CHEMICAL ENGINEERING TRANSACTIONS
VOL. 105, 2023
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.cetjournal.it
Guest Editors: David Bogle, Flavio Manenti, Piero Salatino

ISBN 979-12-81206-04-5; **ISSN** 2283-9216
DOI: 10.3303/CET23105076

Enhancing Computational Performances in Chemical Processes Costs and Emissions Prediction: a Surrogate Modelling Based Approach

Alessandro Di Pretoro*, Stéphane Negny, Ludovic Montastruc

Laboratoire de Génie Chimique, Université de Toulouse, CNRS/INP/UPS, Toulouse, France
alessandro.dipretoro@ensiacet.fr

The increasing amount of variables to be accounted for in chemical processes optimization and the need to have a systemic approach to include all the steps of the industrial production chain implies the exponential growth of the model equations to be solved at the same time. In fact, in order to have an optimal industrial system, the analysis should start from raw materials supply and include demand-side, process side and logistic from the meso- to the macro-scale perspective. Moreover, beside economics, environmental impact, flexibility and scheduling should be coupled in a multi-objective optimization loop. This approach results in a computational effort that is way higher than that required in the past for conventional process optimal design.
Therefore, innovative computational strategies should be implemented in order to ease the optimization loop. During the last decade, surrogate modelling has seen renewed interest for this purpose in chemical process engineering and it has been widely used for feasibility analysis, optimization and optimal scheduling. In this preliminary study we exploit a surrogate modelling approach for costs and emissions calculation for a simple separation process. A distillation unit is simulated by means of ProSimPlus process simulator to retrieve a set of physical and economic data over the operating domain of interest. After that, the sampling strategy is selected according to the suggested standards and adopted to generate a surrogate modelling with a Response Surface Methodology approach by means of ALAMO software. The output variable of interest for this study have been identified as the unit costs and the emissions related to the energy consumption. Despite the complexity of chemical equilibrium in multistage units, the obtained results show good agreement with those generated by the phenomenological models with a computational time whose magnitude is two orders lower. In conclusion, this methodology is worth deeper studies in order to be exploited for more complex systems and have even more benefits with the increasing complexity of the case study when coupling more units in different configurations.

1. Introduction

During the last decade, the majority of the efforts in the research domain are focused on the improvement of computational performances and Artificial Intelligence in order to deal with the Big Data challenge. In chemical and process engineering one of the most established tool is surrogate modelling. The main advantage of this approach on chemical processes is due to the possibility to reduce the computational complexity of non-linear equations involved both in unit operations and in thermodynamic behaviour models as concerns the operations and simulation as well as the early feasibility analysis phase (Bhosekar and Ierapetritou, 2018). In the last years, several applications have been tested on process systems (McBride and Sundmacher, 2019), on the Demand-Side management part (Di Pretoro et al., 2022a) as well as on process dynamics and control (Di Pretoro et al., 2022b). Based on these studies, the main remark concerning the effectiveness of the modelling approach in almost all cases refers to the sampling strategies. In fact, the way the physical systems values are collected and the size of the dataset (Davis et. al, 2018) before the modelling step plays a critical role on the obtained model quality. A further enhancement of the model can also be obtained by including feasibility constraints during the Design of Experiment generation as recently proved by Zinare Mamo et al. (2023). One of the aspects that still lacks of research results in terms of data-driven modelling is that related to environmental impact and process emissions. That is why, in this work, we explore the possibility to model energy consumptions and related

Paper Received: 30 January 2023; Revised: 1 July 2023; Accepted: 9 May 2023
Please cite this article as: Di Pretoro A., Negny S., Montastruc L., 2023, Enhancing Computational Performances in Chemical Processes Costs and Emissions Prediction: a Surrogate Modelling Based Approach, Chemical Engineering Transactions, 105, 451-456 DOI:10.3303/CET23105076

emissions for non-trivial process unit operations such as distillation. In particular, both an ideal mixture case study and a more complex one will be analyzed by coupling simulation, coding and dedicated modelling commercial software in order to assess the performances of technologies that are already available on the market for the users.
In the following sections, further details about the case study and the proposed methodology are respectively described and the obtained outcome is properly presented and commented to have a complete overview about advantages and inconveniences of the proposed approach.

