POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Multi-objective optimization of DR-PSA units using surrogate models

Laurea Magistrale in Chemical Engineering - Ingegneria Chimica

Author: Claudia Pavesi

Advisor: Prof. Renato Rota

Co-advisor: Prof. Giuseppe Storti

Academic year: 2024-2025

1. Introduction

Adsorption processes can separate efficiently components with similar boiling points, azeotropes and thermolabile compounds. The mixture to be separated or purified is fed into the adsorption bed and, after the designed operation time, the solid must be regenerated. Pressure Swing Adsorption (PSA) processes regenerate the adsorbent by lowering the pressure and purging it with clean gas.

Two main process configurations have been proposed in the literature: stripping PSA, based on the Skarstrom cycle, and enriching PSA, based on the Diagne cycle. The first allows to collect only the least adsorbable species at high purity while with the second only the most adsorbable species is collected at high purity.

This is true for every ratio P_{high}/P_{low} and it is due to thermodynamic constraints related to the saturation of the solid surface.

To overcome such a limitation the process has been modified into the so-called Dual Reflux Pressure Swing Adsorption (DR-PSA), in which the two aforementioned PSA configurations are combined. Figure 1 shows one of the possible DR-PSA process configurations, which involves feeding a binary mixture to be separated to the bed at the highest pressure and carrying out the pressure switch between the two columns using a heavy component-rich stream; this process configuration is usually named DR-PH-A process.

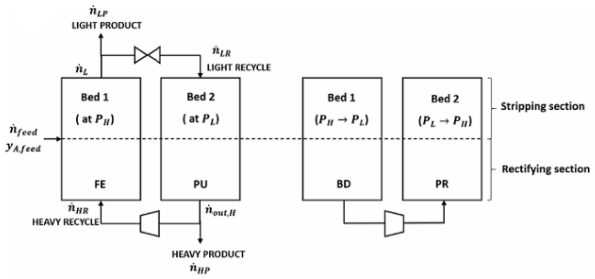

Figure 1: DR-PH-A configuration scheme, adapted from Rossi et al. [3]

The main limitation of DR-PSA processes is that their design methodologies are not well established yet. To be precise, a rigorous way to design DR-PSA units exists, but only under very strict simplifying assumptions based on the so-called Equilibrium Theory, that often implies working under unrealistic conditions. Considering more realistic conditions, trial-and-error approaches have been proposed to design DR-PSA processes. Such heuristic methodologies are heavily based on multiple experiments (or simulations) needed to identify a suitable oper-

ating window for the specific mixture under consideration.
Florit et al. [1] coupled the detailed mathematical model of a DR-PH-A process with an optimizer, obtaining a Pareto front that related the purity of the product stream with the work required by the unit. The limitation of this approach is the intrinsic complexity of the detailed model, that causes the computational time of the optimization to be in the order of weeks or even months. The aim of this Master of science thesis is to do a preliminary analysis in order to understand if the detailed model can be substituted with a surrogate model that is far simpler: in this case the time required to complete the DR-PSA design would drop to much lower values with respect to the current ones. To create the surrogate model the Matlab function $fitrgp$ has been used: this tool exploits gaussian process regression models to create models starting from a data training set. The required training set will be based on the results of detailed simulations. However, given the predictive power of this type of surrogate models, the size of this data set is expected to be smaller than that required using standard methods, thus resulting in a saving of computational effort.

2. State of the art

The main assumptions of the detailed model of DR-PSA processes are:

- ideal gas behavior,
- ideal plug flow (no axial mixing),
- isothermal operation,
- the mass transfer between solid and gas phases is represented by the Linear Driving Force (LDF) model,
- the pressure drops along the column are computed through the Blake-Kozeny equation,
- axial dispersion is neglected, given the very large values of axial Peclet number typical of these units.

