006

Graciano, J.E.A., Le Roux, G.A.C., 2013. Improvements in surrogate models for process synthesis. Application to water network system design. Comput. Chem. Eng. 59, 197–210. doi:10.1016/j.compchemeng.2013.05.024

Guo, B., Ghalambor, A., 2012. Natural gas engineering handbook, 2nd ed. ed. Gulf Pub. Company, Houston, TX.

Haghpanah, R., Majumder, A., Nilam, R., Rajendran, A., Farooq, S., Karimi, I.A., Amanullah, M., 2013. Multiobjective Optimization of a Four-Step Adsorption Process for Postcombustion CO_2 Capture Via Finite Volume Simulation. Ind. Eng. Chem. Res. 52, 4249–4265. doi:10.1021/ie302658y

Häring, H.-W., 2008. Industrial Gases Processing. Wiley-VCH, Weinheim.

Hasan, M.M.F., Baliban, R.C., Elia, J.A., Floudas, C.A., 2012. Modeling, Simulation, and Optimization of Postcombustion CO_2 Capture for Variable Feed Concentration and Flow Rate. 2. Pressure Swing Adsorption and Vacuum Swing Adsorption Processes. Ind. Eng. Chem. Res. 51, 15665–15682. doi:10.1021/ie301572n

Hassibi, B., Stork, D.G., 1993. Second order derivatives for network pruning: Optimal brain surgeon. Adv. Neural Inf. Process. Syst. 164–164.

Haykin, S., 2004. A comprehensive foundation. Neural Netw. 2, 41.

Jackson, S.R., Finn, A.J., Tomlinson, T.R., 2005. New challenges for UK natural gas. Hydrocarb. Eng. 10, 27–30.

Jiang, L., Biegler, L.T., Fox, V.G., 2005. Design and optimization of pressure swing adsorption systems with parallel implementation. Comput. Chem. Eng. 29, 393–399. doi:10.1016/j.compchemeng.2004.08.014

Jiang, L., Biegler, L.T., Fox, V.G., 2003. Simulation and optimization of pressure-swing adsorption systems for air separation. AIChE J. 49, 1140–1157.

Kidnay, A.J., Parrish, W.R., 2006. Fundamentals of natural gas processing, Mechanical engineering. CRC Press, Boca Raton.

Ko, D., Siriwardane, R., Biegler, L.T., 2005. Optimization of Pressure Swing Adsorption and Fractionated Vacuum Pressure Swing Adsorption Processes for CO_2 Capture. Ind. Eng. Chem. Res. 44, 8084–8094. doi:10.1021/ie050012z

Krooss, B.M., Littke, R., Müller, B., Frielingsdorf, J., Schwochau, K., Idiz, E.F., 1995. Generation of nitrogen and methane from sedimentary organic matter: implications on the dynamics of natural gas accumulations. Chem. Geol. 126, 291–318.

Kuo, J.C., Wang, K.H., Chen, C., 2012. Pros and cons of different Nitrogen Removal Unit (NRU) technology. J. Nat. Gas Sci. Eng. 7, 52–59. doi:10.1016/j.jngse.2012.02.004

Leonard, J., Kramer, M.A., 1990. Improvement of the backpropagation algorithm for training neural networks. Comput. Chem. Eng. 14, 337–341. doi:10.1016/0098-1354(90)87070-6

LeVeque, R.J., 2002. Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge; New York.

Lewandowski, J., Lemcoff, N.O., Palosaari, S., 1998. Use of neural networks in the simulation and optimization of pressure swing adsorption processes. Chem. Eng. Technol. 21, 593–597.

Li, S., Feng, L., Benner, P., Seidel-Morgenstern, A., 2014. Using surrogate models for efficient optimization of simulated moving bed chromatography. Comput. Chem. Eng. 67, 121–132. doi:10.1016/j.compchemeng.2014.03.024

Liu, X.-D., Osher, S., Chan, T., 1994. Weighted essentially non-oscillatory schemes. J. Comput. Phys. 115, 200–212.

Liu, Y., Delgado, J., Ritter, J.A., 1998. Comparison of finite difference techniques for simulating pressure swing adsorption. Adsorption 4, 337–344.

Lokhandwala, K.A., Pinnau, I., He, Z., Amo, K.D., DaCosta, A.R., Wijmans, J.G., Baker, R.W., 2010. Membrane separation of nitrogen from natural gas: A case study from membrane synthesis to commercial deployment. J. Membr. Sci. 346, 270–279. doi:10.1016/j.memsci.2009.09.046

MacKenzie, D., Cheta, I., Burns, D., 2002. Removing nitrogen. Hydrocarb. Eng. 7, 57–63.

Madeira, A.C. da F., 2008. AVALIAÇÃO DA TECNOLOGIA DE ADSORÇÃO "PSA" PARA REMOÇÃO DE NITROGÊNIO DO GÁS NATURAL.

Mckay, M.D., Beckman, R.J., Conover, W.J., 2000. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. Technometrics 42, 55–61. doi:10.1080/00401706.2000.10485979

Montgomery, D.C., Runger, G.C., 2014. Applied statistics and probability for engineers, Sixth edition. ed. John Wiley and Sons, Inc, Hoboken, NJ.

MTRI, 2017. Nitrogen Removal from Natural Gas - NRU - Membrane Technology for the Wellhead - Membrane Technology and Research, Inc. [WWW Document]. URL http://www.mtrinc.com/nitrogen_removal.html (accessed 3.22.17).

Nguyen, A.-T., Reiter, S., Rigo, P., 2014. A review on simulation-based optimization methods applied to building performance analysis. Appl. Energy 113, 1043–1058. doi:10.1016/j.apenergy.2013.08.061

Petzold, L.R., 1982. A description of DASSL: A differential/algebraic system solver. Sci. Comput. 1, 65–68.

Powell, M., 1989. A tolerant algorithm for linearly constrained optimization calculations. Math. Program. 45, 547–566.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Kevin Tucker, P., 2005. Surrogate-based analysis and optimization. Prog. Aerosp. Sci. 41, 1–28. doi:10.1016/j.paerosci.2005.02.001

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–538.

Ruthven, D.M., Farooq, S., Knaebel, K.S., 1994. Pressure swing adsorption. VCH Publishers, New York, N.Y.

Sant Anna, H.R., Barreto, A.G., Tavares, F.W., de Souza, M.B., 2017. Machine learning model and optimization of a PSA unit for methane-nitrogen separation. Comput. Chem. Eng. 104, 377–391. doi:10.1016/j.compchemeng.2017.05.006

Sant Anna, H.R., Barreto, A.G., Tavares, F.W., do Nascimento, J.F., 2016. Methane/nitrogen separation through pressure swing adsorption process from nitrogen-rich streams. Chem. Eng. Process. Process Intensif. 103, 70–79. doi:10.1016/j.cep.2015.11.002

Scholes, C.A., Stevens, G.W., Kentish, S.E., 2012. Membrane gas separation applications in natural gas processing. Fuel 96, 15–28. doi:10.1016/j.fuel.2011.12.074

Stocki, R., 2005. A method to improve design reliability using optimal Latin hypercube sampling. Comput. Assist. Mech. Eng. Sci. 12, 393.

Streich, M., 1970. N2 removal from natural gas. Hydrocarb. Process. 49, 86.

Sundaram, N., 1999. Training Neural Networks for Pressure Swing Adsorption Processes. Ind. Eng. Chem. Res. 38, 4449–4457. doi:10.1021/ie9901731

Tagliabue, M., Farrusseng, D., Valencia, S., Aguado, S., Ravon, U., Rizzo, C., Corma, A., Mirodatos, C., 2009. Natural gas treating by selective adsorption: Material science and chemical engineering interplay. Chem. Eng. J. 155, 553–566. doi:10.1016/j.cej.2009.09.010

Webley, P.A., He, J., 2000. Fast solution-adaptive finite volume method for PSA/VSA cycle simulation; 1 single step simulation. Comput. Chem. Eng. 23, 1701–1712.

Chapter II - Methane/nitrogen separation through pressure swing adsorption process from nitrogen-rich streams

This work presents a novel approach for the treatment of nitrogen contaminated natural gas streams. Literature states that for small and medium scale production wells (< 400 MNm^3/day), such stream can be treated by PSA. For larger scale natural gas production, the only mature option available is cryogenic distillation. Computer simulations were performed to evaluate the usage of silicalite in a PSA process for the separation of a nitrogen rich N_2/CH_4 mixture. This unit is designed for coupling to a cryogenic distillation process. The differential algebraic equations system, composed of mass and energy balances as well as other constitutive equations, was solved in the spatial dimension using the finite volumes method. Numerical methods employed were analyzed for numerical dispersion and oscillation. In order to test the numerical approach, parameters and model, data from the literature was correlated. A Rotated Central Composite Design (RCCD) was performed in order to determine optimum operational pressures. In the chosen set of operational values, the process presented nitrogen purity greater than 96% and a recovery beyond 50%. The nitrogen rich outlet stream purity is specified for venting, while the methane rich stream is proper for recycling to the cryogenic process.
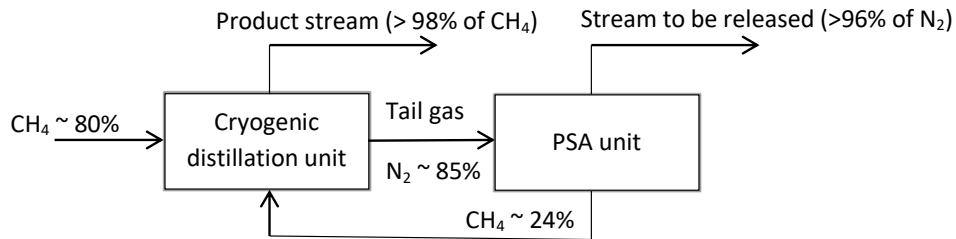
This chapter is based on the published article:

Sant Anna, H.R., Barreto, A.G., Tavares, F.W., do Nascimento, J.F., 2016. Methane/nitrogen separation through pressure swing adsorption process from nitrogen-rich streams. Chem. Eng. Process. Process Intensif. 103, 70–79. doi:10.1016/j.cep.2015.11.002

Chapter II.1 - Introduction

Rejected as a byproduct of oil production until mid-twentieth century, natural gas now accounts for 23% of world energy consumption (Kuo *et al.,* 2012). However, natural gas extracted from the well heads requires the removal of various contaminants such as CO_2, H_2O and N_2 (Kidnay & Parrish 2006). Because CH_4 and N_2 have similar physical and chemical properties, nitrogen separation from natural gas remains as a big challenge the industry faces nowadays, due to the fact that many small and medium production wells shut down for the lack of a suitable economic feasible process (Baker & Lokhandwala 2008). Being an inert gas, excess nitrogen can cause a drastic reduction of the calorific value of natural gas, in addition to increasing compression and transportation costs. About 16% of North American reserves have nitrogen as a contaminant, many of which are commercially infeasible (Kuo *et al.,* 2012, Lokhandwala *et al.,* 2010). The most mature technology for N_2 removal from natural gas is cryogenic distillation, being the only commercially feasible option for gas production in large scales (> 700 MNm^3/day) (Kidnay & Parrish 2006). Using Joule-Thomson expansion, this type of nitrogen removal unit (NRU) performs vapor-liquid equilibrium separation at temperatures as low as 90 K. However, depending on the chosen cryogenic process, as well as the presence of contaminants, the tail gas arising from this process may contain significant amounts of methane. It is known that, greater the desired methane recovery in such processes, the greater is the required energy to compress the gaseous mixtures (Mackenzie, 2002). Furthermore, it is also known that the normal boiling points of CH_4 and N_2 are respectively 111.55 K and 75.36 K. Therefore, it is easy to notice that being the most volatile component, nitrogen condensation in distillation columns requires a larger amount of energy, in order to produce a nitrogen rich tail gas, when compared to the production of methane rich products. The above mentioned translates in a way that compression power needed to operate a cryogenic distillation NRU has a high potential of reduction by relaxing the purity of the tail gas, with its post treatment with another technology such as PSA. From this point view, a PSA should be designed only to improve the nitrogen purity in tail gas stream provided by cryogenic distillation unit. For this purpose, the adsorption unit favors the usage of methane selective solids. The vast majority of the known adsorbents, such as the family of activated carbons, carbon molecular sieves and zeolites exhibit selectivity favoring the retention of methane rather than nitrogen.

Product stream (> 98% of CH$_4$)   Stream to be released (>96% of N$_2$)

CH$_4$ ~ 80% → Cryogenic distillation unit — Tail gas / N$_2$ ~ 85% → PSA unit

CH$_4$ ~ 24%

Figure II.1 – NRU composed by coupling of cryogenic distillation and PSA units used as a motivator case – compositions in molar base.

Therefore, an alternative NRU (Figure II.1) operating with the coupling of cryogenic distillation and PSA unit could lead to the reduction of the energy requirements, in order to deliver two streams with proper specification from the NRU, i.e., a product stream that contain methane specified to be transported (> 98% of methane) and a nitrogen rich stream to be released (> 96% of N$_2$). Figure II.1 shows the kind of coupling between cryogenic distillation and PSA units used in this work as a motivator case. This alternative NRU does not exist in reality, however results available in open literature (Mackenzie, 2002) and results obtained using numerical simulation by our group (data not published) have indicated that a cryogenic distillation unit can succeed to separate large amount of feed gas (greater than 5 MMNm^3/day) with CH$_4$ and N$_2$ (80:20% in molar base), if at least 98% of CH$_4$ at product stream and maximum 85% of N$_2$ at tail gas stream are required. Furthermore, the flowrate of the tail gas stream should be about ten times smaller than the flowrate of the feed gas. The main advantage of this operating condition is that the purification of nitrogen streams using cryogenic distillation is the portion which demands the lowest temperatures, then using adsorbent beds for this task has the potential to increase temperatures in which a cryogenic distillation unit can operate. Therefore, milder temperature conditions and more tolerant to CO$_2$ process, even at trace level, with smaller and less energy intensive units can be achieved by this approach.

There are many possibilities to carry out the coupling of these separation units, involving sizing of each unit, operating condition design and synthesis of the recycling approach. Clearly, the tasks above mentioned are not solved independently and are reached from the formulation and solution of a large MILP (Mixed Integer Linear Programing) constrained problem. However, in this paper we evaluate preliminarily the performance of the PSA unit alone. In this circumstance, a PSA unit should be operated to specify the tail gas stream to be released with at least 96 % of N$_2$. The other stream from the PSA unit with about 25% of CH$_4$ could be recycled into cryogenic distillation unit.

18

List of symbols

b_i : Affinity constant $[m^3 mol^{-1}]$

$b_{i,0}$: Affinity constant at reference temperature $[m^3 mol^{-1}]$

C : Total bulk phase concentration $[mol\ m^{-3}]$

$C_{p,a}$: Adsorbed phase specific heat capacity $[J\ mol^{-1}\ K^{-1}]$

$C_{p,g}$: Bulk specific heat capacity $[J\ mol^{-1}\ K^{-1}]$

$C_{p,g}$: Solid specific heat capacity $[J\ mol^{-1}\ K^{-1}]$

D_{ax} : Axial dispersion coefficient $[m^2\ s^{-1}]$

d_i : Internal diameter $[m]$

h_{in} : Inner heat transfer coefficient $[J\ m^{-2}\ K^{-1}\ s^{-1}]$

ΔH_i : Adsorption enthalpy $[J\ mol^{-1}]$

K_{ax} : Thermal dispersion coefficient $[J\ m^{-1}\ K^{-1}\ s^{-1}]$

k_i : LDF constant $[s^{-1}]$

L : Bed length $[m]$

n : Number of components

P : Pressure $[Pa]$

P_h : High pressure $[Pa]$

P_i : Intermediate pressure $[Pa]$

P_l : Low pressure $[Pa]$

$q_{i,s}$: Saturation adsorbed amount $[mol\ kg^{-1}\ or\ mol\ m^3][x]$

q_i : Adsorbed amount $[mol\ m^3]$

q_i^* : Equilibrium adsorbed amount $[mol\ m^3]$

$\bar{q}$: Equilibrium adsorbed amount $[mol\ kg^{-1}]$

Q_f : Volumetric flow rate $[m^3\ s^{-1}]$

R : Constant of ideal gases $[Pa\ m^3\ mol^{-1}\ K^{-1}]$

r_c : Cristal radius $[m]$

r_{in} : Inner bed radius $[m]$

r_p : Pellet radius $[m]$

T : Temperature $[K]$

T_w : Wall temperature $[K]$

t : Time $[s]$

t_a : Adsorption time $[s]$

t_{cnb} : Counter-current blowdown time $[s]$

t_{cob} : Co-current blowdown time $[s]$

t_p : Purge time $[s]$

v : Velocity $[m\ s^{-1}]$

y_i : Mole fraction of the ith component

Z : Axial coordinate $[m]$

Greek symbols

ε : Porosity

ε_p : Particle porosity

μ : Viscosity $[Pa\ s^{-1}]$

ρ_g : Solid density $[mol\ m^{-3}]$

ρ_p : Particle density $[mol\ m^{-3}]$

ρ_s : Solid density $[mol\ m^{-3}]$

Given the above mentioned, the main focus of this paper is to numerically evaluate the performance of a lab-scale PSA unit, in order to increase the N$_2$ purity on the waste stream and generate another stream with about 25% of CH$_4$. PSA unit sizing is not a focus of this paper. Adsorption processes for methane/nitrogen separation regarded by the literature are mainly focused on removing nitrogen from a methane rich stream. Different from usual application we manage the PSA unit to remove methane from a nitrogen rich stream. In this sense, silicalite was chosen as a selective adsorbent for methane. The PSA operational sequence analogous to a previously sequence available in the literature (Jayaraman *et al.,*, 2004) was chosen to obtain nitrogen rich stream, which is the major compound in the feed gas, on adsorption and co-current blowdown steps and CH$_4$ and N$_2$ mixture

on countercurrent blowdown step. Concerning the accomplishment of this subject, first a numerical analysis of the discretization method was carried out. The well-established adsorption bed mathematical model formed by conservation equations coupled to constitutive equations was discretized by finite volumes using different numerical schemes. The results were analyzed under convergence and computational effort criteria. Secondly, correlation between experimental breakthrough data obtained in the literature (Delgado *et al.,* 2006) and predicted curves was carried out in order to test the numerical approach, parameters and model. Thirdly, the mathematical model tested was rewritten to simulate the dynamic behavior of the different steps inherent to PSA unit, using appropriate initial and boundary conditions. Finally, this mathematical framework was used to evaluate the sensitivity of the performance variables, e.g., purity and recovery of each stream, regarding different pressure levels during pressurization, adsorption and blowdown steps.