2. Case studies

In this research work two case studies (cf Figure 1) of increasing complexity will be discussed in order to compare the same methodology under the perspective of different computational performances.
Further details concerning operating parameters and product specifications are discussed in the corresponding subsections here below in order to provide a complete overview of the systems under analysis.

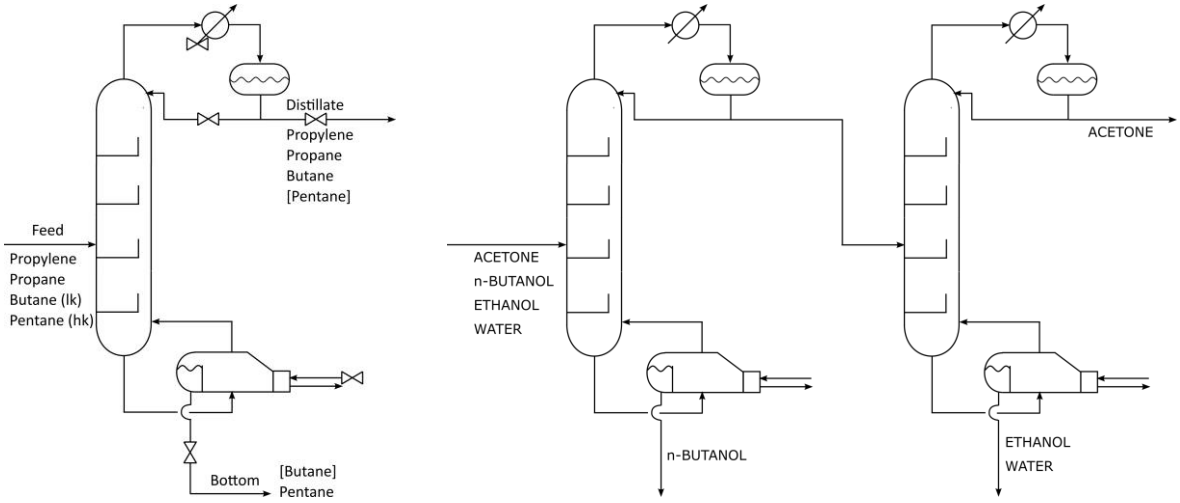

Figure 1: a) Ideal mixture distillation column b) Non-ideal mixture distillation train

2.1 Ideal mixture distillation

The first case study is based on the example proposed by Di Pretoro et al. (2019) for the Butane–Pentane mixture separation. It is an ideal distillation operation whose main purpose is to obtain the two components at the desired purity. The light impurities traces are removed along with the distillate stream and the selected thermodynamic model is Soave–Redlich–Kwong. The feed stream properties, operating parameters and specifications are listed in detail in Table 1 while the specifications, both related to the bottom product stream, are butane molar fraction equal to 0.01786 and 0.97 pentane recovery ratio respectively.

Table 1: Process and variable parameters for the Butane-Pentane separation*

Variable	Symbol	Value	Unit	Min value	Max value	Discretization
Pentane feed flowrate (hk)*	F_5	2.743	mol/s	2.469	3.017	0.1%
Butane feed flowrate (lk)*	F_4	6.863	mol/s	6.177	7.549	0.1%
Propane feed flowrate	F_3	0.053	mol/s			
Propylene feed flowrate	$F_{3=}$	0.055	mol/s			
Feed temperature	F_T	Bubble	K			
Feed pressure	F_P	$15 \cdot 10^5$	Pa			
Number of trays (feed)	N	20 (10)	/			
Column top pressure	P	$4 \cdot 10^5$	Pa			

2.2 Non-ideal mixture distillation

The second case study is based on the distillation train example proposed by Di Pretoro et al. (2020) for the non-ideal Acetone–Butanol–Ethanol and Water mixture. In particular, as discussed in the referred publication, the first column is the one affected by the most relevant perturbations in terms of operating conditions.