Hence, this is composed by the following set of partial differential equations (PDEs):

$$\varepsilon_T \frac{\partial P}{\partial t} + \frac{\partial (uP)}{\partial z} + \rho_B RT \sum_{i=1}^{2} \frac{\partial q_i}{\partial t} = 0 \qquad (1)$$

$$\varepsilon_T \frac{\partial (Py_i)}{\partial t} + \frac{\partial (uPy_i)}{\partial z} + \rho_B RT \frac{\partial q_i}{\partial t} = 0 \qquad (2)$$

$$\frac{\partial q_i}{\partial t} = k_{\mathrm{LDF},i} \left(q_i^* - q_i \right) \qquad (3)$$

$$\frac{\partial P}{\partial z} = -\frac{u}{\frac{150}{4} \frac{1}{r_P^2} \left(\frac{1-\varepsilon_B}{\varepsilon_B} \right)^2 \mu} \qquad (4)$$

Where Eq 2.4 is the global material balance and Eq 2.5 is the material balance of component i. Eq 2.6 is the material balance of component i on the solid surface and Eq 2.7 is the Blake-Kozeny one, used to evaluate the velocity profile along the column axis. The Langmuir isotherm allows to evaluate the concentration of the adsorbed species at equilibrium condition, following the formula presented below:

$$q_i^* = \frac{a_i P y_i}{1 + b_i P y_i} \qquad (5)$$

In this work the integration of these PDEs is carried out using the Finite Volume Method, that has been proved optimal to simulate the DR-PSA process by Rossi et al. [3].

3. Materials and methods

3.1. Surrogate models

Machine learning can be used to construct predictive surrogate models using sufficient and representative training data. The recent improvements in machine learning algorithms enable the modeling of complex multi-variate phenomena as simple input-output relations.
In this thesis work the surrogate models have been created using Gaussian process regression because this method was found to be the best at predicting process outputs, with minimal training effort [2].
A Gaussian process (GP) defines a distribution over functions and it is fully specified by its mean function $m(x)$ and covariance function $k(x, x')$. This is a natural generalization of the Gaussian distribution: while the latter is over vectors, the Gaussian process is over functions. In symbols:

$$f \sim \mathcal{GP}(m, k) \qquad (6)$$

meaning: "the function f is distributed as a GP with mean function m and covariance function k".
On Matlab it is possible to train a Gaussian process regression model using the $fitrgp$ function

and this is what has been done in this thesis work. Once the model is trained, the *predict* Matlab function is used to generate an output starting from the model itself and a new input set.

3.2. Multi-objective optimization

The multi-objective optimization done to obtain the Pareto front has been carried out using the Matlab function *gamultiobj*. This routine uses a variant of the non-dominated sorting genetic algorithm, the NSGA-II, that is particularly robust and can escape local minima.
In this work the optimization revolves around minimizing the work required by the DR-PSA process and maximizing the molar fraction of the heavy component in the product stream. Since the *gamultiobj* can only minimize the objective functions, the two functions considered are *work* and $1/y_H$.

4. Results and discussion

The parameters varied in the optimizations are three: P_{low}, $\dot{n}_{feed}$ and G, but not all combinations of these parameters yield physically meaningful results. Thus, a 3D feasibility region has been identified by testing 5 values of $\dot{n}_{feed}$ (2e-4 mol/s, 2e-3 mol/s, 3e-3 mol/s, 6e-3 mol/s, 8e-3 mol/s) and observing where the model converged. The converged data points were then split into training and validation subsets.
At first, a linear surrogate model was tested for predicting both the heavy component purity and the compression work. The training set has three values of $\dot{n}_{feed}$: 2e-4 mol/s, 3e-3 and 8e-3 mol/s, in order to have a low value, a medium value and a high value. As expected, the auto-validation and validation errors of this surrogate model, calculated as mean absolute percentage errors (MAPE), are high: up to 178% for the work predictions. Consequently the Pareto front generated fails to capture the trade-off between purity and work, especially at lower purity values, as shown in the figure below (the Pareto front obtained by Florit et al. [1] is considered as reference).