Chapter II.2 - Methodology

The adsorption dynamics in a packed bed was carried out by numerical solution of mass, energy and momentum balances, as well as constitutive equations which describe equilibrium and kinetics of adsorption(Haghpanah et al., 2013). Some widespread assumptions in the literature were taken (Ruthven et al., 1994) [1]:

- Axially dispersed plug-flow
- Variable fluid velocity
- Local thermal equilibrium between the solid and the gas inside the pores
- Ideal gas
- Linear driving force kinetics (LDF)
- Negligible radial concentration and temperature gradients
- Constant solid properties along the axial direction

In this context, the component balances is:

$$-\frac{\partial}{\partial Z}\left(CD_{ax}\frac{\partial y_i}{\partial Z}\right) + \frac{\partial(vCy_i)}{\partial z} + \frac{\partial Cy_i}{\partial t} + \frac{(1-\varepsilon)}{\varepsilon}\frac{\partial\overline{q}_i}{\partial t} = 0 \quad i = 1,\dots,n \qquad (1)$$

Where C is the total bulk concentration, Z the axial coordinate, D_{ax} the axial dispersion coefficient, y_i the mole fraction of the ith component, v the velocity, t the time, ε the voidage fraction, $\overline{q}_i$ the

mean concentration in the adsorbed phase, and n the total number of components in the system. Moreover, the total mass balance is [2]:

$$\frac{\partial(vC)}{\partial z} + \frac{\partial C}{\partial t} + \frac{1-\varepsilon}{\varepsilon}\sum_{i=1}^{n}\frac{\partial \overline{q}_i}{\partial t} = 0 \tag{2}$$

Where the total molar concentration C is calculated using the ideal gas law:

$$C = \frac{P}{RT} \tag{3}$$

Where P is the pressure, T the temperature and R the ideal gas constant. The energy balance inside the column is described by [3]:

$$-\frac{K_{ax}}{\varepsilon}\frac{\partial^2 T}{\partial Z^2} + \frac{C_{p,g}}{R}\frac{\partial(vP)}{\partial Z} - \frac{C_{p,g}}{R}\frac{\partial P}{\partial t} + \left[\frac{1-\varepsilon}{\varepsilon}\left(\rho_s C_{p,s} + C_{p,a}\sum_{i=1}^{n}\overline{q}_i\right)\right]\frac{\partial T}{\partial t}$$

$$+ \frac{1-\varepsilon}{\varepsilon}\,C_{p,a}\,T\sum_{i=1}^{n}\frac{\partial \overline{q}_i}{\partial t} + \frac{1-\varepsilon}{\varepsilon}\sum_{i=1}^{n}\left((-\Delta H_i)\frac{\partial \overline{q}_i}{\partial t}\right) - \frac{2h_{in}}{\varepsilon r_{in}}\,(T - T_w) = 0 \tag{4}$$

Where K_{ax} is the thermal axial dispersion coefficient, $C_{p,g}$ the specific heat capacity of the bulk phase, $C_{p,s}$ the solid specific heat capacity, $C_{p,a}$ the adsorbed phase heat capacity, $-\Delta H_i$ the adsorption enthalpy of the ith component, h_{in} the inner heat transfer coefficient, r_{in} the column internal radius and T_w the wall temperature. Pressure drop along the axial dimension was modelled by:

$$-\frac{\partial P}{\partial Z} = \frac{150}{4}\frac{1}{r_p^2}\left(\frac{1-\varepsilon}{\varepsilon}\right)\mu\,v \tag{5}$$

Where r_p is the pellet radius and μ the viscosity.

The solution of the differential algebraic system shown in equations (1) – (5) requires proper initial and boundary conditions. Boundary conditions for breakthrough and PSA cycle steps are shown below.

On a breakthrough, both ends of the bed are opened, with gas flowing in the direction of the Z axis. Boundary conditions on equations (6) − (11) can also be applied on the adsorption step of a PSA cycle. For the component mass balance, Danckwerts boundary condition may apply:

$$D_{ax} \frac{\partial y_i}{\partial Z}\bigg|_{Z=0} = -v|_{z=0}(y_{i,feed} - y_i|_{z=0}) \tag{6}$$

$$\frac{\partial y_i}{\partial Z}\bigg|_{Z=L} = 0 \tag{7}$$

Where $y_{i,feed}$ is the feed mole fraction of each component and L is the bed length. By analogy of mass and heat transfer, boundary conditions for the energy balance are:

$$K_{ax} \frac{\partial T}{\partial Z}\bigg|_{Z=0} = -\varepsilon v|_{z=0}\rho_g C_{p,g}(T_{feed} - T|_{z=0}) \tag{8}$$

$$\frac{\partial T}{\partial Z}\bigg|_{Z=L} = 0 \tag{9}$$

Where ρ_g is the density of the bulk phase and T_{feed} the temperature of the feed gas. Furthermore, pressure and velocity boundary conditions are:

$$P|_{Z=L} = P_h \tag{10}$$

$$v|_{z=0} = v_{feed} \tag{11}$$

Where v_{feed} is the inlet gas feed velocity and P_h can be the operation pressure for both the breakthrough curve and the adsorption step on a PSA cycle. Regarding the pressurization step of the PSA cycle, the end at $Z = 0$ is opened for feed gas admittance, while the other end at $Z = L$ is closed. The boundaries conditions for y_i and T are the same as equations (6) − (9). As for $v|_{z=0}$, such variable is calculated based on the change in pressure on the column inlet, being $v|_{Z=L} = 0$.

Regarding the co-current blowdown step, the end at $Z = 0$ is closed while the end at $Z = L$ is opened, with gas flowing out through the latter end driven by pressure difference. The boundary conditions (7) and (9) are not affected, while (6) and (8) are reduced to:

$$\frac{\partial y_i}{\partial Z}\bigg|_{Z=0} = 0 \tag{12}$$

$$\frac{\partial T}{\partial z}\bigg|_{Z=0} = 0 \tag{13}$$

Since $v|_{Z=0} = 0$, the boundary condition for pressure is:

$$\frac{\partial P}{\partial Z}\bigg|_{Z=0} = 0 \tag{14}$$

While $v|_{z=L}$ is calculated from the change in pressure at the column outlet. For the counter current blowdown step, the boundary conditions are analogous to the latter step.

Table II.1 – Adsorbent, bed and operating conditions for numerical analysis [4].

Adsorbent	
Pellet radius, r_p (m)	0.7×10^{-3}
Particle density, ρ_p (kg m^{-3})	1070
Cristal size, r_c (m)	3×10^{-6}
Particle porosity, ε_p	0.59
PSA and Numerical Analysis bed	
Bed Length, L (m)	2
Bed internal diameter, d_i (m)	0.2
Bed voidage fraction, ε	0.52
Breakthrough operating conditions	
Axial dispersion coefficient, D_{ax} (m^2 s^{-1})	2.56×10^{-4}
Feed volumetric flow, Q_f (m^3 s^{-1})	1.0×10^{-2}
Feed CH$_4$ mole fraction, $y_{CH4,f}$	0.15
Breakthrough pressure, P (bar)	1.01
Breakthrough time, t (s)	60
Radial heat transfer coefficient, h_{in} (kW m^{-1} K^{-1})	2.23×10^{-4}
PSA operating conditions	
Axial dispersion coefficient, D_{ax} (m^2 s^{-1})	2.0×10^{-5}
Feed volumetric flow, Q_f (m^3 s^{-1})	1.0×10^{-2}
Feed CH$_4$ mole fraction, $y_{CH4,f}$	0.15
Pressurization time, t_p (s)	30
Adsorption time, t_a (s)	60
Co-current blowdown time, t_{cob} (s)	10
Counter current blowdown time, t_{cnb} (s)	90

There is a wide variety of methods for solving the above mentioned equations. Recent literature suggests that the finite volumes methods offer robustness and stability for a wide range of problems (LeVeque 2002). This is especially important for adsorption systems, due to the presence of sharp discontinuities alongside the flow, which arise from the propagation of steep mass and energy fronts. Within the framework of finite volumes methods, the present work uses the high resolution

total variation diminishing (TVD) and weighed essentially non oscillatory (WENO) as described in details elsewhere (Haghpanah *et al.,* 2013). The set of equations was converted into a non-dimensional form. The discretized system was then integrated in time using the DASSL integration routine in FORTRAN 90 (Petzold *et al.,* 1982) [5]. The bed was packed with silicalite pellets as described in the literature (Delgado *et al.,* 2006). The equilibrium and LDF constants are shown in Table II.2.

Table II.2 – Adsorption parameters for silicalite pellets (Delgado et al., 2011)

	N_2	CH_4
Affinity constant, $b_{i,0}$ ($10^{-9}Pa^{-1}$)	1.57	1.37
Maximum adsorbed amount, $q_{i,s}$ (mol kg^{-1})	1.72	2.01
Adsorption enthalpy, $-\varDelta H_i$ (kJ mol^{-1})	15.5	18.5

Regarding adsorption equilibrium, Langmuir isotherm was used.

$$q_i = \frac{q_{i,s}b_i P y_i}{1 + \sum_{j=1}^{j=n_{comp}} b_j P y_j} \tag{15}$$

Where q_i is the mass adsorbed in the particle's interface, $q_{i,s}$ is the saturation concentration and b_i the affinity constant of the ith component. Considering mass present inside the particles macropore, the isotherm becomes [6]

$$q_i^* = \frac{P y_i \varepsilon_p}{RT} + \frac{\rho_p q_{i,s} b_i P y_i}{1 + \sum_{j=1}^{j=n_{comp}} b_j P y_j} \tag{16}$$

Where q_i^* is the equilibrium adsorbed concentration in the particle. The affinity constant is temperature dependent:

$$b_i = b_{i,0}\exp\left(-\frac{\Delta H_i}{RT}\right) \tag{17}$$

Where $b_{i,0}$ is the reference affinity constant. Thus, LDF kinetics approach is regarded as:

$$\frac{\partial \overline{q}_i}{\partial t} = k_i(q_i^* - \overline{q}_i) \tag{18}$$

In this approach, k_i was estimated using a correlation presented by Farooq & Ruthven, 1990, modified for cylindrical geometry, evaluated at feed conditions. Such equations evaluate external, macropore and crystalline diffusion respectively.

$$\frac{1}{k_i} = \frac{r_p q_{i,f}}{3k_F C_f y_{i,f}} + \frac{r_p^2 \rho_p q_{i,f} \tau'}{8\varepsilon_p D_{CH_4-N_2} C_f y_{i,f}} + \frac{r_c^2}{15 D_c} \tag{19}$$

Where r_p is the particle radius, k_f is the external mass transfer coefficient, τ' is the tortuosity of the macropore network (assumed as 4), r_c is the crystal radius of zeolites and D_c is the intracrystaline diffusivity (Delgado et al., 2006). The value of constants with the subscript f are determined in the feed state. The above equation showed that the transport mechanism is dominated by diffusion inside the macropores. For the analysis performed in the present work, such parameter settled around $k_{CH_4} \cong 5$ and $k_{N_2} \cong 16$ (Delgado et al., 2006).

Chapter II.3 - Results and discussion

Chapter II.3.1 - Numerical method analysis

The present work compares four spatial discretization methods based on finite volumes schemes for N_2/CH_4 separation in packed beds; Upwind Differentiation Scheme (UDS), Superbee, van Leer and Weighed Essentially Non Oscillatory (WENO), (Haghpanah *et al.,* 2013). In order to choose a suitable method among the four cited, numerical dispersion analysis was first conducted. The column was fed with a binary mixture of nitrogen and methane, being the latter probed at column outlet. The test presented was chosen assuming isothermal profile along the column. A separated analysis regarding thermal effects is presented further in this article. Table II.1 contains parameters used in this analysis. Silicalite parameters were taken elsewhere in the literature (Delgado 2011), while bed dimensions were based in another work (Jayaraman 2004). A Peclet number of 200 was chosen. In order to establish a comparison basis, a curve generated by simulation with 1000 control volumes using WENO method is plotted while other curves were generated using 50 finite volumes. The plot in Figure II.2 shows little numerical dispersion on the finite volume methods for the given system except UDS. Moreover, numerical oscillations were observed neither in the sloped portion nor in the flat region of the breakthrough curve.
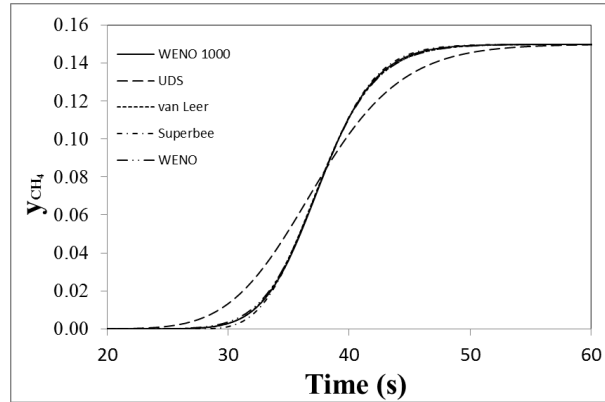
Figure II.2 – Numerical dispersion on breakthrough profiles analysis using four numerical methods: UDS, van Leer, Superbee and WENO using 50 finite volumes. WENO 1000 (WENO method with 1000 finite volumes) is used as a comparison basis for all four remaining curves [7].

Because axial dispersion was not a prohibitive aspect in three out of the four methods presented above, another important analysis concerns the time for each method to converge. Figure II.3 shows computational time for a 50, 100, 150, 200 and 250 finite volumes to simulate the breakthrough curve. It is possible to infer that all four methods present approximately a computational complexity of quadratic order. The quadratic order may be attributed to the time integrator DASSL. As for the discretization schemes, the Superbee method presented the highest computational cost, while UDS performed the fastest simulations. Taking into consideration both axial dispersion and convergence time, van Leer and WENO were the most suitable methods for the present application, being the latter used throughout the rest of the work presented here.
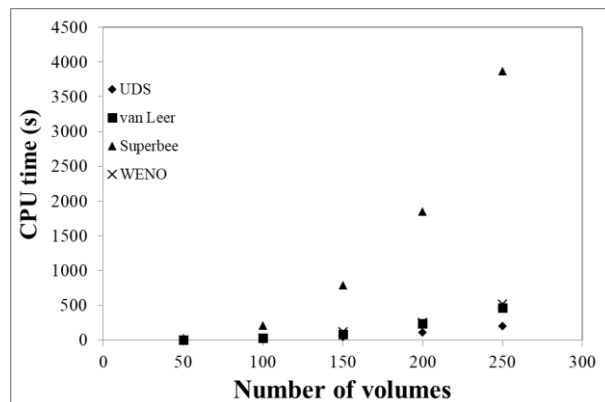
Figure II.3 – Computational time comparison of different finite volumes numerical methods [8].

The previous tests shown so far disregarded the heat generation on the bed caused by adsorption. Because adsorption of gaseous mixtures is an inherently exothermic phenomenon, adsorbed concentration in the surface of a solid tends to decrease as bed temperature increases (Ruthven 1984). Therefore, by considering a non-isothermal bed, the temperature profiles along can modify concentration, pressure and adsorbed amount profiles significantly. Figure II.4 shows a breakthrough analysis, comparing an isothermal bed with a constant wall temperature bed. Data were generated using the WENO method. It is clear in this simulation that the bed temperature decreased significantly the affinity between adsorbent and adsorbates, when compared to the isothermal case, and therefore, the breakthrough occurred faster in such case.
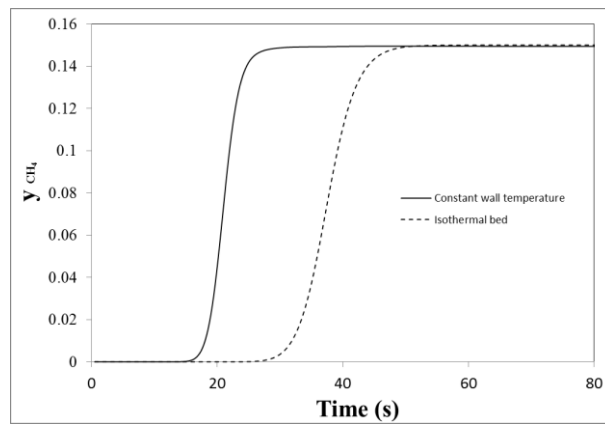
Figure II.4 – Comparison between isothermal and constant wall temperature bed

Chapter II.3.2 - Breakthrough analysis

Besides the convergence analysis, it is important to obtain parameters for real columns and to verify the correctness of the solution. The physical evaluation of the system was carried out comparing experimental breakthrough data available in the literature (Delgado *et al.,* 2006), with the predicted modelling behavior. The chosen simulation cases are shown in Table II.4, as especial column specifications for this specific analysis are present in Table II.3.

Table II.3 – Column dimensions used in the literature to collect experimental data.
Axial dispersion coefficient was calculated with correlations given by the author.
(Extracted from: Delgado et a. 2006)

Diameter, D (m)	0.016
Length, L (m)	0.163
Bed voidage fraction, ε	0.52
Axial dispersion coefficient, D_{ax} (m^2/s)	9.2×10^{-6}

Table II.4 – Feed gas conditions used in the four cases of fixed bed adsorption experiments. (Extracted from: Delgado et al., 2006)

Case	P (bar)	$y_{CH_4,f}$	Q (10^{-7} m^3 s^{-1})	T (K)
1	0.938	0.08	3.24	298
2	0.939	0.17	3.23	308
3	0.951	0.36	3.16	298
4	0.951	0.7	3.02	298

It is possible to see on Figure II.5 that the predicted behavior well fitted the experimental breakthrough data for CH$_4$ and, therefore, it becomes possible to imply that the chosen set of equations as well as the numerical methods chosen can provide a proper way to forecast the dynamics of a packed adsorption bed.
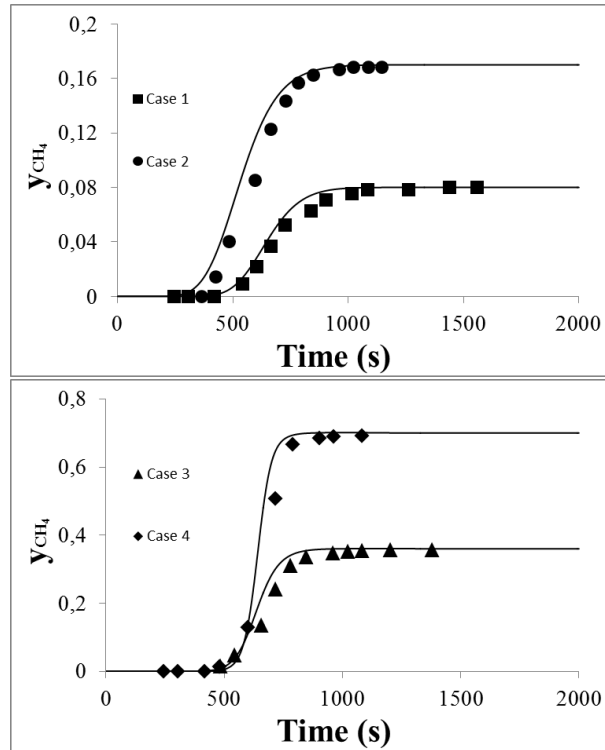
Figure II.5 – CH4 concentration profiles in the outlet of a fixed bed. Comparison between experimental data with predicted modelling behavior in four cases. Experimental points extracted in the literature (Delgado et al., 2006)

Chapter II.3.3 - PSA simulation

After verifying the system, PSA simulations were carried out in order to evaluate a bed packed with silicalite for the binary mixture separation in large scale. After preliminary studies, the chosen operational sequence, as shown schematically in Figure II.6, is:

- Pressurization: gaseous mixture is fed to the bed, with the outlet valve closed until operational pressure is reached
- Adsorption: Outlet valve is opened, letting the feed gas flow through the bed, preferentially retaining the most adsorbed species (methane)
- Co-current blowdown: having the inlet valve been closed, the bed is depressurized in the direction of the adsorption flow
- Counter current blowdown: bed is further depressurized in the opposite direction of the adsorption flow, preparing it to another sequence