Therefore, this study will only focus on it to ease both the modelling and the analysis steps. For this example, the NRTL thermodynamic model was used to take into account non-ideal behaviour between water and alcohols. All related parameters are listed in Table 2 while the specifications for the n-Butanol product purity are 0.99 mass fraction and 0.9604 recovery ratio.

Table 2: Process and variable parameters for the ABE/W mixture*

Variable	Symbol	Value	Unit	Min value	Max value	Discretization
Acetone feed flowrate	F_A	12.030	mol/s			
n-Butanol feed flowrate*	F_B	61.328	mol/s	55.195	67.461	0.01%
Ethanol feed flowrate	F_E	3.839	mol/s			
Water feed flowrate*	F_W	12.428	mol/s	11.185	13.671	0.01%
Feed temperature	F_{T2}	361.26	K			
Feed pressure	F_{P2}	$1.013 \cdot 10^5$	Pa			
Number of trays (feed)	N_2	16 (10)	/			
Column top pressure	P_2	$1.013 \cdot 10^5$	Pa			

3. Methodology

3.1 Unit simulation and sampling

The distillation units have been modelled in both cases by the standard distillation module available in ProSim process simulator over the entire uncertain domain by means of the sensitivity analysis tool. After each run, the obtained results were stored in a .csv file that is then imported in Matlab for the sampling step. However, before performing the DoE phase, the total energy consumption of the system is assessed and converted into the associated emissions. For this case study, the selected environmental impact indicator is the Global Warming Potential (GWP) whose unit is the kg-eq$_{CO2}$. For the particular case of distillation units, its value is calculated according to the hypothesis and assumptions discussed in Gadalla et al. (2006) by means of the equations:

$$CO_2 = \left(\frac{Q_{fuel}}{NHV}\right) \cdot \left(\frac{C\%}{100}\right) \cdot \alpha \quad (1)$$

$$Q_{fuel} = \left(\frac{Q_{reb}}{\lambda_{steam}}\right) \cdot (h_{steam} - 419) \cdot \left(\frac{T_F - T_0}{T_F - T_S}\right) \quad (2)$$

where Q_{fuel} and Q_{reb} are the fuel and reboiler duties, NHV is the fuel Net Heating Value, $C\%$ is the carbon content of the fuel, α is the *CO_2-to-C* molar mass ratio equal to 3.67, λ_{steam} and h_{steam} are the latent and sensitive enthalpy for steam referred to 419 kJ/kg at 100 °C and, finally, T_F, T_S and T_0 are respectively the flame, stack and standard temperature in K. Furthermore, the operating costs function associated with the energy consumption is assessed as well considering $7.78 \cdot 10^6$ €/kJ.

The selected sampling approach for this work is the Latin Hypercube Sampling (LHS), conceived to have a reduced variance sample with a good space-filling capacity. A dedicated function is already present in Matlab for defined hyperspace size and overall number of points. The obtained sample is finally exported in a new .csv file that will be processed by the surrogate modelling software as discussed in the next subsection.

3.2 ALAMO modeling

The modeling step was performed via ALAMO (Automatic Learning of Algebraic MOdels) whose interface is shown in Figure 2 with a particular focus on the base functions selection window. ALAMO was developed in 2014 by Cozad, Sahinidis and Miller, and its purpose is to address the problem of derivative-free optimization. To be more precise, it aims at generating the simplest and most accurate algebraic surrogate model of black-box systems, for which an experimental set-up or a simulator is currently available. It is based on a three-step iterative process that, in the order, interrogates the initial DoE, creates a model by optimizing the coefficients of an eventually selected subset of simple basis functions among the available ones and making a linear combination of them and, finally, by updating the sample points where the model lacks of accuracy by means of an adaptive-sampling methodology based on derivative-free optimization solvers. Once the model is obtained, the resulting Response Surface function is then implemented in Matlab to perform the data analysis step and assess the quality of the obtained model with respect to function values that were not exploited for the previous modeling phase in order to have unbiased results.