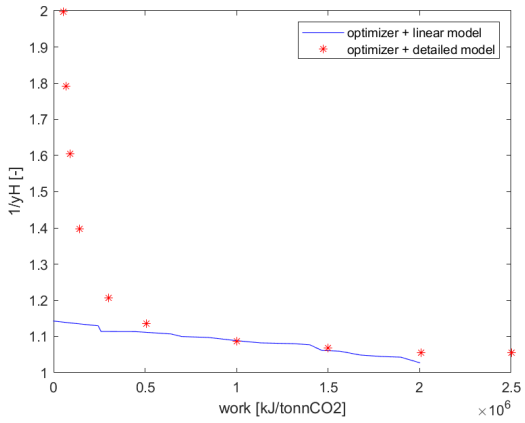

Figure 2: Pareto front using a linear surrogate model.

A more complex surrogate model is needed, and so the *fitrgp* function is used to generate new surrogate models. The model created giving to this Matlab function the same training set provided to the linear fitting is here called surrogate model 1.
The auto-validation and validation errors are significantly lower with respect to the linear surrogate case, as shown in the following table.

Table 1: Auto-validation of surrogate model 1

MAPE (yH model) [%]	MAPE (work model) [%]
2.8e-3	0.1

Table 2: Validation of surrogate model 1, with respect to $\dot{n}_{feed}$=2e.3, 6e-3 mol/s

MAPE (yH model) [%]	MAPE (work model) [%]
10.5	24
27	102

The Pareto front obtained coupling surrogate model 1 with the optimizer (*gamultiobj*) is presented below:

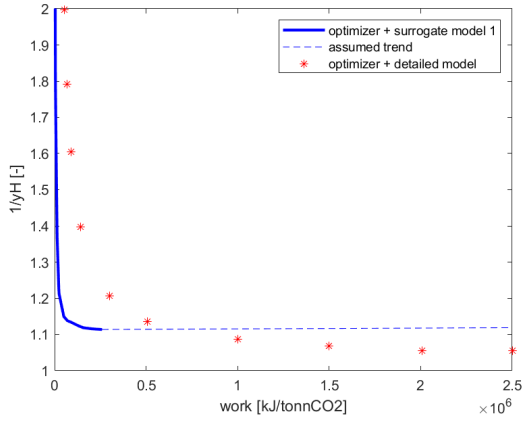

Figure 3: Pareto front using surrogate model 1

This Pareto front is better than the one generated using linear surrogate models, because it does not predict negative values of work and it follows the trend of the reference front both in the higher branch of the plot and in the region where the trend inverts. Note that the Pareto front obtained using the surrogate model stops for $work \sim$ 0.25e6 kJ/tonnCO2. This could be because for higher work values the $1/y_H$ values predicted by the surrogate model could be slightly higher than the ones predicted using the detailed model. Hence they would be higher also than the values predicted by surrogate model itself when work is lower than 0.25e6 kJ/tonnCO2, and so there would be points with lower purity and higher work required. They would not be optimal points and so the Pareto front would stop.
Anyway, the lower branch of the Pareto front has a lower practical interest because it comprehends the operative points that require to provide high works to the DR-PSA process.
A way to increase the accuracy of the surrogate models could be to increase the size of their training sets. Here this is done without using again the detailed model, to have shorter computational times: three two-parameter surrogate models are built, one for each value of $\dot{n}_{feed}$ present in the current training set. These models have low errors and they are more accurate thanks to their simplicity (just two parameters are varied), hence they are used to generate points that are added to the training set ones. In this way, a new training set containing a low, medium and high value for all the three parameters is created. Giving this new training set to the $fitrgp$ Matlab function a new surrogate model is generated, here called surrogate model 2. Its errors are similar to the ones of surrogate model 1 and the Pareto front that it generates is almost identical to the one generated by surrogate model 1.
Similar results can be obtained by using a new training set created by applying the Box-Behnken design on the training set of surrogate model 2. The surrogate generated is called surrogate model 3 and this experiment has been carried out to demonstrate that the Box-Behnken design is capable of choosing a representative data subset.
To prove that increasing the number of points in the training set the accuracy of the surrogate models increase, a new training set has been considered. This set is made by all the points which have been simulated with the detailed model and the surrogate model generated using it is here called surrogate model 4. Its errors are low and satisfying as the ones of the previous surrogate models and the Pareto front is improved, as shown in the figure below:

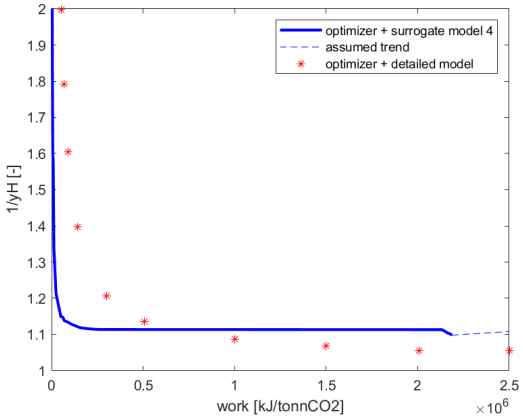

Figure 4: Pareto front using surrogate model 4

The last test that can be done using the available detailed model data is to consider 5 two-parameter surrogate models, one for any value of $\dot{n}_{feed}$ considered in this thesis work. For each of the 5 two-parameters surrogate models a Pareto front has been generated; the final Pareto front is generated by the envelope of these 5 fronts and it is shown in the following figure:

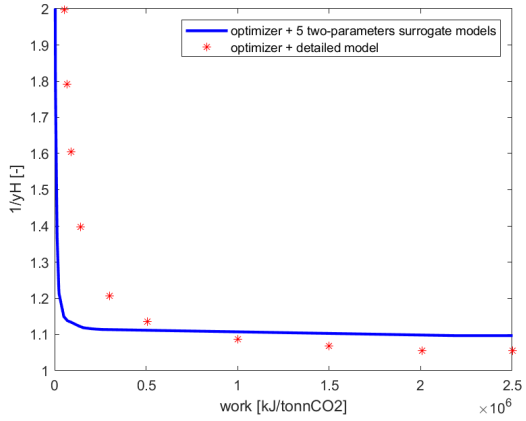

Figure 5: Pareto front using surrogate model 5

This Pareto front is the best among the ones generated with surrogate models in this thesis work; this is a further confirmation that surrogates with a lower number of input parameters are more accurate.

5. Conclusions

In this thesis it has been demonstrated that the surrogate models built with *fitrgp* to simulate the DR-PSA process provide similar results with respect to the detailed model, and they do so requiring much lower computational times. Indeed in the work of Florit et al. [1] hundreds of detailed model simulations have been carried out, while in this work the number of detailed model simulation is in the order of tens.
It also has been verified that increasing the size of the training sets provided to the *fitrgp* function the accuracy of the surrogate models built increases.
These preliminary results suggest that surrogate models built using the Matlab function *fitrgp* can substitute the detailed model of DR-PSA processes, decreasing significantly the computational time required to design such processes. However, more work is required both to confirm such a conclusion as well as to identify the optimal distribution in the space of parameters of the training set data.

References

[1] F. Florit, L. Dalla Giovanna, G. Storti, and R. Rota. On the design of a dual reflux pressure swing adsorption process. *Industrial Engineering Chemistry Research*, 63(26):11626–11636, 2024.

[2] K. N. Pai, V. Prasad, and A. Rajendran. Experimentally validated machine learning frameworks for accelerated prediction of cyclic steady state and optimization of pressure swing adsorption processes. *Separation and Purification Technology*, 241:116651, 2020.

[3] E. Rossi, M. Paloni, G. Storti, and R. Rota. Modeling dual reflux-pressure swing adsorption processes: Numerical solution based on the finite volume method. *Chemical Engineering Science*, 203:173–185, 2019.