Such sequence is based on a sequence presented by Jayaraman *et al.* (2004). The main objective of the present process is to decrease nitrogen concentration, since the feed stream arises from an impure tail gas of a cryogenic unit for methane concentration. Nevertheless, methane product purity in the PSA is only of secondary interest, because this stream will be later recycled to the cryogenic process in order to be further concentrated as a methane pure product. Therefore, in the designed PSA sequence, nitrogen product purity and recovery are given as priority.
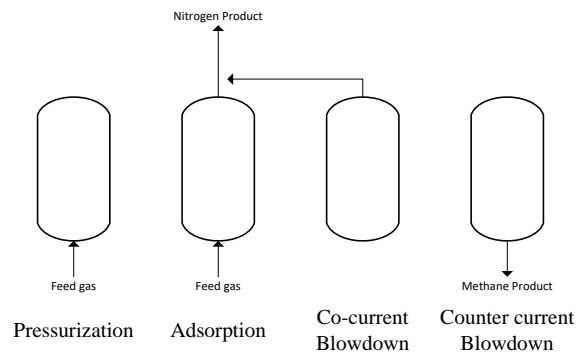
Figure II.6 – PSA sequence (Adapted from Jayaraman et al., 2004)

For the sake of demonstrating the operation of a PSA cycle, the behavior of the pressure, temperature as well as nitrogen and methane concentration on the first two steps are plotted in Figure II.7. The bed is initially filled with an inert gas at P_I, being such gas expelled from the column during the first adsorption step. As it can be seen, there is a considerable difference between the second and the first cycle. Such phenomena happens successively characterizing the equilibration phase of a PSA cycle. However, when two consecutive steps present approximately the same behavior, the Ciclic Steady State (CSS) is reached, and appreciable variation is not observed throughout the following steps. During here presented PSA simulations, it took 10 equilibration cycles in order to reach CSS.
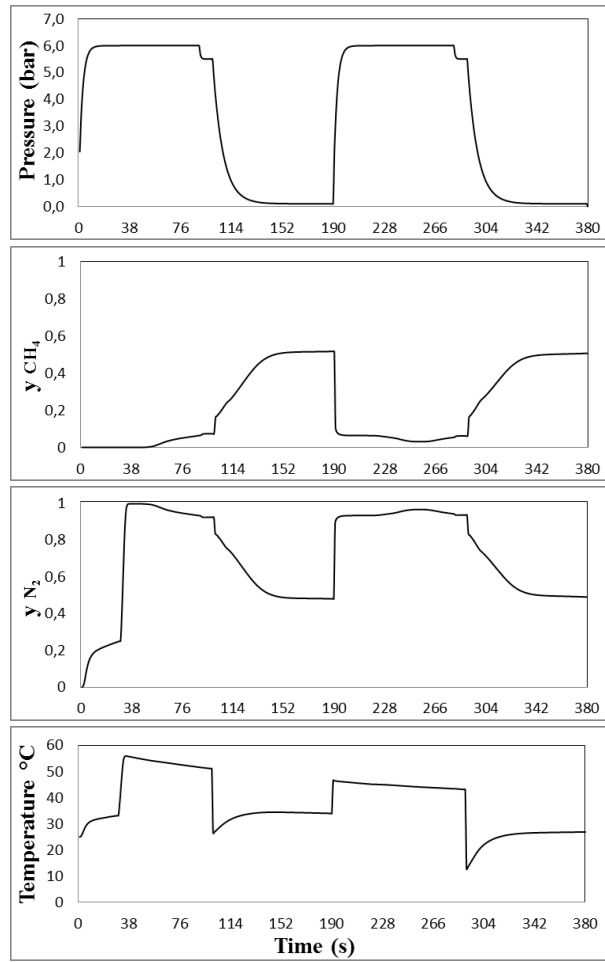
Figure II.7 - Profiles of four process variables probed at the outlet of the bed at each step: Pressure, CH4 mole fraction, N2 mole fraction and temperature. Figure shows the first two complete cycles of a PSA operation.

Table II.1 shows the operational variables of such process. Case studies were carried out in different operational conditions, in order to analyze the adsorption separation performance. The PSA cycle presented here is analyzed at room temperature. In the work by Jayaraman *et al.,* (2004), nitrogen selective adsorbent was used on a methane rich stream to remove N_2. The present work shows an analogous approach, however a methane selective adsorbent is used to treat a nitrogen rich stream to remove CH_4. For this purpose, we adapted the PSA cycle proposed by Jayaraman *et al.,* (2004) to investigate its performance using silicalite pellets. Therefore, no change in bed size, and step duration was proposed. However, a more detailed analysis regarding the cycle's pressure profile was carried out in order to maximize nitrogen purity and recovery.

Global optimization is one of the best approaches in order to maximize the objective function containing the desired dependent variables. However, PSA simulations are computationally

demanding, and this approach would require several PSA evaluations in order to find global maximum. To search over the dependent variable space for a good solution with a low amount of evaluations with a rational statistical approach, a rotatable central composite design (RCCD) (Hinkelmann & Kempthorne 2005) was employed. This approach is used when the response variables are influenced by various independent variables, in order to trace a quadratic response surface and analyze its curvature with a few set of points. This method is often used in experimental works however it is useful in several applications and was here for computational analisys. Table II.5 presents the regularized independent variables and levels as Table II.6 displays simulation generated performance points for the RCCD approach.

Table II.5 – Independent variables for the rotatable central composite design.

Independent variables	Levels				
	-1.68	-1	0	1	1.68
High Pressure, P_h (bar)	5.33	5.50	5.75	6.00	6.17
Intermediate pressure, P_i (bar)	4.33	4.50	4.75	5.00	5.17
Low pressure, P_l (bar)	0.07	0.10	0.15	0.20	0.23

Table II.6 – Dependent variables showing performance of the PSA cycles, using 16 sets of independent variables generated using RCCD criteria. The analysis was performed in order to maximize N2 purity.

P_h	P_i	P_l	CH_4 purity	CH_4 Recovery	N_2 Purity	N_2 Recovery
-1	-1	-1	0.2587	0.8687	0.9609	0.5602
-1	-1	1	0.2506	0.8249	0.9489	0.5638
-1	1	-1	0.2445	0.8871	0.9634	0.5158
-1	1	1	0.2376	0.8479	0.9517	0.5190
1	-1	-1	0.2720	0.8524	0.9589	0.5970
1	-1	1	0.2628	0.8058	0.9469	0.6004
1	1	-1	0.2575	0.8713	0.9614	0.5563
1	1	1	0.2496	0.8292	0.9497	0.5594
-1.68	0	0	0.2427	0.8622	0.9564	0.5245
1.68	0	0	0.2639	0.8326	0.9530	0.5897
0	-1.68	0	0.2658	0.8281	0.9523	0.5959
0	1.68	0	0.2428	0.8636	0.9568	0.5242
0	0	-1.68	0.2618	0.8894	0.9667	0.5571
0	0	1.68	0.2479	0.8150	0.9461	0.5627
0	0	0	0.2537	0.8470	0.9547	0.5596
0	0	0	0.2537	0.8470	0.9547	0.5596

Table II.7 – Statistical analysis correlating linear and quadratic effects of the independent variables on nitrogen purity. Std. Err stands for standard error, t(6) stands for Student's t value with 6 degrees of freedom, -95% Cnf. Lim. stands for the lower limit of the effect value with 95% confidence and 95% Cnf. Lim. is analogous to the latter. P value is a function of the statistical set used in hypothesis tests.

	Effect	Std. Err.	$t(6)$	p-value	-95% Cnf. Lim.	95% Cnf. Lim.
Mean	0.954678	0.000109	8731.445	0.000000	0.954410	0.954945
P_h Linear	-0.002001	0.000084	-23.847	0.000000	-0.002207	-0.001796
P_h Quad	-0.000002	0.000102	-0.022	0.982901	-0.000252	0.000247
P_i Linear	0.002654	0.000084	31.627	0.000000	0.002449	0.002860
P_i quad	-0.000116	0.000102	-1.138	0.298479	-0.000365	0.000133
P_l linear	-0.012028	0.000084	-143.307	0.000000	-0.012233	-0.011822
P_l quad	0.001196	0.000102	11.732	0.000023	0.000946	0.001445
P_h X P_i	0.000010	0.000110	0.095	0.927523	-0.000258	0.000279
P_h X P_l	0.000005	0.000110	0.043	0.966924	-0.000264	0.000273
P_i X P_l	0.000168	0.000110	1.532	0.176389	-0.000100	0.000436

Table II.8 – Effects estimate ignoring statistically irrelevant factors. Column labels are analogous to the ones shown on Table II.7.

	Effect	Std. Err.	$t(11)$	P value	-95% Cnf. Lim.	95% Cnf. Lim.
Mean	0.954612	0.000051	18606.36	0.000000	0.954499	0.954725
Ph Linear	-0.002001	0.000080	-25.13	0.000000	-0.002177	-0.001826
Pi Linear	0.002654	0.000080	33.33	0.000000	0.002479	0.002830
Pl linear	-0.012028	0.000080	-151.01	0.000000	-0.012203	-0.011852
Pl quad	0.001231	0.000084	14.69	0.000000	0.001046	0.001415

On a first analysis the null hypothesis, stating that the response variable does not alter its value with a change in the independent variable within the range of the tests, is statistically analyzed with a degree of confidence of 95%. With this procedure, it is possible to determine the extension and also the signal of the effect.

In Table II.8 is possible to reject the null hypothesis on four factors: high, intermediate and low pressures upon their linear factors as well as the quadratic factor of low pressure. Such factors are taken into consideration in order to construct the quadratic response surface. All other factors can be ignored due to the fact that their P values are higher than the chosen significance limit of 0.05. The final effects estimates are present in Table II.8.

According to the effect estimates given on Table II.8, Low and High pressures negatively affects N_2 purity, while intermediate pressure positively affects such dependent variable. This way, the best condition can be found by having low and high pressures in the lower bound, while keeping intermediate pressure in the upper bound. These effects can be used to model a quadratic surface of responses, giving:

$$Purity = 0.48 - 1.0 * 10^{-3}P_h + 1.3 * 10^{-4}P_i - 6.0 * 10^{-3}P_l + 6.2 * 10^{-4}P_l^2 \qquad (20)$$

As it can be seen in Figure II.8, the quadratic factor in low pressure effect is not high enough to give a local stationary point in the range of the tests, thus influencing the surface plot only with a slight curvature. It is also worthy to mention that nitrogen purity can be further increased by lowering P_l even more. However, vacuum purging the column spends a great amount of energy, so it was arbitrated that 0.1 bar will be a lower pressure limit for such step.

Furthermore, it can also be observed that the closer P_h and P_i are, the greater is N_2 purity. Nonetheless, it would be more appropriate to eliminate co-current blowdown step, leading to a work of process synthesis, which is out of the scope of the present paper. So, for the purpose of this simulation, it was also arbitrated that a minimum offset between such pressures should be 0.5 bar.
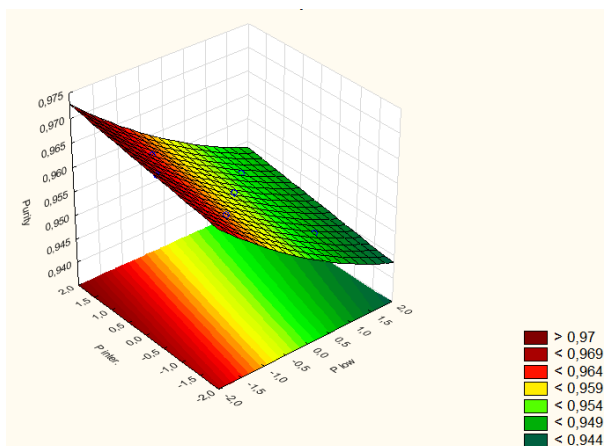
Figure II.8 – Surface generated by the quadratic model given in equation (20) keeping Ph on the lower level (6.0 bar).

Table II.9 – PSA performance

Operational conditions	
High pressure, Ph (bar)	6.0
Intermediate pressure, Pi (bar)	5.5
Low pressure, Pl (bar)	0.1
Feed velocity, vf (m/s)	0.085
Performance variables	
N_2 Purity (%)	96.61%
N_2 Recovery (%)	50.93%
N_2 Productivity (kg N_2 kg Ads^{-1} hr^{-1})	0.0417
CH_4 Purity (%)	24.42%
CH_4 Recovery (%)	89.81%
CH_4 Productivity (kg CH_4 kg Ads^{-1} hr^{-1})	0.0074

With the interest in further refining the results presented so far, the performance was evaluated by sweeping P_i values between 3.5 and 7.5 bar, while keeping P_h 0.5 bar greater than the latter variable. As it can be seen in Figure II.9, five PSA simulations were carried out, demonstrating that there is a local maximum of N_2 purity around 5.5 bar, and a local minimum N_2 recovery around 4.5 bar.
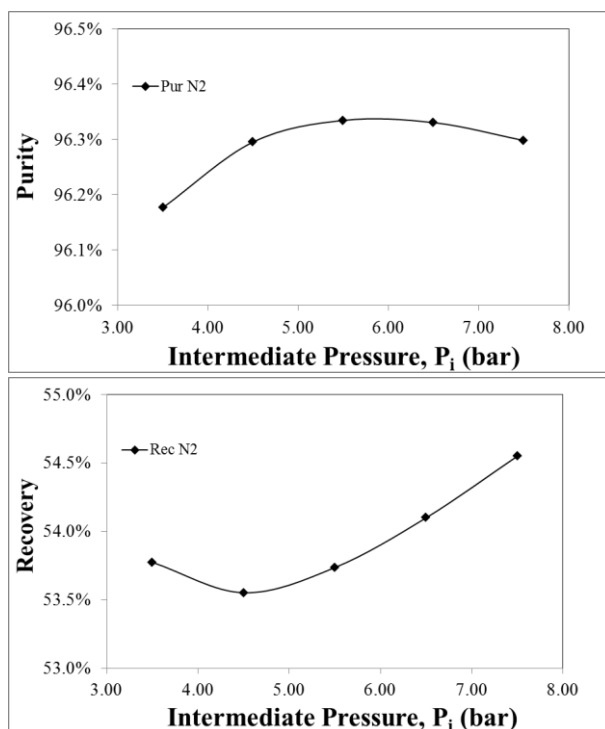
Figure II.9 – Influence of intermediate pressure on N2 purity and recovery. Each point was generated by a different PSA calculation. Lines are spline fitting of the data to improve visualization.

Another important analysis comprises the effect of feed velocity on response variables describing separation performance. Figure II.10 shows purity and recovery responses of both species by varying the given independent variable. It is possible to visualize that there is a tradeoff between N_2 purity and recovery, where the first decreases while the latter increases with inlet velocity. It is possible to conclude that the value of 0.085 m/s is a satisfactory value for allowing nitrogen purity to be greater than 96%, while keeping recovery beyond 50%. A full set of performance variables on that point can be found on Table II.9, which represents the best PSA operational conditions given the here presented analysis.
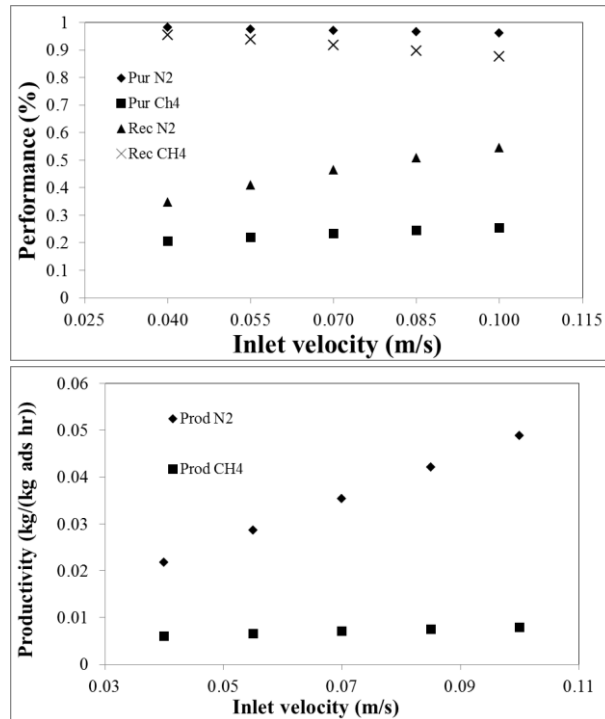
Figure II.10 – Simulated PSA cycles on different inlet velocity conditions. Influence on purity, recovery and productivity of both methane and nitrogen.

Chapter II.4 - Conclusion

In the present paper, an adsorption process was analyzed in order to mitigate the methane concentration arising from a cryogenic nitrogen removal unit tail gas. The differential algebraic equations system was verified by correlation with experimental points present elsewhere in the literature, showing good agreement between calculated and experimental data. A RCCD method was employed in order to find the best set of operational pressures to conduct a PSA separation of the binary mixture. A feed velocity analysis was also conducted, showing the existence of a tradeoff between nitrogen purity and recovery across the range of inlet flow test. Upon an inlet feed with 85% nitrogen and the selected values of pressures and velocity, the process presented nitrogen purity greater than 96% while allowing a recovery beyond 50%.

Chapter II.5 - Appendix I

The non-dimensional form of the component mass balance (21) is achieved by substituting equation (3) in (1) and also by substituting the dimensionless variables. Furthermore, non-dimensional

equations regarding total mass (22), particle (23) and energy balance (24), as well as pressure drop (25) can be found below.

$$-\frac{1}{Pe}\frac{\partial}{\partial z}\left(\frac{\overline{P}}{\overline{T}}\frac{\partial y_i}{\partial z}\right)+\frac{\partial}{\partial z}\left(\frac{\overline{P}}{\overline{T}}\,y_i\overline{v}\right)+\frac{\overline{P}}{\overline{T}}\frac{\partial y_i}{\partial\tau}+\frac{1}{\overline{T}}\,y_i\frac{\partial\overline{P}}{\partial\tau}-\frac{\overline{P}}{\overline{T}^2}\frac{\partial\overline{T}}{\partial\tau}+\psi\frac{\partial x_i}{\partial\tau}=0 \tag{21}$$

$$\frac{\partial}{\partial z}\left(\frac{\overline{P}}{\overline{T}}\,\overline{v}\right)+\psi\sum_{i=1}^{n_{comp}}\frac{\partial x_i}{\partial\tau}-\frac{\overline{P}}{\overline{T}^2}\frac{\partial\overline{T}}{\partial\tau}+\frac{1}{\overline{T}}\frac{\partial\overline{P}}{\partial\tau}=0 \tag{22}$$

$$\frac{\partial x_i}{\partial\tau}=\alpha_i(x_i^*-x_i) \tag{23}$$

$$-\Omega_1\frac{\partial^2\overline{T}}{\partial z^2}+\Omega_2\frac{\partial}{\partial z}(\overline{z}\,\overline{P})+\Omega_3\,\overline{T}\sum_{i=1}^{n_{comp}}\frac{\partial x_i}{\partial\tau}-\sum_{i=1}^{n_{comp}}\left(\sigma_i\frac{\partial x_i}{\partial\tau}\right)+\Omega_4(\overline{T}-\overline{T}_w)+\Omega_2\frac{\partial\overline{P}}{\partial\tau}+\frac{\partial\overline{T}}{\partial\tau} \tag{24}$$
$$=0$$

$$-\frac{\partial\overline{P}}{\partial z}=\frac{150}{4}\frac{1}{r_p^2}\left(\frac{1-\varepsilon}{\varepsilon}\right)^2\frac{v_0 L}{P_0}\mu\overline{v} \tag{25}$$

Dimensionless variables:

$$\overline{P}=\frac{P}{P_0}\qquad\qquad \overline{T}=\frac{T}{T_0}\qquad\qquad \overline{T}_w=\frac{T_w}{T_0}\qquad\qquad x_i=\frac{q_i}{q_{s,0}}$$

$$\overline{v}=\frac{v}{v_0}\qquad\qquad z=\frac{Z}{L}\qquad\qquad \tau=\frac{tv_0}{L}\qquad\qquad \alpha_i=\frac{k_i L}{v_0}$$

Dimensionless groups

$$Pe=\frac{v_0 L}{D_{ax}} \tag{26}$$

$$\psi=\frac{RT_0 q_{s,0}}{P_h}\frac{(1-\varepsilon)}{\varepsilon} \tag{27}$$

$$\Omega_1=\frac{\dfrac{K_{ax}}{v_0\varepsilon L}}{\dfrac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps}+q_{s,0}C_{pa}\sum_{i=0}^{n_{comp}}x_i\right)} \tag{28}$$

$$\Omega_2 = \frac{\frac{C_{pg}}{R}\frac{P_0}{T_0}}{\frac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps} + q_{s,0} C_{pa} \sum_{i=0}^{n_{comp}} x_i\right)} \tag{29}$$

$$\Omega_3 = \frac{C_{pa} q_{s,0}}{\left(\rho_s C_{ps} + q_{s,0} C_{pa} \sum_{i=0}^{n_{comp}} x_i\right)} \tag{30}$$

$$\Omega_4 = \frac{\frac{2h_{in}}{r_{in}}\frac{L}{v_0}}{\frac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps} + q_{s,0} C_{pa} \sum_{i=0}^{n_{comp}} x_i\right)} \tag{31}$$

$$\sigma_i = \frac{\frac{q_{s,0}}{T_0}(-\Delta H_i)}{\left(\rho_s C_{ps} + q_{s,0} C_{pa} \sum_{i=0}^{n_{comp}} x_i\right)} \tag{32}$$

Chapter II.6 Addendum

[1]   All of the equations, as well as most of the simplifications were all taken from the work of Haghpanah et al., 2014 that, in turn, were based upon the development presented by Ruthven, 1984. For the sake of CPU speed, the simplification made by this work is the one of constant wall temperature, while most of the literature in this field considers a refrigerating coat.