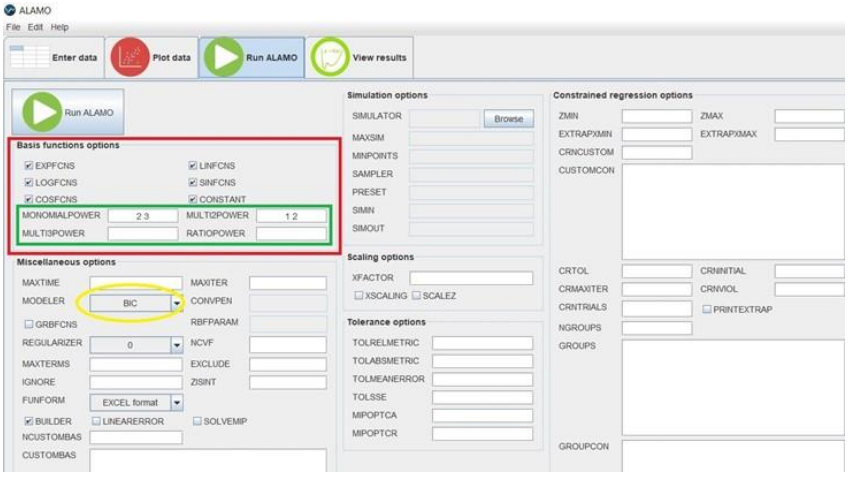

Figure 2: ALAMO software: Interface window *for base functions selection*

In particular, the procedure was performed twice accounting, in the second case, for smaller subsamples as suggested by Zinare Mamo et al. (2023) in order to compare the two performances. The selected indicators for this study are the normalized Maximum Absolute Error (MAE) and mean absolute error (mae) expressed as a percentage. Moreover, the amount of variables, i.e. the amount of coefficients for the basis functions is reported as well in order to give an idea of the complexity of the resulting model.

4. Results

The modelling procedure has been then carried out according to the methodology discussed in the previous section. The obtained results are discussed in the following subsections according to the specific case study.

4.1 Ideal mixture distillation

The first case study was simulated and the related model derived. Due to the selected discretization accuracy, 10201 runs were required to cover the entire domain of interest with an overall computational time (modelling phase included) of about 12 minutes. As expected, the most relevant parameter in terms of energy consumption and, thus, emissions is the reboiler duty. The phenomenological model behaviour and that obtained by Response Surface Methodology are compared in Figure 3.

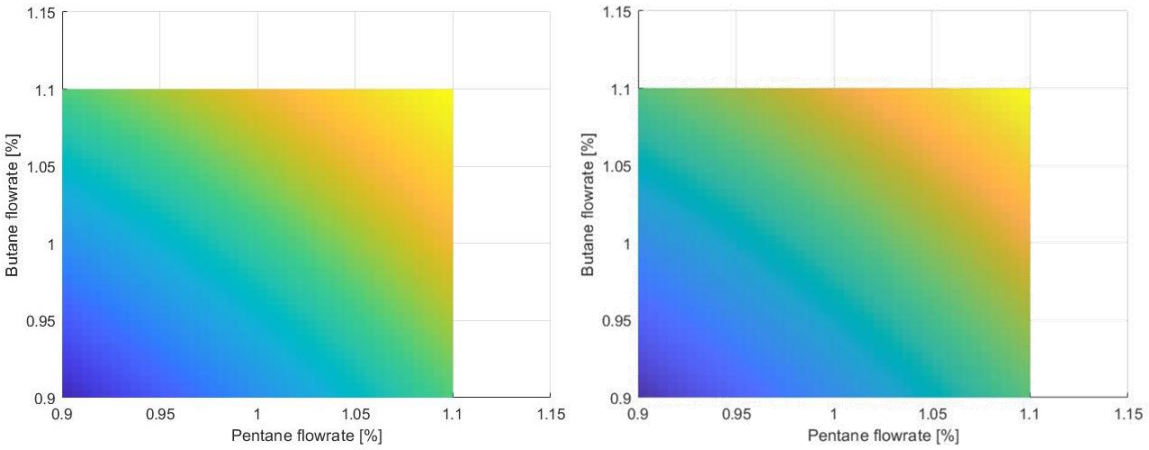

Figure 3: Reboiler duty emissions [kg-eq CO_2] for ideal distillation a) simulation values b) surrogate model

The first conclusion that can be drawn from these graphs is that the compliance between the simulation model and the surrogate one is almost complete. Energy consumption and emissions increase, as expected, almost linearly with the increasing of the overall system flowrates. A further confirmation about the quality of the model can be obtained by looking at the performance indicators resumed, for both case studies, in Table 3.

In general, the mae does not overcome the 1% value. The reboiler heat duty has the main impact on the CO_2 emissions and, thus, a higher wight on the average total duty value model. The amount of parameters is generally moving in the interval between 5 and 8 and the involved functions are mainly linear and quadratic. It is worth remarking that trigonometric base functions have been excluded since they could incur into aliasing effect and bias the model accuracy with respect to the data set.

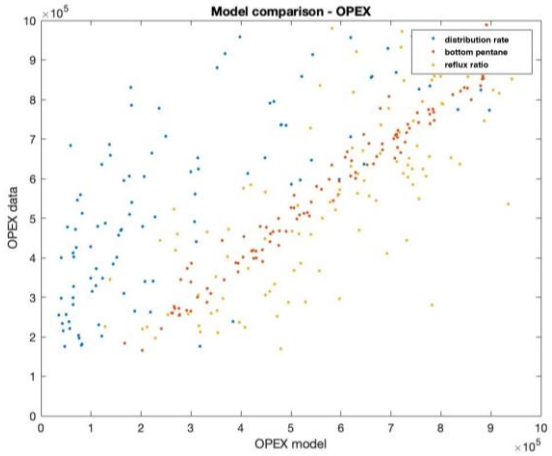

Figure 4: OPEX comparison between surrogate model and simulation data

Finally, for the Butane-Pentane column, the calculation for OPerating EXpenses was performed as well as shown in Figure 4 based on the subsets organization discussed in Zinare Mamo et al. (2023) ordered according to the distribution product rate, bottom pentane partial flowrate and reflux ratio. As it can be noticed, the second and the third parameters have a higher impact on the OPEX function due to their closer relationship with the energy duty and, thus, the operating expenses. Therefore, in a longer perspective, if the algorithm needs to be improved, these parameters are those worth better investigation.

4.2 Non-ideal mixture distillation

The same approach is then followed for the ABE/W mixture. The obtained results for emissions are presented in Figure 5 for a one subset approach. As it can be noticed, there is good agreement between the left and right graphs with some discrepancies on the borders. As it can be noticed, there is good agreement between the left and right graphs with some discrepancies on the borders. Duties and, thus, emissions are higher in the regions where the operating conditions are more severe, i.e. when flowrates are more relevant and in proximity of the distillation boundaries due to the fact that the separation is approaching the azeotropic species composition.