[2]   The total mass balance is the sum of all component mass balances, considering $\sum y_i = 1$. However, in the formulation of the problem there is a third component not mentioned anywhere else in this work, which is the inert component that fills the bed at $t = 0$ in the beginning of all simulations. For the purposes of this work, such inert has negligible interaction with the adsorbent.

[3]   In this equation, the thermal advection is:

$$C_{p,g}\frac{\partial(v \cdot \rho_g \cdot T)}{\partial z}$$

After applying the ideal gas law $\rho_g = P/(RT)$, we obtain the second term on the right-hand side of the energy balance equation.

[4]   We extracted the solid properties as well as parameters of the Langmuir adsorption isotherm from the work of Delgado et al., 2006 and Delgado et al., 2011. While our published article in the journal reports a D_{ax} of 1.4×10^{-2} m^2 s^{-1} for the breakthrough case, the correct value used in the simulations was 2.56×10^{-4}. We also used the bed dimensions based upon these works. We further analyze these design parameters on the next chapter. We determined the axial dispersion coefficient using a correlation provided by Delgado et al., 2006. The thermal axial dispersion coefficient comes from the work of Farooq & Ruthven 1990. The determination of the Radial heat transfer coefficient considers the axial heat transfer inside the column.

[5]   While we used the DASSL subroutine as provided by the author, we implemented our own code for the calling subroutine, which included the FVM, the boundary conditions and the PSA cycles. Regarding the DASSL subroutine, we chose an absolute tolerance of 1.0 x 10[-8] and a relative tolerance of 1.0 x 10[-6]. When using DASSL, each time step consists of a new call to the integrator, which, in our case, occurred for every 3.0 x 10[-5] seconds in simulated time. All of the remaining parameters used were DASSL's default options.

[6]   The maximum adsorbed concentration $q_{s,i}$ as provided in the table has dimension of $mol\ kg^{-1}$. However, it is important to notice that for the sake of dimensional consistency of equations (1), (2) and (4), we multiply this quantity by the solid specific mass in (16), obtaining an adsorbed concentration with dimensions of $mol\ m^{-3}$.

[7]   One of the necessities of using TVD methods and WENO is the steep discontinuity during the transition of PSA steps on the inlet. Due to the axial dispersion, the concentration profile on the outlet is already fairly dispersed. However, most methods fail to converge due to abrupt changes in all state variables in the inlet boundary condition.

[8]   On unpublished results, we compared various mesh sizes to find the better suited for our case. We found that 20-30 finite volumes presents a good convergence. Therefore, we use 50 volumes for the sake of security.

References

ARGAWAL, A. Advanced strategies for optimal design and operation of pressure swing adsorption processes, PhD thesis, Carnegie Mellon University, PA, USA, May 2010.

BAKER, W. LOKHANDWALA, K. Natural Gas Processing with Membranes: An Overview, Industrial & Engineering Chemistry Research, Vol. 47, Issue 7, 2008, Pgs. 2109-2121

DELGADO, J.A. UGUINA M.A. SOTELO, J.L. RUÍZ, B. Modelling of the fixed-bed adsorption of methane/nitrogen mixtures on silicalite pellets, Separation and Purification Technology, Vol. 50, 2006, Pgs. 192-203

DELGADO, J.A. UGUINA, M.A. SOTELO, J.L. ÁGUEDA, V.I. GÓMEZ, P. Numerical simulation of a three-bed PSA cycle for the methane/nitrogen separation with silicalite, Separation and Purification Technology, Vol. 77, Issue 1, 2011, Pgs. 7-17.

FAROOQ, S. RUTHVEN, D.M. Heat Effects in Adsorption Column Dynamics. 2. Experimental Validation of the One-Dimensional Model, Industrial & Engineering Chemistry Research, Vol. 29, 1990, Pgs. 1084-1090

HAGHPANAH, R. MAJUMDER, A. NILAM, R. RAJENDRAN, A. FAROOQ, S. KARIMI, I. A. AMANULLAH, M. Multiobjective Optimization of a four-step adsorption process for postcombustion CO2 capture via finite volume simulation, Industrial & Engineering Chemistry Research, Issue 52, 2013, Pgs. 4249–4265

HINKELMANN, K. KEMPTHORNE, O. Design and analysis of experiments. USA, Wiley-Interscience, 2008, 631 pgs.

JAYARAMAN, A. HERNANDEZ-MALDONADO, A.J. YANG, R.T. CHINN, D. MUNSON, C.L. MOHR, D.H. Clinoptilolites for nitrogen/methane separation, Chemical Engineering Science, Vol. 59, Issue 12, June 2004, Pgs. 2407-2417

KIDNAY, A.; PARRISH, W., Fundamentals of Natural Gas Processing, Taylor and Francis Group, LLC, 2006.

KUO, J.C. WANG, K.H. CHEN, C. Pros and cons of different Nitrogen Removal Unit (NRU) technology, Journal of Natural Gas Science and Engineering, Vol. 7, July 2012, Pgs. 52-59

LEVEQUE, R.J. Finite volumes methods for hyperbolic problems, UK, Cambridge University Press, 2002, 558 pgs.

LOKHANDWALA, K.A. PINNAU, I. HE, Z. AMO, K.D. DACOSTA, A. E. WIJMANS, J.G. BAKER, R.W. Membrane separation of nitrogen from natural gas: A case study from membrane synthesis to commercial deployment, Journal of Membrane Science, Vol. 346, Issue 2, 15 January 2010, Pgs. 270-279

MACKENZIE, D. CHETA, I. BURNS, D. Removing Nitrogen. Hydrocarbon Engineering, November 2002.

PETZOLD, L.R. A description of DASSL: a differential/algebraic system solver, Sandia National Laboratories, Livermore, USA, September 1982.

RUTHVEN, D.M. Principles of adsorption and adsorption processes. USA, Wiley-Interscience, 1984. 433 Pgs.

Chapter III - Machine Learning Model and Optimization of a PSA Unit for Methane-Nitrogen Separation

In this work we study the separation of N2/CH4 with a bed packed with silicalite. Pressure swing adsorption (PSA) is a competitive technology for this task. Predicting PSA performance is a time consuming computational intensive problem. Direct optimization of the system of differential algebraic equations (DAE) describing the phenomena takes an impractical amount of time. We then analyze the suitability of using artificial neural networks (ANN) as a surrogate model to predict and optimize the PSA performance. Using the ANN surrogate model, optimization time decreased from 15.7 hours to 50 seconds. We demonstrate that the PSA cycle proposed can achieve an optimized 99.5% nitrogen purity stream from an 85% inlet stream and a 50% purity stream from a 10% inlet stream. We also show that nitrogen recovery can be at most 90%. We further carry out a multi-objective optimization to demonstrate the tradeoff curve between nitrogen purity and recovery.

This chapter is based on the published article:

Chapter III.1 - Introduction

Among several contaminants that may be present in natural gas, there is a great interest in removing nitrogen, an inert component, from wellhead streams (Kidnay and Parrish, 2006; Kuo et al., 2012). Two scenarios require the usage of Nitrogen Removal Units (NRU): gas upgrade to meet pipeline specifications and Enhanced Oil Recovery (EOR). Pipeline specification for the nitrogen content is usually up to 4% (Lokhandwala et al., 2010). Different authors report high N_2 concentration wells usually ranging from 6% to 14% (Jackson et al., 2005; Madeira, 2008; Streich, 1970), although reservoirs with higher nitrogen content may occur (Krooss et al., 1995). The presence of this inert molecule on natural gas may cause several problems like heating value reduction, energy waste in compression systems and pipeline inefficiency. On the other hand, EOR is a technique to increase oil production by injecting nitrogen into stimulation wells, increasing the nitrogen content of the natural gas throughout the reservoir production cycle (Kidnay and Parrish, 2006). Usual nitrogen content for EOR applications can range from 4% to 85% (MacKenzie et al., 2002). Physicochemical separation of nitrogen and methane – the predominant component in natural gas – is a difficult task because their molecular properties such as kinetic diameter, dipole moment and critical temperature are similar (Tagliabue et al., 2009). The most frequently used NRU technology is cryogenic distillation. However, such process operates under temperatures as low as -180°C, which makes it both capital and energy intensive. Moreover, this process is economically feasible for reservoirs producing more than 400 MSm3/d. Membrane permeation and adsorption are two alternative processes employed for 15 – 700 MSm3/d and 60 – 400 MSm3/d natural gas flow rate respectively (Lokhandwala et al., 2010).

Phenomenological modeling of pressure swing adsorption (PSA) systems for gaseous separation is an established subject in the literature. Mass, energy and momentum balances, as well as constitutive equations regarding adsorption equilibrium and kinetics form a system of algebraic differential equations (DAE) (Ruthven et al., 1994). In most cases, when simplifications are infeasible, there is a need to use numerical methods to simulate gas flowing through a packed bed for separation. Since the adsorption phenomena depend on both time and space, there are two main approaches to solve the DAE system: discretizing time and space derivatives and solving them simultaneously or using a two-step method of lines (Biegler et al., 2005). This work uses the second approach by first discretizing the spatial derivatives using a finite volume method (FVM) and solving it in the time domain by a standard time integration routine (Sant Anna et al., 2016). We explain this

approach in more details throughout this text. Owing to the nature of the conservation equations, sharp fronts of concentration, temperature and pressure may propagate along the adsorption column (Haghpanah et al., 2013). These steep fronts may cause significant numerical problems, such as numerical dispersion (smearing) and oscillation since the resulting set of ODEs becomes stiff and requires a numerically stable integration routine (Biegler et al., 2005). Numerical integration of the DAE system becomes a computationally intensive problem. As we will present later, this is not due to the number of equations and control volumes, but to the complexity of the spatial discretization scheme as well as the number of steps required by the time integration routine to simulate one second of operation, in order to capture accurately the sharp fronts mentioned above.

The amount of time it takes to make multiple PSA simulations can be sometimes intractable, especially for the case of optimization, which requires multiple calls to the DAE model. In this case, one can use surrogate models (also known as substitute models, meta-models or reduced-order models). Surrogate models are functional approximations of response variables constructed with samples of the input-output space, which can arise from experimental or simulation data. There has been recently an extensive usage of such approach in engineering research (Capitanescu et al., 2015; Eason and Cremaschi, 2014; Fahmi and Cremaschi, 2012; Graciano and Le Roux, 2013). Many surrogate model techniques have been developed over the past few decades such as Polynomial Surface Response Models (PRSM), Kriging, Radial Basis Functions, Support Vector Regression and Artificial neural networks (Garud et al., 2017). A comprehensive comparison of these techniques for engineering modeling, design and optimization is available in the literature (Forrester et al., 2008; Forrester and Keane, 2009; Nguyen et al., 2014; Queipo et al., 2005). In the field of adsorptive separation processes, proper orthogonal decomposition (POD) has been analyzed for simulated moving beds model and optimization (Li et al., 2014), Kriging has been used in the design and optimization of PSA for CO_2/N_2 separation (Beck et al., 2012, 2015). Artificial Neural Networks (ANN) surrogate models have been used for optimizing the separation of N_2/Air and H_2/CO gaseous mixtures (Lewandowski et al., 1998; Sundaram, 1999). In this work, we use ANN surrogate models to optimize the PSA separation of N_2/CH_4 mixtures.

List of symbols:

b_0 : Langmuir isotherm affinity constant

cov : Covariance

C_p : Heat capacity

D_c : Crystal diffusivity

D_{col} : Column diameter

df : Degrees of freedom

$D_{i,j}$: Gaseous diffusivity

$\mathbb{E}$: Expected value

E_{BV} : Bias+variance error

E_{LT} : Repeated learning-testing error

E_{out} : Out of sample error

e_q : Unit vector in weight space

E_{test} : Testing error

E_{train} : Training error

E_{val} : Validation error

F_0 : F-value (fisher test)

$f_{1-\alpha, df_{ANN}, df_{linear}}$: F distribution percentage point

$g^{(D)}$: repeated learning-testing regression function

$\bar{g}(x)$: mean repeated learning-testing regression function

H : Hessian matrix of errors with respect to the weights

h_{ANN} : ANN regression model

k_g : Gaseous conductivity

L_q : Error increase with removal of w_q

L : Column length

lb : Lower bound (sampling. optimization)

M_w : Mole weight

$n_{i,j}$: Mole flow

P_h : Adsorption pressure

P_l : Counter-current blowdown pressure

P_m : Co-current blowdown pressure

Pur : Purity

Q : Volumetric flow

q_{max} : Langmuir isotherm maximum adsorbed concentration

r_c : Crystal radius

Rec : Recovery

r_p : Pellet radius

s^2 : Sample variance

s_j : Node activation

SS : Sum of squares

T : Temperature

t : Time

t_{ad} : Adsorption time

t_{co} : Compression time

t_{dj} : Co-current desorption time

t_{dr} : Counter-current desorption time

ub : Upper bound (sampling, optimization)

$w_{i,j}$: Connection weight

W^T : Weight matrix

x_i' : Dimensionless variable

x_j : Input vector

y_j : Target output vector

$y_{N_2,f}$: feed nitrogen mole fraction

z_j : Sanple

Z : Sample domain

Greek letters:

Δ_j : Equation error

$-\Delta H$: Adsorption enthalpy

ε : Bed voidage fraction

ε_p : Particle porosity

μ : Viscosity

$\rho_{y_n, h(x_n)}$: Pearson correlation

ρ_p : Particle density

$\hat{\sigma}^2$: Model prediction variance

σ : Standard deviation

$\theta(s_j)$: Activation function

Optimization of PSA processes is a subject of great interest by the literature over the last years (Agarwal et al., 2009; Boukouvala et al., 2017; First et al., 2014; Hasan et al., 2012; Jiang et al., 2003, 2005; Ko et al., 2005). In order to summarize the work of PSA optimization, most methods fall into one of the four categories (Biegler et al., 2005):

- Surrogate based optimization (SBO): A simplified (surrogate) model interpolates data from plant or from a more detailed model. An optimization procedure is then performed on the simplified model and the results are compared with the more detailed one.

- Black-box optimization: The optimizer searches the optimization direction on the arguments through successive calls to an input-output function, which returns the objective function evaluation. The PSA model runs inside an inner loop each time the optimizer calls the black-box function and the gradient matrix arises from finite differences.

- Complete discretization based: In this approach, the complete discretization strategy is applied where bed equations, objective functions and constraints are solved simultaneously.

- Simultaneous Tailored Optimization: In this approach, bed models are solved implicitly in order to obtain sensitivities as well as values of the constraints and objective function.

On a previous work, we have proposed a hybrid process, combining cryogenic distillation with PSA (Sant Anna et al., 2016). In this arrangement, placing a PSA cycle to purify the waste nitrogen stream could lead to reductions in energy consumption of the cryogenic process. Due to the nature of the current leaving the cryogenic unit Figure III.1, the main goal of the PSA unit was to purify the waste nitrogen stream from a methane contaminated inlet stream. We also compared three finite volumes methods (Van Leer, Superbee and WENO) for PSA modeling with each other and with the Upwind Difference Scheme, where we concluded that WENO was the most stable, robust and fastest.
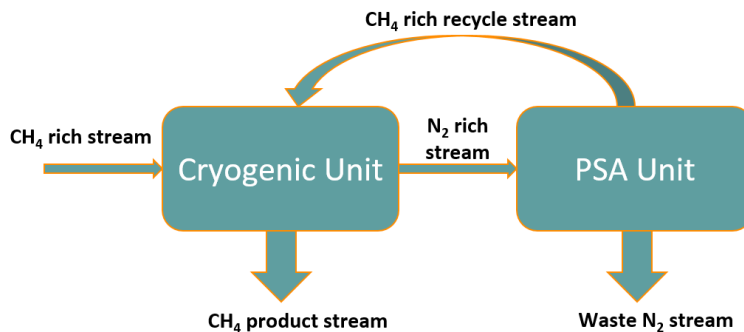
Figure III.1 – Coupled cryogenic distillation-PSA scheme

The novelty of this work relies on the construction of a surrogate model to analyze a single bed vacuum pressure swing adsorption process (VPSA) packed with silicalite zeolite for the separation of CH_4 and N_2 on a wide range of mole fractions. Here, our main interest is in producing a purified waste nitrogen stream. We divide this paper as follows: In section 2 we present details about the

DAE, numerical methods and PSA sequence used in the phenomenological modeling, including the adsorbent specifications as well as adsorption isotherms. In section 3 we present the methodology employed for the ANN surrogate model generation, including the sampling technique and the pruning algorithm to avoid overfitting. In section 4, we discuss the ideal number of samples to obtain a proper ANN surrogate model and make different error estimation procedures. In section 5, we present the optimization strategy, with comments on the performance and convergence of the selected approach as well as the s tradeoff curves on the performance variables.

Chapter III.2 - Phenomenological adsorption modeling

Pressure swing adsorption is a cyclic dynamic process comprising two major steps:

- Adsorption (Production): A gaseous mixture stream permeates the fixed bed, where the solid packing selectively retains the molecule with more affinity, while the molecule with less affinity moves quicker in the direction of the stream, producing a product purified on the second component.
- Desorption (Recovery): The pressure inside the column decreases in order to remove the retained substances in the gas/solid interface, regenerating the bed for another adsorption step.

During its operation, all system variables are constantly changing through time and space in cycles. When two or more consecutive cycles perform in a similar way, the process achieves a cyclic steady state (CSS). We found that CSS occurs on average under 5 cycles, therefore, all simulations use 20 cycles for a good safety margin. To model this phenomenon, simplifications and assumptions widely used in the adsorption field are necessary such as axially dispersed plug-flow, variable fluid velocity, local thermal equilibrium between the solid and the gas inside the pores, ideal gas, linear driving force kinetics (LDF), negligible radial concentration and temperature gradients. The system of differential algebraic equations is composed of the component mass balance, total mass balance, ideal gas law, energy balance, pressure drop, adsorption isotherm and linear driving force (LDF) mass transfer. We presented and discussed the complete set of equations as well as boundary conditions for such a problem on our previous work (Sant Anna et al., 2016) and we provide them on Appendix A, for the sake of reproducibility.