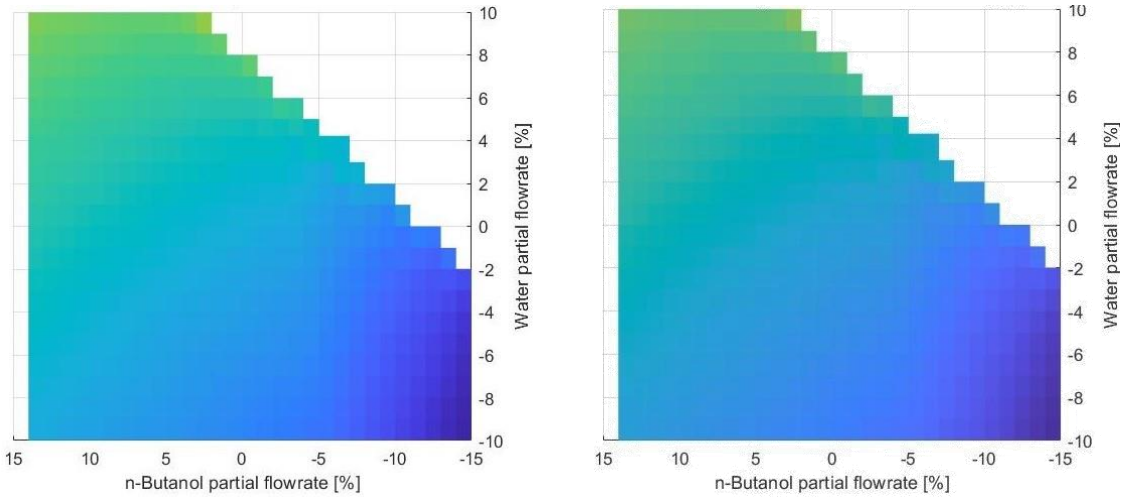

Figure 5: Reboiler duty emissions [kg_{eq} CO_2] for non-ideal distillation a) simulation values b) surrogate model

Table 3: Performance indicators for the obtained models

	Variable	MAE [%]	mae [%]	n parameters	subsets
Case study #1	Reboiler heat duty	3.12	1.03	6	
	Condenser heat duty	2.28	0.85	5	
	Total heat duty	2.75	0.92	8	
		1.96	0.51	5 (avg)	yes
Case study #2	Reboiler heat duty	4.29	1.62	7	
	Condenser heat duty	3.19	1.27	4	
	Total heat duty	3.84	1.53	7	
		2.17	0.89	8 (avg)	yes

Also for this example, as reported in Table 3, the reboiler is the main part of the unit in terms of environmental impact, the mae is no lower than 2% for both the utilities and the use of subsets allows to further reduce these values below the 1%.

5. Conclusions

The presented research work explores the opportunity to predict energy consumption based unit operations emissions and related costs by means of surrogate modelling exploiting the Surface Response Methodology by means of the ALAMO software. The outcome of the proposed methodology on the distillation unit case study allows to accurately define the emissions of a distillation unit both for a simple thermodynamic unit operation and for a more complex one involving an azeotropic species over the operating domain of interest.
Moreover, the partition of the domain into subsets allowed to further increase the model accuracy without relevant drawbacks affecting the computational performances.
Based on this results, the surrogate modelling approach for emissions estimation is a tool worth further investigation in future research. Applications of particular interest could be more complex systems, such as distillation trains, as well as processes involving reactors or also power plants.

References

Bhosekar A., Ierapetritou M., 2018. Advances in surrogate based modeling, feasibility analysis, and optimization: A review. Computers & Chemical Engineering 108, 250–267.

Cozad A., Sahinidis N. V., Miller D. C., 2014. Learning surrogate models for simulation-based optimization, AIChE Journal 60, 2211–2227.

Di Pretoro A., Montastruc L., Manenti F., Joulia X., 2019. Flexibility Analysis of a Distillation Column: Indexes Comparison and Economic Assessment, Computers & Chemical Engineering 124, 93–108.

Di Pretoro A., Montastruc L., Manenti F., Joulia X., 2020. Flexibility assessment of a biorefinery distillation train: Optimal design under uncertain conditions. Computers & Chemical Engineering, 138, pp.106831.

Di Pretoro A., Bruns B., Negny S., Grünewald M., Riese J., 2022a. Demand response scheduling using derivative-based dynamic surrogate models. Comput. & Chem. Eng. 160, 107711.

Di Pretoro A., Tomaselli A., Manenti F., Montastruc L., 2022b. Dynamic Surrogate Modeling for Continuous Processes Control Applications, Computer Aided Chemical Engineering 51, 91-96.

Davis S. E., Cremaschi S., Eden M. R., 2018. Efficient Surrogate Model Development: Impact of Sample Size and Underlying Model Dimensions. Comput. Aided Chem. Eng. 44, 979-984.

Gadalla, M., Olujić, Ž., Jobson, M., Smith, R. 2006. Estimation and reduction of CO2 emissions from crude oil distillation units. Energy 31, 2398–2408.

McBride K., Sundmacher K., 2019. Overview of Surrogate Modeling in Chemical Process Engineering. Chemie Ingenieur Technik 91, 228–239.

Zinare Mamo T., Di Pretoro A., Chiari V., Montastruc L., Negny S., 2023. Benefits of feasibility constrained sampling on unit operations surrogate model accuracy. Comput. & Chem. Eng., 108210.