This system applies over the cylindrical shape of the packed bed. We use the method of lines to solve the system in space and time. There is a wide variety of techniques to calculate the spatial

distribution of the process variables such as orthogonal collocation, finite differences, finite elements and finite volumes (Webley and He, 2000). Since packed bed adsorption presents sharp wave fronts of concentration and temperature, numerical methods must be accurate in order to avoid unrealistic numerical oscillation and dispersion (Jiang et al., 2003). Higher order methods are especially good to capture sharp variation fronts but are prone to oscillate over flat regions. Low order methods can avoid oscillation but at the cost of smearing the high gradient fronts. Recent studies suggest that total variation diminishing (TVD) finite volume methods with flux limiters (LeVeque, 2002) and weighted essentially non-oscillatory (WENO) (Liu et al., 1994) schemes can address both of the previously presented shortcomings, which are characteristic problems of hyperbolic conservation laws. Another paper in the literature compares a simple upwind differentiation scheme (UDS), TVD methods with two types of flux limiters (Van Leer and Superbee) and WENO (Haghpanah et al., 2013). On our previous work (Sant Anna et al., 2016), we tested these four methods for N_2/CH_4 separation. WENO was the strongest candidate because of its stability, absence of smearing and computational performance. Therefore, we use such finite volumes technique (Appendix A) throughout the course of this work.

In regards to time variations of the process variables, we use DASSL time integration routine (Petzold, 1982) for initial value DAE problems with implicit variables. DASSL is a code for solving index zero and one systems of differential algebraic equations of the form:

$$F(t, y, y') = 0 \tag{1}$$

$$y(t_0) = y_0 \tag{2}$$

$$y'(t_0) = y_0' \tag{3}$$

where F, y and y' are N-dimensional vectors. This method uses a variable step size variable order fixed leading coefficient implementation of backward differentiation formulas (BDF) to advance the solution from one time step to the next. DASSL also employs an interpolant to compute the solution between mesh points (Brenan et al., 1996). These capabilities favor the convergence of time integration for stiff problems such as hyperbolic conservation laws and fast computations of advancement in time. In fact, another work compared the performance of DASSL against successive substitution method (SS) and block LU decomposition (BLUD) procedure for PSA air drying and PSA solvent vapor recovery. The author found that DASSL was twice as fast as BLUD and one order of magnitude faster than SS (Liu et al., 1998).

To summarize the two-step approach, we first discretized the axial differential terms on Appendix A while keeping the time differential terms intact. On the right-hand side of each equation, we substitute the zero term with an error term ($\Delta_j \, j = 1 \dots NEQ$). We then call the time integration routine providing exact initial values for y and y'. On the second step, the inner calculation algorithm of DASSL takes care of using the proper BDF approximation on the time derivatives in order to decrease the value of Δ_j until it satisfies the relative and absolute tolerances defined in the calling function (Brenan et al., 1996; Petzold, 1982). We implemented the first step in Fortran 90, while DASSL's authors implemented the second step in Fortran 77. Although not required by the author, it is polite to disclose that DASSL is a piece of software available in the public domain. Finally, we present on Table III.1 the simulation parameters used in the finite volumes scheme and in the calling of DASSL subroutine. The number of control volumes was formally determined in our previous work (Sant Anna et al., 2016). Since DASSL's algorithm is of variable step size, the specified time step serves the purpose of returning an equally spaced time grid. Therefore, although we retrieve $3.3 \, 10^5$ time steps per second of simulation, the routine may evaluate the DAE on even more time steps. Absolute and relative errors were determined by our accumulated experience in simulating this system, in order to guarantee stability and accuracy. These facts indicate that the main reason for the computational burden in simulating PSA is not an elevated number of equations, but a small step size and low tolerances of the DASSL time integrator in order to avoid numerical anomalies.

Table III.1 - Numerical simulation parameters used in the DAE system

Components	2
Finite Volumes	50
Equations	300
Time Step (s)	$3.0 \, 10^{-05}$
Absolute Tolerance	$1.0 \, 10^{-08}$
Relative tolerance	$1.0 \, 10^{-06}$

Chapter III.2.1 - PSA process

The process operates under four steps as on Figure III.2, namely:

- Pressurization: pressure inside the bed increases from the minimum to the highest established value, P_L and P_H respectively using the feed gas at the column inlet.
- Adsorption: feed gas with a methane mole fraction of $y_{CH4,f}$ enters the bed with a Q volumetric flow rate. This step generates the desired product, which is a purified nitrogen gas.

- Co-current blowdown: gas leaves the column in the direction of the adsorption flow by decreasing the inner pressure from P_H to P_M.
- Counter-current blowdown: the purging of the column occurs in the counter-current direction by decreasing the pressure from P_M to P_L.
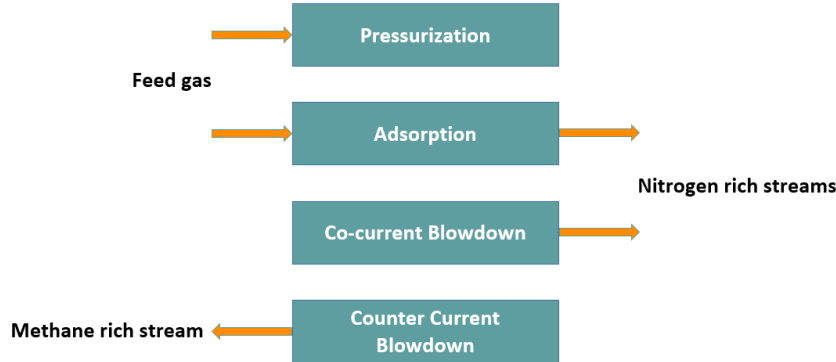
Figure III.2 – Sequence of steps performed on a pressure swing adsorption (PSA) cycle

Besides the simulation parameters, phenomenological modeling requires physical constants of the solid adsorbent and the gaseous mixture as well as variables regarding the operational parameters. The solid adsorbent used in this analysis is pelletized silicalite. Its necessary physical characteristics are pellet radius, particle density, crystal size, particle porosity and tortuosity. The packed bed requires length, diameter and voidage fraction parameters. The gaseous mixture comprises of methane and nitrogen in several mole fractions. In order to calculate the axial dispersion coefficient, as well as the LDF mas transfer constant used in the model, we need physical constants from the gases such as molar weight, viscosity, gaseous diffusivity and micropore diffusivity. Finally, in order to calculate heat transfer parameters we need thermal coefficients such as the solid and gas heat capacities as well as the thermal conductivity of the gases. We present each parameter above-mentioned on Table III.2.

Table III.2 – Solid and gases parameters for phenomenological simulation (Delgado et al., 2006)

Solid parameters	
Adsorbent	Silicalite
Pellet radius, r_p (m)	0.7×10^{-3}
Particle density, ρ_p (kg m^{-3})	1070
Cristal size, r_c (m)	3.0×10^{-6}

Particle porosity, ε_p	0.59
Tortuosity, τ'	4.00
Bed voidage fraction, ε	0.52
Specific heat, C_p (J mol K^{-1})	1000
Gas parameters (at 293K)	
N2	
Molar weight, M_w (10^{-3} kg mol^{-1})	28
Viscosity, μ (10^{-5} Pa s)	1.74
Specific heat, Cp (J mol K^{-1})	29
Gaseous diffusivity, $D_{N2/CH4}$ (10^{-5} m^2 s^{-1})	2.105
Micropore diffusivity, D_c (10^{-9} m^2 s^{-1})	1
Thermal conductivity, k_g (W m^{-1} K^{-1})	0.0255
CH4	
Molar weight, M_w (10^{-3} kg mol^{-1})	16
Viscosity, μ (10^{-5} Pa s)	1.07
Specific heat, Cp (J mol K^{-1})	36
Gaseous diffusivity, DCH4/N^2 (10^{-5} m^2 s^{-1})	2.105
Micropore diffusivity, Dc (10^{-9} m^2 s^{-1})	0.9
Thermal conductivity, kg (W m^{-1} K^{-1})	0.0333

We obtained those data from another work in the literature (Delgado et al., 2006). The necessary equations to calculate the dispersion coefficients, as well as the LDF constant can be found elsewhere in the literature (Delgado et al., 2006; Farooq and Ruthven, 1990) and we present them on Appendix B.

Adsorption of N_2 and CH_4 on silicalite pellets is an equilibrium controlled phenomenon (Delgado et al., 2006, 2011; Sant Anna et al., 2016). To describe the adsorption equilibrium we used Langmuir isotherms. Table III.3 presents the adsorption parameters used throughout this work.

Table III.3 - Langmuir equilibrium isotherm (Delgado et al., 2011)

	b_0 (10^{-9} Pa^{-1})	$-\Delta H$ (kJ mol^{-1})	q_{max} (mol kg^{-1})
N_2	1.57	15.5	1.72
CH_4	1.37	18.5	2.01

Figure II.7 presents profiles of pressure and temperature as well as nitrogen and methane concentrations during 20 PSA cycles. The bed is initially filled with an inert gas at 1 bar, which is expelled from the column during the first adsorption step. As Figure III.3 shows, there is a considerable difference between first and second cycles and the Cyclic Steady State (CSS) occurs after about 5 cycles for mole fractions and about 10 cycles for temperature.
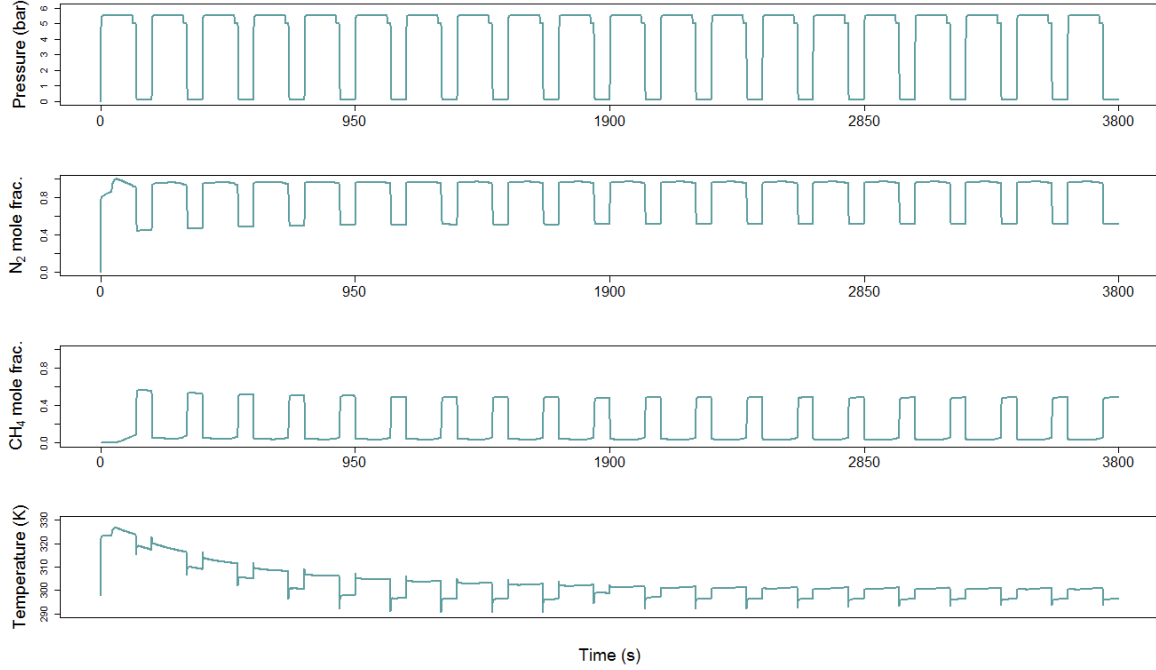
Figure III.3 - Profiles of pressure, CH4 mole fraction, N2 mole fractions and temperature at the outlet of the bed at each step.

There are two separation performance indicators in this case: purity (4) and recovery (5) for the ith component, where i is either N_2 or CH_4. We are interested in evaluating nitrogen purity and recovery at the adsorption phase outlet, as this stream is a final byproduct of the process, which will proceed to atmospheric venting. Methane purity and recovery are of secondary interest, since this stream will be recycled to the cryogenic distillation process.

$$Pur_i = \frac{\int_0^{t_{ad}} n_{i,out} dt}{\int_0^{t_{ad}} (n_{N_2,out} + n_{CH_4,out}) dt} \tag{4}$$

$$Rec_i = \frac{\int_0^{t_{ad}} n_{i,out} dt + \int_0^{t_{dj}} n_{i,out} dt}{\int_0^{t_{co}} n_{i,in} dt + \int_0^{t_{ad}} n_{i,in} dt} \tag{5}$$

Chapter III.3 - Machine learning

In this work, we use triple layered feed-forward artificial neural networks (ANN) as the surrogate model [1]. A neural network is a massively parallel-distributed processor made of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use (Haykin, 2004). The aim is to find a powerful synaptic modification rule that will allow an arbitrarily connected neural network to develop an internal structure that is appropriate for a particular task domain. The task is specified by giving the desired state vector of the output units

for each state vector of input units (Rumelhart et al., 1986). We use a sequential workflow in order to obtain our ANN model for PSA separation of N_2/CH_4 (Figure III.4). We sample the input space using a Latin Hypercube Sampling (LHS) technique and then calculate the corresponding output vectors with the DAE model presented earlier in this work. We determine the proper number of samples for our network by using training-testing error analysis. After that, we make statistical analysis to determine its predictability (repeated learning-testing cross-validation, bias+variance decomposition, out of sample validation and outlier detection). We further compare this model against N-dimensional linear regression to assure that a nonlinear model is necessary.
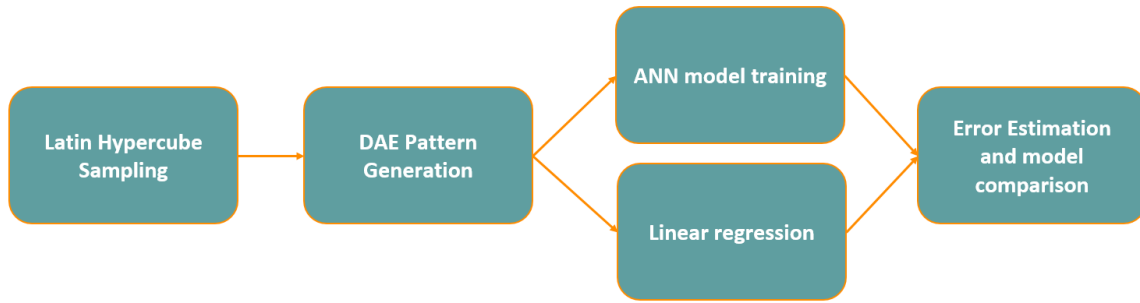
Figure III.4- Machine Learning workflow

Chapter III.3.1 - Input Space Sampling

The state vector of input units consists of twelve process variables $x_j = x_{i,j}, i \in I = \{1,2,...,12\}$, where seven units ($x_{j,1} - x_{j,7}$) represent the operational conditions of the plant (P_h, P_m, $P_{l,}$, t_{ad}, t_{dj}, t_{dr}, t_{co}), three units ($x_{j,8,}$, $x_{j,9}$ and $x_{j,12}$) represent the inlet gas specifications (Q, T and y_{N_2}), and two units ($x_{j,10}$ and $x_{j,11}$) represent the process design specifications (L and D_c). The chosen parameters described above are the necessary constants, i.e. apart from the bed, solid and gases constants, to simulate a PSA process using the DAE model. The aim of this work is to provide a surrogate model exhaustive in its variables. Therefore, we avoid making simplifications in order reduce the input vector size. Some simplifications that are worthy to mention is that we consider the inlet gas pressure equal to the adsorption pressure and the column wall temperature equal to the inlet gas temperature. Moreover, to our understanding, all the remaining variables (e.g. interstitial velocity, cross section area, Péclet number etc.) derive from the twelve previously mentioned variables and we rule them out from our analysis. These considerations by no means eliminate the possibility of any correlation between any subset of variables in the state vector of

inputs. Instead, we assume that they are not relevant, since our interest rests on the predictability of the model rather than extracting information from model parameters.

Table III.4 – Upper and lower bounds of input variables for the sampling plan

	Ph (bar)	Pm (bar)	Pl (bar)	T_{ad} (s)	T_{dj} (s)	T_{dr} (s)	T_{co} (s)	Q (m^3/s)	T (K)	L (m)	Dc (m)	Y_f
Lower	4.00	4.00	0.10	20.00	10.00	10.00	20.00	$1.00 \ 10^{-3}$	273.00	1.00	0.10	0.10
Upper	6.00	6.00	0.20	60.00	50.00	50.00	30.00	$3.00 \ 10^{-3}$	323.00	3.00	0.30	0.90

We chose 1300 samples $x_j, j \in J = \{1,2,\dots,1300\}$ from the input state space X (Table III.4) using Latin Hypercube Sampling (LHS) with genetic algorithm (Mckay et al., 2000; Stocki, 2005). Using this approach, we intend to generate a stratified sampling plan satisfying the Latin Hypercube condition, i.e. the projection of each sample hypercube on each of the twelve axes must not overlap the projection of another hypercube, while maximizing the S optimality condition. S-optimality seeks to maximize the mean distance from each design point to all the other points in the design, so the points are as spread as possible. With the samples from the input state space, we generated 1300 corresponding samples $y_j, j \in J = \{1,2,\dots,1300\}$ of the output state space Y, where each unit in the state vector of outputs represents respectively N_2 purity, N_2 recovery, CH_4 purity and CH_4 recovery. With this sampling plan, we obtained a dataset containing 1300 input-output samples $z_j = (x_j, y_j), j \in J = \{1,2,\dots,1300\}$ of the input-output space $Z = (X, Y)$.

Chapter III.3.2 - Neural networks

In order to capture the non-linearity in the data, we use multilayered feed-forward artificial neural networks. The network contains one input layer, one output layer and one hidden layer as presented in Figure III.5 (Abu-Mostafa et al., 2012; Baughman and Liu, 1995).
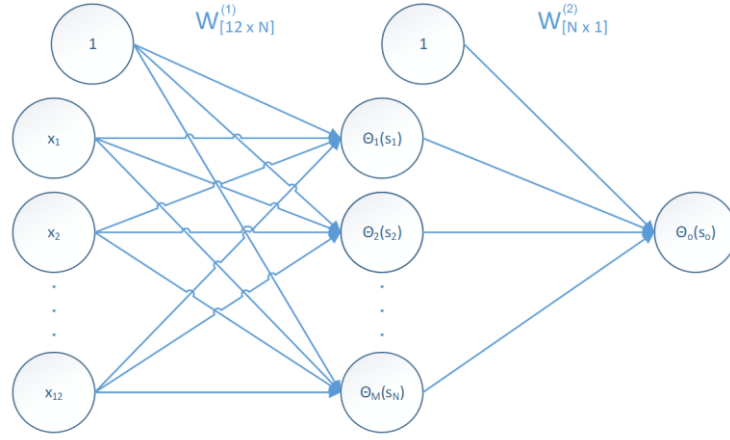
Figure III.5 – Diagram showing the neural network layout.

The input layer represents a receiver for the normalized input variables, as no mathematical operation happens inside such nodes. We normalize each unit in the state vectors of inputs in the dataset between zero and one according to (6).

$$x'_{i,j} = \frac{x_{i,j} - \min_{j \in J} x_{i,j}}{\max_{j \in J} x_{i,j} - \min_{j \in J} x_{i,j}} \tag{6}$$

Each node in the hidden layer performs a sequence of two mathematical operations. The first operation calculates the node activation as a weighted sum of the units in the state vector of inputs (7),

$$s_k(x_j) = \sum_{i=0}^{12} w_{i,k}^{(1)} x_{i,j} \tag{7}$$

where $k \in K = 1, 2, \ldots, N^{hidden}$ is the position of the neuron in the hidden layer. One entry is added to the state vector of inputs corresponding to $x_{0,j} = 1$, which is used to determine the node bias. The second operation transforms the result of the activation into the hidden layer output using a logistic activation function (8).

$$\theta(s_k) = \frac{1}{1 + e^{-s_k}} \tag{8}$$

The neural networks considered contains a single node on the output layer with a logistic activation function. Therefore, we developed one neural network per unit in the state vector of outputs. The neural network model can be compactly written as its hypothesis set (9).

$$h_{ANN}(x_j) = \theta \left(\sum_{k=0}^{N^{hidden}} w_{k,1}^{(2)} \theta(s_k(x_j)) \right) \tag{9}$$

Again, one entry is $\theta\left(s_0(x_j)\right) = 1$ added to the vector arising from the hidden layer to account for the output neuron bias.

Chapter III.3.3 - Training and testing the neural networks

We trained a single different ANN model for each unit in the state vector of outputs (N_2 purity, N_2 recovery, CH_4 purity and CH_4 recovery). The learning algorithm used in this work is the gradient descent (steepest descent) method. The standard backpropagation (backward propagation of errors) algorithm calculates the partial derivatives of the mean squared error between each target value and the network output regarding the weight values (Rumelhart et al., 1986). The gradient descent technique updates the weight values to minimize the mean squared error (MSE) (10). Gradient descent is an optimization procedure that relies on first order derivatives to minimize the MSE. It is possible to use second order derivatives in the optimization procedures, as is the case of the Davidson-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods. However, the hessian updates in such methods cast a heavy computational burden to ANN training (Leonard and Kramer, 1990).

$$E(h_{ANN}) = \frac{1}{N^{sanples}} \sum_{n=1}^{N^{samples}} (h(x_n) - y_n)^2 \tag{10}$$

$$\rho_{y_n,h(x_n)} = \frac{cov\left(y_n, h(x_n)\right)}{\sigma_{y_n}\sigma_{h(x_n)}} \tag{11}$$

In order to evaluate the convergence and generalization of the learning algorithm, we train the network with 80% of the available samples to calculate the approximation error $E_{train}(h(x_n))$ or simply E_{train}. With the remaining 20% of the dataset, we test the neural network by calculating the test error $E_{test}(h(x_n))$ or simply E_{test}. Both error computations use equation (10). In some cases, we also analyze the Pearson correlation coefficient between the predicted and target values (11).

Chapter III.3.4 - Pruning the neural networks

A central problem in data-driven learning is choosing the adequate number of neurons and connections, which directly influence the number of model parameters. A model with fewer parameters than the ideal may not make accurate predictions, while an over-parameterized model

may fit the noise in the dataset. The latter situation is also known as overfitting and its main consequence is the lack of model generalization, i.e. the model makes good predictions inside the training dataset but makes poor predictions outside of it. There are two main approaches to avoid overfitting. The first approach is regularization, which is a set of methods to optimize the MSE imposing constraints to the weight values (Abu-Mostafa et al., 2012). The second is pruning, a technique to remove non-useful connections (weights) or even entire neurons in the network. We use the optimal brain surgeon technique to prune the neural network (Hassibi and Stork, 1993). Figure III.6 shows the algorithm to remove connections on a network. Here, $H \equiv \partial^2 E / \partial W^2$ is the hessian matrix computed in the optimum point of the training algorithm and $L_q = \frac{1}{2} \frac{w_q^2}{[H^{-1}]_{qq}}$ is the error increase by the removal of the weight w_q (saliency). On step 4, one updates the weights using $\delta W = -\frac{w_q}{[H^{-1}]_{qq}} H^{-1} \cdot e_q$ where e_q is the unit vector in the weight space corresponding to the (scalar) weight w_q.

```
1. Train a "reasonably large" network to minimum error.
2. Compute H⁻¹.
3. Find the q that gives the smallest saliency. If this candidate
   error increase is much smaller than E, then the qᵗʰ weight
   should be deleted, and we proceed to step 4; otherwise, go to
   step 5.
4. Use the q from step 3 to update all weights. Go to step 2.
5. No more weights can be deleted without a large increase in E.
   (At this point it may be desirable to retrain the network.)
```

Figure III.6 - Optimal Brain Surgeon pruning algorithm (Hassibi and Stork, 1993)

We used the RSNNS package (Bergmeir and Benítez, 2012) with the R programming language, to train test, prune and analyze different ANN topologies.

Chapter III.4 - Results and discussion

Chapter III.4.1 - Ideal number of training samples

The first machine learning analysis we perform is the determination of the ideal number of samples to generate accurate models to describe the PSA dynamics. In order to accomplish that, we generated 26 random subsets $S^{(l)}$ from the dataset z_j, where $l = 50, 100, \dots, 1300$ is the number of samples inside each subset. We then applied the learning-testing procedure described above over each subsample to generate Figure III.7. It is possible to see a significant error reduction until $S^{(550)}$.

On the $S^{(600)}$ the error slightly increases and from there on it remains relatively stable. Therefore, we use $S^{(650)}$ over the next sections. It is important to disclose that there is no guarantee that this subsample is stratified, i.e. it is possible that some regions are undersampled while other are oversampled. Section Chapter III.4.2.3 discusses the positive implications of this approach.
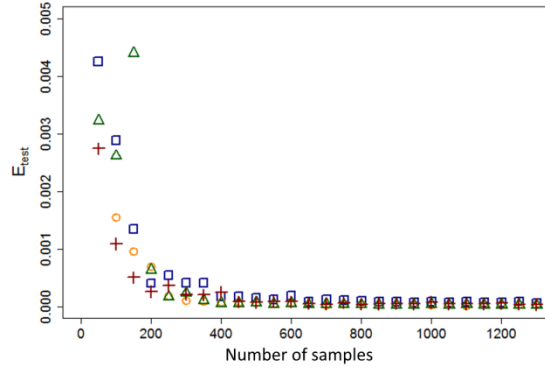

Figure III.7 – Test error for the prediction of N2 purity (squares), N2 recovery (circles), CH4 purity (triangles) and CH4 recovery (crosses) as a function of the number of samples.

We applied the pruning algorithm on all neural networks displayed above, beginning with 50 fully connected neurons in the hidden layer. This procedure not only removed connections (by zeroing the weights) but also removed entire neurons, rendering different incompletely connected ANN for each dataset and each output variable. Table III.5 presents the resulting number of neurons and connections for $S^{(650)}$. The reader can find the complete dataset and the complete weight matrixes in the supplementary material of this article.

Table III.5 – Comparison between initial and pruned neural network topologies

Number of…	N_2 purity	N_2 recovery	CH_4 Purity	CH_4 Recovery
training samples	520	520	520	520
testing samples	130	130	130	130
initial hidden units	50	50	50	50
initial connections	650	650	650	650
initial parameters	701	701	701	701
final hidden units	22	14	16	14
final connections	112	68	61	77
final parameters	135	83	78	92

Chapter III.4.2 - Out of sample error estimation

During ANN training, E_{train} is the approximation error of a given ANN. However, an important type of error is the out of sample error or generalization error E_{out}. This error measures the mean

quadratic difference between the predicted and the target functions over the entire domain. Since its exact calculation requires knowing the behavior of the target function over the whole domain, we here make three analysis to approximate and understand the value of E_{out}; repeated learning-testing cross-validation (LT), bias+variance analysis (BV) and validation with the unused dataset (Val).

Chapter III.4.2.1 - Repeated learning-testing cross-validation

Another work in the literature introduced this method and analyzed its stability (Burman, 1989). We repeatedly split the data randomly into two parts, a learning set of size 520 and a test set of size 130. For each split, we conduct the ANN training with the learning set. After the training, we calculate the MSE over the test set. Since we split the data 10 times over, the repeated learning-testing error E_{LT} is the average over the 10 E_{test}:

$$E_{LT} = \frac{1}{10} \sum_{\alpha=1}^{10} \left\{ \frac{1}{130} \sum_{i=1}^{130} (h(x_i) - y_i)^2 \right\} \tag{12}$$

Since the pruning algorithm is a part of the learning algorithm, the number of hidden units and connections of the ANN models may vary for each data split. This impairs our ability to analyze the statistics for each individual weight. However, we make several different analysis with the ANN predicted values to ensure the consistency and stability of this model.

Chapter III.4.2.2 - Bias+variance decomposition

The bias+variance decomposition of the out of sample error also depends on squared error measures. Let D be the collection of dataset splits $D_i, i = 1,2,...,10$. Let also $g^{(D_i)}(x)$ be the ANN model trained on a specific dataset. The out of sample error estimation can be expressed as:

$$\mathbb{E}_D[E_{out}(g^{(D)})] = \mathbb{E}_D\left[\mathbb{E}_x\left[\left(g^{(D)}(x) - y\right)^2\right]\right] \tag{13}$$

for any x. Now, let $\bar{g}(x) \cong 1/10 \sum_{i=1}^{10} g^{(D_i)}(x)$ approximate the average function over all possible datasets. We can rewrite the right-hand side of (X) as:

$$\mathbb{E}_D[E_{out}(g^{(D)})] = \mathbb{E}_x\left[\mathbb{E}_D\left[\left(g^{(D)}(x) - \bar{g}(x)\right)^2\right] + (\bar{g}(x) - y)^2\right] \tag{14}$$

$$= \mathbb{E}_x\left[\mathbb{E}_D\left[\left(g^{(D)}(x) - \bar{g}(x)\right)^2\right]\right] + \mathbb{E}_x\left[(\bar{g}(x) - y)^2\right] \tag{15}$$

$$= variance + bias \tag{16}$$

Here, the bias term indicates the squared distance between the average function and the target function while the variance term indicates how spread are the other approximation functions from the average function. The reader can find the complete derivation of the bias+variance decomposition elsewhere in the literature (Abu-Mostafa et al., 2012). Notice that the difference between the repeated learning-testing error E_{LT} and the bias+variance error $E_{BV} = \mathbb{E}_D\big[E_{out}\big(g^{(D)}\big)\big]$ is that, while the first one used only the test dataset, the second uses the whole dataset for each estimation.

Chapter III.4.2.3 - Validation error

Another way to estimate E_{out} is by using a validation set, i.e. a dataset unused for training and testing. Since we first generated 1300 samples and only used 650, we calculate E_{val} by taking the MSE over the unused dataset.

Table III.6 - Summary of the out of sample error estimators for ANN regression

	N_2 purity	N_2 recovery	CH_4 Purity	CH_4 Recovery	average
E_{train}	3.56E-05	3.32E-05	2.33E-05	2.97E-05	3.05E-05
E_{test}	1.33E-04	4.10E-05	2.67E-05	4.15E-05	6.05E-05
E_{LT}	1.07E-04	4.67E-05	3.76E-05	7.22E-05	6.59E-05
E_{BV}	6.68E-05	9.08E-05	3.18E-05	2.02E-04	9.79E-05
E_{val}	9.93E-05	4.72E-05	3.99E-05	5.12E-05	5.94E-05
Bias	3.45E-05	3.59E-05	2.21E-05	6.50E-05	3.94E-05
Variance	3.23E-05	5.49E-05	9.72E-06	1.37E-04	5.86E-05

On section 4.1, we mentioned that subsampling a stratified dataset could not necessarily generate a stratified dataset. However, since the validation dataset is the complement of the training-testing dataset, undersampled regions on the former find an equivalent oversampled region in the latter and vice versa. Therefore, it becomes clear that the here presented validation error is a conservative estimation of E_{out}. Table III.6 presents a summary of each E_{out} estimator for each output variable. We can argue that on average $E_{BV} > E_{LT} > E_{test} > E_{val} > E_{train}$ in the present analysis. Therefore, our most pessimistic E_{out} estimator is the bias+variance decomposition. The bias+variance decomposition also gave us evidence that our ANN model is stable because of the relatively low bias and variance across the different dataset splits. We can also discuss, based on all estimators for E_{out} that this value has an expected order of magnitude of -4.

Chapter III.4.3 - Outlier detection

We make use of the bias+variance analysis in order to find outliers. Since $\bar{g}(x)$ is the average function for any x, we assume g^D to be an independent identically distributed random variable where mean and variance are given by:

$$\mathbb{E}_D[g^D(x)] \cong \bar{g}(x) \tag{17}$$

$$s_D^2[g^D(x)] \cong \mathbb{E}_x\left[\mathbb{E}_D\left[\left(g^{(D)}(x) - \bar{g}(x)\right)^2\right]\right] \tag{18}$$

with the randomness arising from the randomness in the training dataset. We then use a student-t test on each z_j, to test the null hypothesis $H_0: \bar{g}(x_i) = y_i$ the alternative hypothesis is $H_0: \bar{g}(x_i) \neq y_i$ with 99% confidence. Therefore, for each z_i we test if

$$\bar{g}(x_i) - t_{\alpha,\nu} * s_D[g^D(x)] < y_i < \bar{g}(x_i) + t_{\alpha,\nu} * s_D[g^D(x)] \tag{19}$$

where $\alpha = 0.005$, $\nu = \infty$ ant $t_{0.005,\infty} = 2.576$. Hence, the points who fail to pass this test are considered outliers. Table III.7 presents the number of outliers in the training-testing and validation datasets. There is an acceptable relative number of outliers for all output variables as they are all less than 10%.

Table III.7 – Outliers in the training-testing, validation and total samples for the ANN regression

	N2 purity	N2 recovery	CH4 Purity	CH4 Recovery
Number of training-testing sample outliers	14	19	47	0
	2.2%	2.9%	7.2%	0.0%
Number of validation sample outliers	45	31	64	1
	6.9%	4.8%	9.8%	0.2%
Total outliers	59	50	111	1
	4.5%	3.8%	8.5%	0.1%

Chapter III.4.4 - Comparison against a linear model

In this section, we compare the ANN surrogate model against a multivariate linear regression in order to analyze if a less complex model can fit the data with the same level of error. Multivariate linear regression is a data-driven learning technique which extends the simple linear regression (regression towards mediocrity) (Galton, 1886; Montgomery and Runger, 2014).

The hypothesis set for this model is:

$$h_{lin}(x_j) = \sum_{i=0}^{12} w_i x_{j,i} = W^T x_j \qquad (20)$$

We train the model using $S^{(650)}$ and its respective training-testing split. We use the method of least squares to obtain the model coefficients, which is an equivalent to performing ANN training. We apply the same treatment to estimate E_{out} as the one applied to ANN on section Chapter III.4.2. Table III.8 presents the out of sample error estimators. It is notable that they are about 100 times larger than the ANN regression errors. This translates into a poor fitting as shown on the outlier analysis results (Table III.9). We expected this result since PSA separation of gases is a highly nonlinear phenomenon. Therefore, a linear model is not suited for this application. Although simpler nonlinear surrogate models may exist, making an exhaustive comparison among them is out of the scope of this work.

Table III.8 – Summary of the out of sample error estimators for ANN regression

	N_2 purity	N_2 recovery	CH_4 Purity	CH_4 Recovery	Average
E_{train}	4.91E-03	1.01E-03	1.47E-03	3.89E-03	2.82E-03
E_{test}	5.33E-03	1.11E-03	1.51E-03	4.02E-03	2.99E-03
E_{LT}	5.26E-03	1.07E-03	1.51E-03	4.07E-03	2.98E-03
E_{BV}	5.00E-03	1.03E-03	1.48E-03	3.92E-03	2.86E-03
E_{val}	5.00E-03	1.27E-01	2.01E-01	3.94E-02	9.30E-02
Bias	4.98E-03	1.03E-03	1.48E-03	3.90E-03	2.85E-03
Variance	2.32E-05	7.70E-06	6.63E-06	1.64E-05	1.35E-05

Table III.9 - Outliers in the training-testing, validation and total samples for the ANN regression

	N2 purity	N2 recovery	CH4 Purity	CH4 Recovery
Number of training-testing sample outliers	579	540	586	565
	89.1%	83.1%	90.2%	86.9%
Number of validation outliers	572	515	567	563
	88.0%	79.2%	87.2%	86.6%
Total outliers	1151	1055	1153	1128
	88.5%	81.2%	88.7%	86.8%

Chapter III.5 - Optimization

The literature presents a diversity of Surrogate Based Optimization (SBO) strategies such as the basic unconstrained SBO, multiple surrogates SBO, approximation model management framework (Queipo et al., 2005) and the trust region framework (Forrester and Keane, 2009; Li et al., 2014). In this work, we make use the surrogate model global accuracy (SMGA) assumption (Forrester and Keane, 2009). The evidence we use for this assumption is the similarity among all error calculation

strategies, the small surrogate model variance in respect to 10 different training datasets, and the relatively low number of outliers.

We then perform the One Shot SBO (Queipo et al., 2005) as described in Figure III.8. The first step of the One Shot SBO consists of the ANN model training and evaluation described in Section Chapter III.4. In order to perform the second step of the One Shot SBO, we use the optimization algorithm TOLMIN which employs a Sequential Quadratic Programming (SQP) method over an arbitrary objective function and is capable of handling linear equality and inequality constraints (Powell, 1989) as well as box constraints. Therefore, the TOLMIN algorithm makes successive calls to the ANN surrogate model (objective function evaluations) to accomplish the minimization. The third step of the One Shot SBO (checking phase) makes a full PSA simulation using the DAE system to compare the target value against the ANN predicted value in the optimal condition.

```
1. Construct a surrogate model
   from a set of known data
   points.

2. Estimate the function
   minimizer using the
   surrogate function.

3. Evaluate the true function
   value at the estimated
   minimum (checking phase).
```

Figure III.8 –One Shot SBO (Queipo et al., 2005)

In addition to the SBO, we also perform a black-box optimization in order to compare the performance of both methods. Across the black-box optimization steps, the TOLMIN algorithm makes successive calls to a black-box function providing, in each call, a proper state vector of inputs and fetching the objective function as well as the gradient vector. In each call to the black-box function a full PSA simulation using the DAE system occurs, as described in Section Chapter III.2. In both optimization schemes, TOLMIN estimates the gradient vector using finite differences.

Chapter III.5.1 - Single Objective Optimization

Due to the nature of the adsorption process, our first analysis consists of optimizing the nitrogen product purity. The standard format of the optimization problem is:

$$\text{Minimize: } -Pur_{N_2} \tag{21.a}$$
$$\text{Subject to: } P_H - P_M > 0 \tag{21.b}$$
$$T = 298.15\ K \tag{21.c}$$
$$L = 2\ m \tag{21.d}$$
$$D_c = 0.2\ m \tag{21.e}$$
$$y_{N_2,f} = 0.85 \tag{21.f}$$
$$lb < x_i < ub \tag{21.g}$$

Where lb and ub are the lower and upper bounds for each input variable (box constraints) in Table III.10. In the proposed problem, both adsorption and co-current desorption pressures can lie in a range between 4 and 6 bar. Therefore, (21.b) keep the first one always larger than the second.

Table III.11 summarizes the results regarding the objective function. In all table columns except "Step 2", each function call corresponds to a full PSA simulation using the DAE system. For "Step 2", each function call corresponds to a feed-forward ANN calculation, which is cheaper in terms of computational costs, and therefore, has a smaller duration. Since the optimization problem has 12 decision variables, each function evaluation takes 1 call to the objective function plus 12 function calls to obtain the gradient vector. Therefore, the number of objective function evaluations is the number of function calls divided by 13. Our first observation is that, in the One Shot SBO, the checking phase optimum was close to the predicted ANN optimum with a difference of only 0.6%. Moreover, the SBO found a higher value of the optimum with less objective function evaluations than the black-box optimization. A possible explanation is the ability of a non-overfitted ANN to generate a smooth hypersurface in the input-output space $Z = (X, Y)$ in the presence of noise inside the dataset (Abu-Mostafa et al., 2012), which facilitates the search for the optimum. One evidence to confirm this statement is that P_m is closer to P_h and P_l is at the lower bound in the arguments of the maxima (Table III.10) for the One Shot SBO. This is an expected observation, since decreasing the purging pressure leads to a cleaner bed for the better operation of the posterior adsorption cycle and a higher co-current blowdown pressure leads to a smaller amount of CH_4 desorbing and going to the nitrogen product stream. The black-box optimization clearly was not

able to reach such frontiers, which may be due to small residue oscillations close to the optimum. However, the relative difference between the black-box and the SBO optima was small (about 1.6%).

Regarding the optimizer performance (Table III.11), simulating a PSA cycle comprising the pressurization, adsorption, co-current blowdown and counter-current blowdown takes on average 2.5 seconds on a personal computer (Intel Core I7-2670QM microprocessor with 2.20 GHz and 8Gb of RAM). Since the CPU time varies for different computer configurations, we use the dimensionless time unit $cycles = \dfrac{solution\ time}{average\ single\ cycle\ CPU\ time}$.

Table III.10 – Lower & upper bounds, initial guess and arguments of the maxima in the black-box optimization and the One Shot SBO.

Variables	Ph (bar)	Pm (bar)	Pl (bar)	T_{ad} (s)	T_{dj} (s)	T_{dr} (s)	T_{co} (s)	Q (m³/s)
Lower	4.00	4.00	0.10	20.0	10.0	10.0	20.0	0.0010
Upper	6.00	6.00	0.20	60.0	50.0	50.0	30.0	0.0030
Guess	5.20	5.00	0.15	40.0	30.0	30.0	25.0	0.0020
black-box optim. argmax	5.42	4.31	0.12	20.0	45.1	34.3	30.0	0.0017
One Shot SBO argmax	5.58	5.51	0.10	20.0	50.0	50.0	30.0	0.0010

The black-box optimization took 22620 cycles (15.7 hours) while the Step 2 of the One Shot SBO (the core of the optimization procedure) took 0.02 cycles (about 50 milliseconds). However, it took 26000 cycles (18 hours) to generate the 1300 samples (Step 1), and the ANN checking phase (Step 3) takes another 20 cycles (50 seconds). Therefore, the SBO lasts 15% longer than the complete One Shot SBO formulation. However, once Step 1 is done, it is possible to perform numerous SBO procedures at the time cost of 20 cycles per optimization, or even 0.02 cycles if one trusts the model enough to dismiss the checking phase. This approach also opens possibilities to implement real-time simulation and optimization of processes like ANN nonlinear model predictive control (NMPC) for the PSA process.

Table III.11 – Optimization results and performance. The function calls takes into account both the objective function evaluation and gradient vector generation.

	black-box optim.	One Shot SBO		
		Step 1	Step 2	Step 3
Optimum N_2 purity (mole/mole)	0.9768	-	0.9891	0.9951
Number of function calls	1131	1300	403	1
Duration (cycles)	22620	26000	0.02	20

Chapter III.5.2 - Multi-objective optimization

Maximizing N_2 purity leads to a very poor nitrogen recovery. Conversely, maximizing N_2 recovery leads to a very poor N_2 purity. We then perform a bi-objective optimization in order to understand the tradeoff between both variables. The standard form of the optimization problem is now:

$$\text{Minimize: } -\left(\alpha \, Pur_{N_2} + (1-\alpha)\, Rec_{N_2}\right) \quad \alpha \in [0,1] \tag{22.a}$$
$$\text{Subject to: } P_H - P_M > 0 \tag{22.b}$$
$$T = 298.15 \; K \tag{22.c}$$
$$L = 2 \; m \tag{22.d}$$
$$D_c = 0.2 \; m \tag{22.e}$$
$$y_{N_2,f} = 0.85 \tag{22.f}$$
$$lb < X'_i < ub \tag{22.g}$$

We conduct the multi-objective optimization using the One Shot SBO described in the previous section. Figure III.9 shows the optima estimated by the ANN surrogate models (Step 2) as well as their respective checking phase values (Step 3). We can state that the ANN slightly underestimates the nitrogen purity and slightly overestimates nitrogen recovery. Over the 100 evaluated points composing the Pareto front, we observed a maximum relative difference between the estimated (Step 2) and target (Step 3) value of 1.4% for N_2 purity and 4% for N_2 recovery. This indicates that the One Shot SBO is also a proper technique for multi-objective optimization. Table III.12 shows five selected optimal points of the multi-objective optimization. Therefore, we argue that all of the 8 independent variables have an influence on the search for the optima.

Table III.12 – Five selected optimal points of the multi-objective optimization

α	Ph	Pm	Pl	T_{ad}	T_{dj}	T_{dr}	T_{co}	Q	N_2 pur.	N_2 rec.	CH_4 pur.	CH_4 rec.
0	6.00	4.00	0.15	60.00	10.00	29.94	25.00	0.0030	0.9078	0.7570	0.3256	0.6914
0.01	6.00	4.00	0.10	60.00	10.00	48.25	20.00	0.0030	0.9181	0.7570	0.3299	0.7136
0.82	6.00	4.00	0.10	46.82	17.59	43.56	20.00	0.0022	0.9515	0.6695	0.2934	0.8339
0.99	5.80	5.73	0.10	20.00	50.00	50.00	30.00	0.0010	0.9891	0.2148	0.1809	0.9865
1.00	5.58	5.50	0.10	20.00	50.00	50.00	30.00	0.0010	0.9892	0.2111	0.1800	0.9867
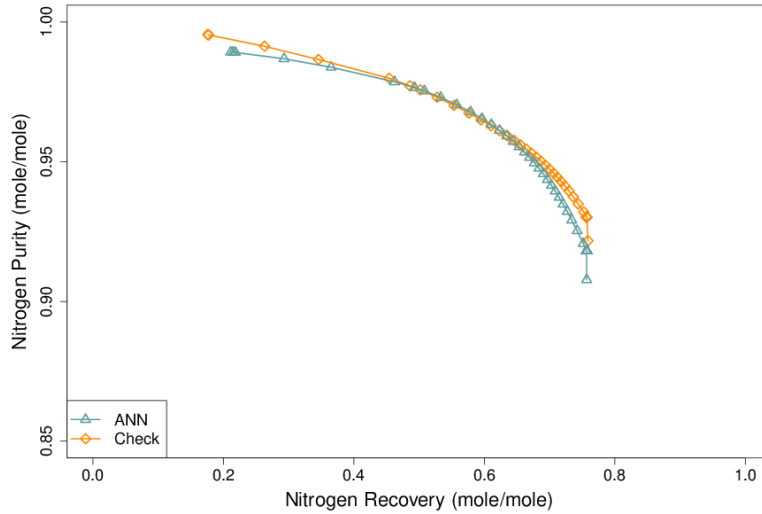
Figure III.9 – Tradeoff curve between nitrogen purity and nitrogen recovery maximization. The values predicted by the ANN models are in blue and the DAE check of the objective function is in orange.

Figure III.10 shows how each of the fixed variables (temperature, length diameter and inlet mole fraction) influence the optimization curve. Temperature has the smallest influence in the maximum nitrogen purity and maximum nitrogen recovery while inlet N_2 mole fraction has the greatest influence on the overall shape and position of the Pareto front. In all four cases, although the process can reach near 100% N_2 purity, there seems to exist an upper limit to N_2 recovery of around 90%.

Regarding performance, we made 100 One Shot SBO procedures to generate Figure III.9, which took 1.4 computational hours. Generating Figure III.10, took 400 One Shot SBO procedures, but this time we disregarded the checking phase (Step 3), taking about 20 computational seconds to accomplish. This is a proper approach, since we are interested in describing the qualitative behavior of the four values fixed by optimization constraints.
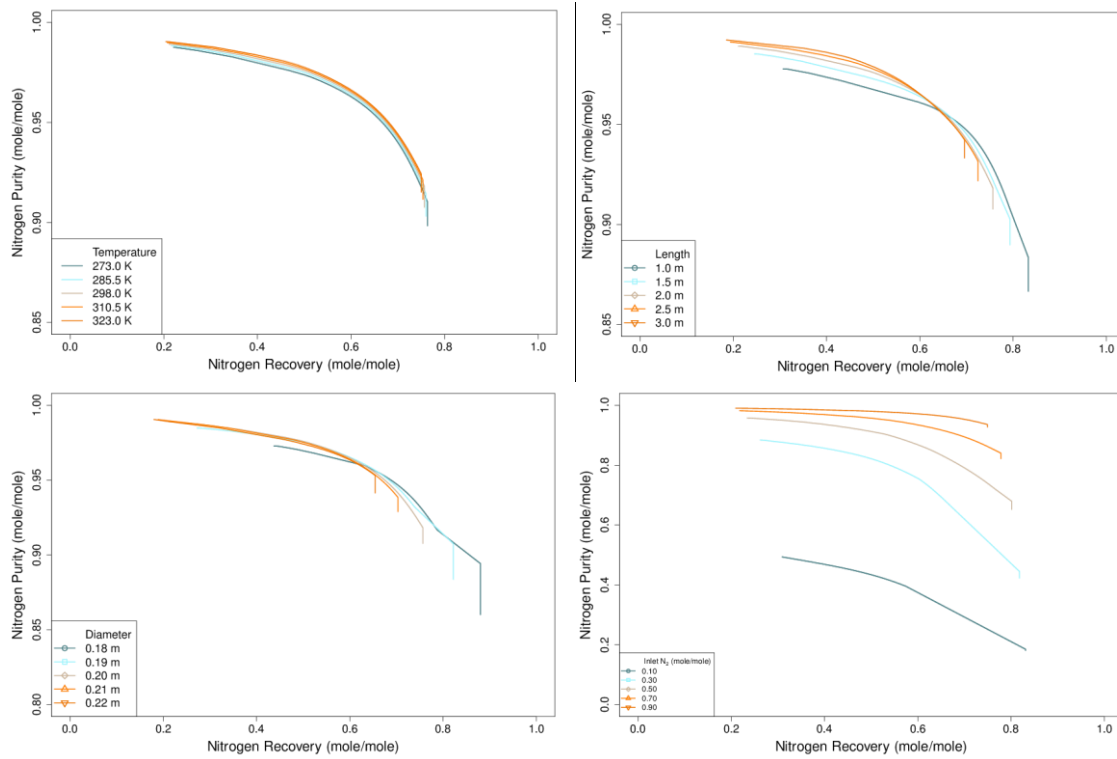
Figure III.10 – Influence of each fixed variable in the multi-objective optimization

Chapter III.6- Conclusion

This work demonstrates the advantage of using artificial neural networks surrogate models to analyze the PSA separation of N_2/CH_4 mixtures using silicalite. Neural networks presented an out of sample error estimation about 100 times lower than linear regression. Moreover, pruning the ANN decreased the model complexity, avoiding overfitting problems. We demonstrated that optimizing the differential algebraic system of equations takes up to 15.7 hours on a desktop PC (Intel Core I7-2670QM microprocessor with 2.20 GHz and 8Gb of RAM). Black-box optimization with neural network models resulted in an optimization time of 50 seconds considering the checking phase (One Shot SBO Step 3) and 50 milliseconds disregarding it. This method was also able to find a higher confirmed nitrogen purity than black-box optimization on single objective optimization, around 99.5%. Multi-objective optimization showed a performance cap to nitrogen recovery at about 90%. Among the fixed design variables, temperature was the only variable to show small influence on the overall process performance. The approach described in this work can be both a process simulator and a process optimization scheme. Although we chose to fix four decision variables in this problem, one can fix any desired amount of process variables in the optimization procedure, depending on the intended analysis.

References

Abu-Mostafa, Y.S., Magdon-Ismail, M., Lin, H.-T., 2012. Learning from data. AMLBook New York, NY, USA:

Agarwal, A., Biegler, L.T., Zitney, S.E., 2009. Simulation and Optimization of Pressure Swing Adsorption Systems Using Reduced-Order Modeling. Ind. Eng. Chem. Res. 48, 2327–2343. doi:10.1021/ie071416p

Baughman, D.R., Liu, Y.A., 1995. Neural networks in bioprocessing and chemical engineering. Academic Press, San Diego.

Beck, J., Friedrich, D., Brandani, S., Fraga, E.S., 2015. Multi-objective optimisation using surrogate models for the design of VPSA systems. Comput. Chem. Eng. 82, 318–329. doi:10.1016/j.compchemeng.2015.07.009

Beck, J., Friedrich, D., Brandani, S., Guillas, S., Fraga, E.S., 2012. Surrogate based optimisation for design of pressure swing adsorption systems, in: Proceedings of the 22nd European Symposium on Computer Aided Process Engineering. pp. 1217–1221.

Bergmeir, C.N., Benítez, J.M., 2012. Neural networks in R using the Stuttgart neural network simulator: RSNNS.

Biegler, L.T., Jiang, L., Fox, V.G., 2005. Recent Advances in Simulation and Optimal Design of Pressure Swing Adsorption Systems. Sep. Purif. Rev. 33, 1–39. doi:10.1081/SPM-120039562

Boukouvala, F., Hasan, M.M.F., Floudas, C.A., 2017. Global optimization of general constrained grey-box models: new method and its application to constrained PDEs for pressure swing adsorption. J. Glob. Optim. 67, 3–42. doi:10.1007/s10898-015-0376-2

Brenan, K.E., Campbell, S.L., Petzold, L.R., 1996. Numerical solution of initial-value problems in differential-algebraic equations, Classics in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia.

Burman, P., 1989. A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. Biometrika 503–514.

Capitanescu, F., Ahmadi, A., Benetto, E., Marvuglia, A., Tiruta-Barna, L., 2015. Some efficient approaches for multi-objective constrained optimization of computationally expensive black-box model problems. Comput. Chem. Eng. 82, 228–239. doi:10.1016/j.compchemeng.2015.07.013

Delgado, J.A., Uguina, M.A., Sotelo, J.L., Águeda, V.I., Gómez, P., 2011. Numerical simulation of a three-bed PSA cycle for the methane/nitrogen separation with silicalite. Sep. Purif. Technol. 77, 7–17. doi:10.1016/j.seppur.2010.11.004

Delgado, J.A., Uguina, M.A., Sotelo, J.L., Ruíz, B., 2006. Modelling of the fixed-bed adsorption of methane/nitrogen mixtures on silicalite pellets. Sep. Purif. Technol. 50, 192–203. doi:10.1016/j.seppur.2005.11.026

Eason, J., Cremaschi, S., 2014. Adaptive sequential sampling for surrogate model generation with artificial neural networks. Comput. Chem. Eng. 68, 220–232. doi:10.1016/j.compchemeng.2014.05.021

Fahmi, I., Cremaschi, S., 2012. Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. Comput. Chem. Eng. 46, 105–123. doi:10.1016/j.compchemeng.2012.06.006

Farooq, S., Ruthven, D.M., 1990. Heat effects in adsorption column dynamics. II: Experimental validation of the one-dimensional model. Ind. Eng. Chem. Res. 29, 1084–1090.

First, E.L., Hasan, M.M.F., Floudas, C.A., 2014. Discovery of novel zeolites for natural gas purification through combined material screening and process optimization. AIChE J. 60, 1767–1785. doi:10.1002/aic.14441

Forrester, A.I.J., Keane, A.J., 2009. Recent advances in surrogate-based optimization. Prog. Aerosp. Sci. 45, 50–79. doi:10.1016/j.paerosci.2008.11.001

Forrester, A.I.J., Sóbester, A., Keane, A.J., 2008. Engineering design via surrogate modelling: a practical guide. J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ.

Galton, F., 1886. Regression towards mediocrity in hereditary stature. J. Anthropol. Inst. G. B. Irel. 15, 246–263.

Garud, S.S., Karimi, I.A., Kraft, M., 2017. Smart Sampling Algorithm for Surrogate Model Development. Comput. Chem. Eng. 96, 103–114. doi:10.1016/j.compchemeng.2016.10.006

Graciano, J.E.A., Le Roux, G.A.C., 2013. Improvements in surrogate models for process synthesis. Application to water network system design. Comput. Chem. Eng. 59, 197–210. doi:10.1016/j.compchemeng.2013.05.024

Haghpanah, R., Majumder, A., Nilam, R., Rajendran, A., Farooq, S., Karimi, I.A., Amanullah, M., 2013. Multiobjective Optimization of a Four-Step Adsorption Process for Postcombustion CO_2 Capture Via Finite Volume Simulation. Ind. Eng. Chem. Res. 52, 4249–4265. doi:10.1021/ie302658y

Hasan, M.M.F., Baliban, R.C., Elia, J.A., Floudas, C.A., 2012. Modeling, Simulation, and Optimization of Postcombustion CO_2 Capture for Variable Feed Concentration and Flow Rate. 2. Pressure Swing Adsorption and Vacuum Swing Adsorption Processes. Ind. Eng. Chem. Res. 51, 15665–15682. doi:10.1021/ie301572n

Hassibi, B., Stork, D.G., 1993. Second order derivatives for network pruning: Optimal brain surgeon. Adv. Neural Inf. Process. Syst. 164–164.

Haykin, S., 2004. A comprehensive foundation. Neural Netw. 2, 41.

Jackson, S.R., Finn, A.J., Tomlinson, T.R., 2005. New challenges for UK natural gas. Hydrocarb. Eng. 10, 27–30.

Jiang, L., Biegler, L.T., Fox, V.G., 2005. Design and optimization of pressure swing adsorption systems with parallel implementation. Comput. Chem. Eng. 29, 393–399. doi:10.1016/j.compchemeng.2004.08.014

Jiang, L., Biegler, L.T., Fox, V.G., 2003. Simulation and optimization of pressure-swing adsorption systems for air separation. AIChE J. 49, 1140–1157.

Kidnay, A.J., Parrish, W.R., 2006. Fundamentals of natural gas processing, Mechanical engineering. CRC Press, Boca Raton.

Ko, D., Siriwardane, R., Biegler, L.T., 2005. Optimization of Pressure Swing Adsorption and Fractionated Vacuum Pressure Swing Adsorption Processes for CO_2 Capture. Ind. Eng. Chem. Res. 44, 8084–8094. doi:10.1021/ie050012z

Krooss, B.M., Littke, R., Müller, B., Frielingsdorf, J., Schwochau, K., Idiz, E.F., 1995. Generation of nitrogen and methane from sedimentary organic matter: implications on the dynamics of natural gas accumulations. Chem. Geol. 126, 291–318.

Kuo, J.C., Wang, K.H., Chen, C., 2012. Pros and cons of different Nitrogen Removal Unit (NRU) technology. J. Nat. Gas Sci. Eng. 7, 52–59. doi:10.1016/j.jngse.2012.02.004

Leonard, J., Kramer, M.A., 1990. Improvement of the backpropagation algorithm for training neural networks. Comput. Chem. Eng. 14, 337–341. doi:10.1016/0098-1354(90)87070-6

LeVeque, R.J., 2002. Finite volume methods for hyperbolic problems. Cambridge University Press, Cambridge; New York.

Lewandowski, J., Lemcoff, N.O., Palosaari, S., 1998. Use of neural networks in the simulation and optimization of pressure swing adsorption processes. Chem. Eng. Technol. 21, 593–597.

Li, S., Feng, L., Benner, P., Seidel-Morgenstern, A., 2014. Using surrogate models for efficient optimization of simulated moving bed chromatography. Comput. Chem. Eng. 67, 121–132. doi:10.1016/j.compchemeng.2014.03.024

Liu, X.-D., Osher, S., Chan, T., 1994. Weighted essentially non-oscillatory schemes. J. Comput. Phys. 115, 200–212.

Liu, Y., Delgado, J., Ritter, J.A., 1998. Comparison of finite difference techniques for simulating pressure swing adsorption. Adsorption 4, 337–344.

Lokhandwala, K.A., Pinnau, I., He, Z., Amo, K.D., DaCosta, A.R., Wijmans, J.G., Baker, R.W., 2010. Membrane separation of nitrogen from natural gas: A case study from membrane synthesis to commercial deployment. J. Membr. Sci. 346, 270–279. doi:10.1016/j.memsci.2009.09.046

MacKenzie, D., Cheta, I., Burns, D., 2002. Removing nitrogen. Hydrocarb. Eng. 7, 57–63.

Madeira, A.C. da F., 2008. AVALIAÇÃO DA TECNOLOGIA DE ADSORÇÃO "PSA" PARA REMOÇÃO DE NITROGÊNIO DO GÁS NATURAL.

Mckay, M.D., Beckman, R.J., Conover, W.J., 2000. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. Technometrics 42, 55–61. doi:10.1080/00401706.2000.10485979

Montgomery, D.C., Runger, G.C., 2014. Applied statistics and probability for engineers, Sixth edition. ed. John Wiley and Sons, Inc, Hoboken, NJ.

Nguyen, A.-T., Reiter, S., Rigo, P., 2014. A review on simulation-based optimization methods applied to building performance analysis. Appl. Energy 113, 1043–1058. doi:10.1016/j.apenergy.2013.08.061

Petzold, L.R., 1982. A description of DASSL: A differential/algebraic system solver. Sci. Comput. 1, 65–68.

Powell, M., 1989. A tolerant algorithm for linearly constrained optimization calculations. Math. Program. 45, 547–566.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Kevin Tucker, P., 2005. Surrogate-based analysis and optimization. Prog. Aerosp. Sci. 41, 1–28. doi:10.1016/j.paerosci.2005.02.001

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature 323, 533–538.

Ruthven, D.M., Farooq, S., Knaebel, K.S., 1994. Pressure swing adsorption. VCH Publishers, New York, N.Y.

Sant Anna, H.R., Barreto, A.G., Tavares, F.W., do Nascimento, J.F., 2016. Methane/nitrogen separation through pressure swing adsorption process from nitrogen-rich streams. Chem. Eng. Process. Process Intensif. 103, 70–79. doi:10.1016/j.cep.2015.11.002

Schwaab, M., Pinto, J., 2007. Análise de dados experimentais I.

Stocki, R., 2005. A method to improve design reliability using optimal Latin hypercube sampling. Comput. Assist. Mech. Eng. Sci. 12, 393.

Streich, M., 1970. N2 removal from natural gas. Hydrocarb. Process. 49, 86.

Sundaram, N., 1999. Training Neural Networks for Pressure Swing Adsorption Processes. Ind. Eng. Chem. Res. 38, 4449–4457. doi:10.1021/ie9901731

Tagliabue, M., Farrusseng, D., Valencia, S., Aguado, S., Ravon, U., Rizzo, C., Corma, A., Mirodatos, C., 2009. Natural gas treating by selective adsorption: Material science and chemical engineering interplay. Chem. Eng. J. 155, 553–566. doi:10.1016/j.cej.2009.09.010

Webley, P.A., He, J., 2000. Fast solution-adaptive finite volume method for PSA/VSA cycle simulation; 1 single step simulation. Comput. Chem. Eng. 23, 1701–1712.

Chapter III.7 - Appendix A: Complete list of equations

Component mass balance (A1), total mass balance (A2), ideal gas law (A3), energy balance (A4), LDF mass transfer (A5), Ergum equation (A6), Langmuir equation (A7) and Langmuir isotherm temperature dependence (A8).

$$-\frac{\partial}{\partial Z}\left(CD_{ax}\frac{\partial y_i}{\partial Z}\right)+\frac{\partial(vCy_i)}{\partial z}+\frac{\partial Cy_i}{\partial t}+\frac{(1-\varepsilon)}{\varepsilon}\frac{\partial \overline{q}_i}{\partial t}=0 \quad i=1,\dots,n \tag{A1}$$

$$\frac{\partial(vC)}{\partial z}+\frac{\partial C}{\partial t}+\frac{1-\varepsilon}{\varepsilon}\sum_{i=1}^{n}\frac{\partial \overline{q}_i}{\partial t}=0 \tag{A2}$$

$$C=\frac{P}{RT} \tag{A3}$$

$$-\frac{K_{ax}}{\varepsilon}\frac{\partial^2 T}{\partial Z^2}+\frac{C_{p,g}}{R}\frac{\partial(vP)}{\partial Z}-\frac{C_{p,g}}{R}\frac{\partial P}{\partial t}+\left[\frac{1-\varepsilon}{\varepsilon}\left(\rho C_{p,s}+C_{p,a}\sum_{i=1}^{n}\overline{q}_i\right)\right]\frac{\partial T}{\partial t}$$

$$+\frac{1-\varepsilon}{\varepsilon}C_{p,a}T\sum_{i=1}^{n}\frac{\partial \overline{q}_i}{\partial t}+\frac{1-\varepsilon}{\varepsilon}\sum_{i=1}^{n}\left((-\Delta H_i)\frac{\partial \overline{q}_i}{\partial t}\right)-\frac{2h_{in}}{\varepsilon r_{in}}(T-T_w)=0 \tag{A4}$$

$$\frac{\partial \overline{q}_i}{\partial t}=k_i(q_i^*-\overline{q}_i) \tag{A5}$$

$$-\frac{\partial P}{\partial Z}=\frac{150}{4}\frac{1}{r_p^2}\left(\frac{1-\varepsilon}{\varepsilon}\right)\mu\,v \tag{A6}$$

$$q_i=\frac{Py_i\varepsilon_p}{RT}+\frac{q_{max,i}b_iPy_i}{1+\sum_{j=1}^{j=n_{comp}}b_jPy_j} \tag{A7}$$

$$b_i=b_{i,0}\exp\left(-\frac{\Delta H_i}{RT}\right) \tag{A8}$$

Boundary conditions.

Adsorption

$$D_{ax} \frac{\partial y_i}{\partial Z}\Big|_{Z=0} = -v|_{z=0}(y_{i,feed} - y_i|_{z=0})$$

(A9)

$$\frac{\partial y_i}{\partial Z}\Big|_{Z=L} = 0$$

(A10)

$$K_{ax} \frac{\partial T}{\partial Z}\Big|_{Z=0} = -\varepsilon v|_{z=0}\rho_g C_{p,g}(T_{feed} - T|_{z=0})$$

(A11)

$$\frac{\partial T}{\partial Z}\Big|_{Z=L} = 0$$

(A12)

$$P|_{Z=L} = P_h$$

(A13)

$$v|_{z=0} = v_{feed}$$

(A14)

Blowdown:

$$\frac{\partial y_i}{\partial Z}\Big|_{Z=0} = 0$$

(A15)

$$\frac{\partial y_i}{\partial Z}\Big|_{Z=L} = 0$$

(A16)

$$\frac{\partial T}{\partial z}\Big|_{Z=0} = 0$$

(A17)

$$\frac{\partial T}{\partial Z}\Big|_{Z=L} = 0$$

(A18)

$$\frac{\partial P}{\partial Z}\Big|_{Z=0} = 0$$

(A19)

$$v|_{Z=0} = 0 \tag{A20}$$

Pressurization:

$$D_{ax} \left.\frac{\partial y_i}{\partial Z}\right|_{Z=0} = -v|_{z=0}(y_{i,feed} - y_i|_{z=0}) \tag{A21}$$

$$\left.\frac{\partial y_i}{\partial Z}\right|_{Z=L} = 0 \tag{A22}$$

$$K_{ax} \left.\frac{\partial T}{\partial Z}\right|_{Z=0} = -\varepsilon v|_{z=0}\rho_g C_{p,g}(T_{feed} - T|_{z=0}) \tag{A23}$$

$$\left.\frac{\partial T}{\partial Z}\right|_{Z=L} = 0 \tag{A24}$$

$$v|_{z=0} = f(P|_{z=0}) \tag{A25}$$

$$v|_{Z=L} = 0 \tag{A26}$$

Dimensionless equations:

$$-\frac{1}{Pe}\frac{\partial}{\partial z}\left(\frac{\bar{P}}{\bar{T}}\frac{\partial y_i}{\partial z}\right) + \frac{\partial}{\partial z}\left(\frac{\bar{P}}{\bar{T}}y_i\bar{v}\right) + \frac{\bar{P}}{\bar{T}}\frac{\partial y_i}{\partial \tau} + \frac{1}{\bar{T}}y_i\frac{\partial \bar{P}}{\partial \tau} - \frac{\bar{P}}{\bar{T}^2}\frac{\partial \bar{T}}{\partial \tau} + \psi\frac{\partial x_i}{\partial \tau} = 0 \tag{A27}$$

$$\frac{\partial}{\partial z}\left(\frac{\bar{P}}{\bar{T}}\bar{v}\right) + \psi\sum_{i=1}^{n_{comp}}\frac{\partial x_i}{\partial \tau} - \frac{\bar{P}}{\bar{T}^2}\frac{\partial \bar{T}}{\partial \tau} + \frac{1}{\bar{T}}\frac{\partial \bar{P}}{\partial \tau} = 0 \tag{A28}$$

$$-\Omega_1\frac{\partial^2 \bar{T}}{\partial z^2} + \Omega_2\frac{\partial}{\partial z}(\bar{z}\,\bar{P}) + \Omega_3\,\bar{T}\sum_{i=1}^{n_{comp}}\frac{\partial x_i}{\partial \tau} - \sum_{i=1}^{n_{comp}}\left(\sigma_i\frac{\partial x_i}{\partial \tau}\right) + \Omega_4(\bar{T} - \bar{T}_w) + \Omega_2\frac{\partial \bar{P}}{\partial \tau}$$
$$+ \frac{\partial \bar{T}}{\partial \tau} = 0 \tag{A29}$$

$$\frac{\partial x_i}{\partial \tau} = \alpha_i(x_i^* - x_i) \tag{A30}$$

$$-\frac{\partial \overline{P}}{\partial z} = \frac{150}{4}\frac{1}{r_p^2}\left(\frac{1-\varepsilon}{\varepsilon}\right)^2\frac{v_0 L}{P_0}\mu\overline{v} \tag{A31}$$

Dimensionless variables:

$$\overline{P} = \frac{P}{P_0} \qquad \overline{T} = \frac{T}{T_0} \qquad \overline{T}_w = \frac{T_w}{T_0} \qquad x_i = \frac{q_i}{q_{s,0}} \qquad \overline{v} = \frac{v}{v_0} \qquad z = \frac{Z}{L} \qquad \tau = \frac{tv_0}{L} \qquad \alpha_i = \frac{k_i L}{v_0}$$

Dimensionless groups

$$Pe = \frac{v_0 L}{D_{ax}} \tag{A32}$$

$$\psi = \frac{RT_0 q_{s,0}}{P_h}\frac{(1-\varepsilon)}{\varepsilon} \tag{A33}$$

$$\Omega_1 = \frac{\dfrac{K_{ax}}{v_0 \varepsilon L}}{\dfrac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps} + q_{s,0} C_{pa}\sum_{i=0}^{n_{comp}} x_i\right)} \tag{A34}$$

$$\Omega_2 = \frac{\dfrac{C_{pg}}{R}\dfrac{P_0}{T_0}}{\dfrac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps} + q_{s,0} C_{pa}\sum_{i=0}^{n_{comp}} x_i\right)} \tag{A35}$$

$$\Omega_3 = \frac{C_{pa} q_{s,0}}{\left(\rho_s C_{ps} + q_{s,0} C_{pa}\sum_{i=0}^{n_{comp}} x_i\right)} \tag{A36}$$

$$\Omega_4 = \frac{\dfrac{2h_{in}}{r_{in}}\dfrac{L}{v_0}}{\dfrac{(1-\varepsilon)}{\varepsilon}\left(\rho_s C_{ps} + q_{s,0} C_{pa}\sum_{i=0}^{n_{comp}} x_i\right)} \tag{A37}$$

$$\sigma_i = \frac{\dfrac{q_{s,0}}{T_0}(-\Delta H_i)}{\left(\rho_s C_{ps} + q_{s,0} C_{pa}\sum_{i=0}^{n_{comp}} x_i\right)} \tag{A38}$$

Finite Volumes discretization (Haghpanah et al., 2013)

Notice that the j subscript considers values evaluated inside the finite volume while the $j + 0.5$ subscript considers values evaluated in the interface.

$$\frac{1}{Pe}\frac{1}{\Delta z}\left(\frac{\bar{P}}{\bar{T}}\Big|_{j+0,5}\frac{y_{i,j+1}-y_{i,j}}{\Delta z}-\frac{\bar{P}}{\bar{T}}\Big|_{j-0,5}\frac{y_{i,j}-y_{i,j-1}}{\Delta z}\right)-\frac{1}{\Delta z}\left(\frac{y_i\bar{P}}{\bar{T}}\bar{v}\Big|_{j+0,5}-\frac{y_i\bar{P}}{\bar{T}}\bar{v}\Big|_{j-0,5}\right)$$

$$-\psi\frac{\partial x_{i,j}}{\partial\tau}-\frac{y_{i,j}}{\bar{T}_j}\frac{\partial\bar{P}_j}{\partial\tau}+\frac{\bar{P}_jy_{i,j}}{\bar{T}_j^{\,2}}\frac{\partial\bar{T}_j}{\partial\tau}-\frac{\bar{P}_j}{\bar{T}_j}\frac{\partial y_{i,j}}{\partial\tau}=0 \tag{A39}$$

$$-\frac{1}{\Delta z}\left(\frac{\bar{P}}{\bar{T}}\bar{v}\Big|_{j+0,5}-\frac{\bar{P}}{\bar{T}}\bar{v}\Big|_{j-0,5}\right)-\psi\sum_{i=1}^{n_{comp}}\left(\frac{\partial x_{i,j}}{\partial\tau}\right)+\frac{\bar{P}_j}{\bar{T}_j^{\,2}}\frac{\partial\bar{T}_j}{\partial\tau}-\frac{1}{\bar{T}_j}\frac{\partial\bar{P}_j}{\partial\tau}=0 \tag{A40}$$

$$\Omega_{1,j}\frac{1}{\Delta z}\left(\frac{\bar{T}_{j+1}-\bar{T}_j}{\Delta z}-\frac{\bar{T}_j-\bar{T}_{j-1}}{\Delta z}\right)-\Omega_{2,j}\frac{1}{\Delta z}\left(\bar{v}\bar{P}|_{j+0,5}-\bar{v}\bar{P}|_{j-0,5}\right)-\Omega_{3,j}\bar{T}_j\sum_{i=1}^{n_{comp}}\frac{\partial x_{i,j}}{\partial\tau}$$

$$+\sum_{i=1}^{n_{comp}}\left(\sigma_{i,j}\frac{\partial x_{i,j}}{\partial\tau}\right)-\Omega_{4,j}(\bar{T}_j-\bar{T}_{W,j})-\Omega_{2,j}\frac{\partial\bar{P}_j}{\partial\tau}-\frac{\partial\bar{T}_j}{\partial\tau}=0 \tag{A41}$$

$$\frac{\partial x_{i,j}}{\partial t}=k_i(x_{i,j}^*-x_{i,j}) \tag{A42}$$

$$\bar{v}_{j+0,5}=-\frac{1}{\Delta z}\frac{4}{150}\left(\frac{\varepsilon}{1-\varepsilon}\right)^2 r_p^2\frac{P_0}{\mu v_0 L}(\bar{P}_{j+1}-\bar{P}_j) \tag{A43}$$

WENO FVM (Haghpanah et al. 2013)

$$f_{j+0,5}=\frac{\alpha_{0,j}}{\alpha_{0,j}+\alpha_{1,j}}\left[\frac{1}{2}(f_j+f_{j+1})\right]+\frac{\alpha_{1,j}}{\alpha_{0,j}+\alpha_{1,j}}\left[\frac{3}{2}f_j-\frac{1}{2}f_{j-1}\right] \tag{A44}$$

$$\alpha_{0,j}=\frac{\frac{2}{3}}{\left(f_{j+1}-f_j+\delta\right)^4}\qquad\alpha_{1,j}=\frac{\frac{1}{3}}{\left(f_j-f_{j-1}+\delta\right)^4} \tag{A45}$$

Chapter III.8 - Appendix B: Phenomenological coefficient calculations

LDF kinetic constant

$$\frac{1}{k_i}=\frac{r_p q_{i,f}}{3k_F P y_{i,f}}+\frac{r_p^2\rho_p q_{i,f}\tau'}{8\varepsilon_p D_{N_2,CH_4}P y_{i,f}}+\frac{r_c^2}{15D_c} \tag{B1}$$

$$k_F=\frac{v}{Sc^{\frac{2}{3}}}\left(\frac{0.765}{(Re\varepsilon)^{0.82}}+\frac{0.365}{(Re\varepsilon)^{0.386}}\right) \tag{B2}$$

Axial dispersion

$$D_{ax} = \frac{D_{N_2,CH_4}}{\varepsilon}\,(0.26 + 0.5ScRe) \qquad\text{(B3)}$$

Thermal axial dispersion

$$K_{ax} = k_g(10 + 0.5RePr) \qquad\text{(B4)}$$

Dimensionless groups

$$Re = \frac{Dv\rho}{\mu} \qquad Sc = \frac{\mu}{\rho D} \qquad Pr = \frac{C_p\mu}{k_g}$$

Chapter III.9 Addendum

[1]   For the purposes of training, pruning and error analysis, we used the R package with the RSNNS neural networks library. However, for optimization purposes, we implemented a feed forward neural network calculations frim scratch in Fortran 90 language.

Chapter IV – General Conclusions and Suggestions for Future Work

Chapter IV.1- General Conclusions

The first part of this work showed that the chosen set of equations, as well as the numerical methods, were able to describe properly the experimental breakthrough curve. This article also showed that the optimization of the process pressures alone were able to give a maximum nitrogen purity of 96.6% from an inlet current with 15% methane. This was an indication that other process variables should also be regarded in the optimization process. However, optimizing the DAE model is a computationally intensive task, inspiring us to come up with the ANN surrogate model. In the second work, we showed that directly optimizing the DAE model took about 15 hours, an undesirable amount of time for process design purposes. With the aid of an ANN surrogate model we were able to perform an optimization procedure in less than a second if the checking phase is disregarded and about fifty seconds otherwise. As compared to the optimization procedure in the first article, we were able to find a maximum nitrogen purity of 99.5%. Therefore, we conclude that a PSA with methane selective adsorbent is able to purify the waste nitrogen stream from a cryogenic distillation unit.

Chapter IV.2 - Suggestion for future work

For future work, we recommend the following analysis.

- Comparative study of various numerical methods, including Central Differentiation Schemes (CDS), higher order finite differences and UDS with more finite volumes.
- Make rigorous mesh convergence analysis, including mesh Courent and Péclet numbers.
- Behavior of the reduced order model in the presence of noise and uncertainty.
- ANN training using experimental pilot/actual plant data.
- Recurrent neural networks to model the adsorption/desorption dynamics.
- Development of a general purpose ANN in order to take into account material properties.
- Testing and screening with multiple adsorbents.
- Integration between the reduced order model and process simulator.
- Testing combined adsorption, membrane and distillation processes.