Systems & Control Transactions

Research Article - Peer Reviewed Conference Proceeding
ESCAPE 35 - European Symposium on Computer Aided Process Engineering
Ghent, Belgium. 6-9 July 2025
Jan F.M. Van Impe, Grégoire Léonard, Satyajeet S. Bhonsale, Monika E. Polańska, Filip Logist (Eds.)

PSE PRESS

A Bayesian optimization approach for data-driven Petlyuk distillation column

Alexander Panales-Pérez[a], Antonio Flores-Tlacuahuac[b]*, Luis Fabián Fuentes-Cortés[c], Miguel Angel Gutierrez-Limon[d] and Mauricio Sales-Cruz[e]

[a] Tecnológico Nacional de México, Instituto Tecnológico de Celaya, Departamento de Ingeniería Química, Celaya, Guanajuato, Mexico, 38010.
[b] Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Campus Monterrey Ave. Eugenio Garza Sada 2501, Monterrey, N.L, 64849, México.
[c] Department of Energy Systems and Environment, IMT Atlantique, GEPEA, rue Alfred Kastler, Nantes, 44000, France.
[e] Departamento de Energía, Universidad Autónoma Metropolitana-Azcapotzalco, Av. San Pablo 180. C.P. 02200, Ciudad de México, México.
[d] Departamento de Procesos y Tecnología, Universidad Autónoma Metropolitana-Cuajimalpa, Av. Vasco de Quiroga 4871. C.P. 05348, Ciudad de México, México.
* Corresponding Author: antonio.flores.t@tec.mx

ABSTRACT

Recently, the focus on increasing process efficiency to reduce energy consumption has driven the adoption of alternative systems, such as Petlyuk distillation columns. It has been proven that, when compared to conventional distillation columns, these systems offer significant energy and cost savings. From an economic standpoint, achieving high-purity products alone does not ensure the feasibility of a process. Instead, balancing the trade-off between product purity and cost necessitates multi-objective optimization. While conventional optimization methods are effective, novel strategies like Bayesian optimization offer distinct advantages for handling complex systems. Bayesian optimization requires no explicit mathematical model and can efficiently optimize even when starting from a single initial point. However, as a black-box method, it demands a detailed analysis of hyperparameters, such as the acquisition function and the number of initial points, to ensure optimal performance. This work presents a case study of a Petlyuk distillation column, focusing on the influence of key hyperparameters in Bayesian optimization. Additionally, the role of uncertainty in the optimization process is explored, as it is a critical factor in real-world applications. By simulating uncertainty, this study provides insights into its impact on the optimization process.

Keywords: Process Design, Distillation, Artificial Intelligence, Aspen Plus.

INTRODUCTION

In chemical engineering, complex systems such as Petlyuk distillation columns have been analyzed to achieve considerable energy savings compared to traditional distillation columns, due to its configuration, a Petlyuk column consists of a prefractionator and a main column, where vapor and liquid streams are thermally coupled, thus, unlike conventional sidestream columns, it uses a single reboiler and a single condenser. Therefore, various methodologies have been employed for the operation and design of Petlyuk columns, including nonlinear programming (NLP) [5] and mixed-integer nonlinear programming (MINLP) [10]. Traditional solvers for NLP and MINLP problems frequently face convergence challenges. In practice, these solvers often fail to identify the global optimum due to the non-convex nature of the problem. Alternative solution approaches, such as genetic algorithms, have been explored. However, their performance improvements over the past two decades have been limited [9], highlight the need to explore alternative methods, such as data-driven optimization. In this manner, Bayesian optimization (BO), a black-box method, does not require an explicit mathematical model of the problem suited to this problem. The core of this method is the construction of a surrogate probabilistic model, most commonly a Gaussian process (GP), which efficiently guides the sequential acquisition of new data through an acquisition function. Whose role is to balance the trade-off between exploration and exploitation. In practical terms, exploration targets points where the surrogate model variance is high, while exploitation focuses on points with a high surrogate model mean. This balance allows BO to intelligently select evaluation points for the objective function, thereby guiding the search toward the optimal value effectively. Moreover, the optimal solution is reached either after a specified number of iterations or when a stopping condition based on the objective function is achieved. This method is especially suited for handling expensive objective functions, where “expensive” refers to high computational or experimental costs. Additionally, the objective function can be expressed as a multi-objective function [4]. The problem to be addressed can be formally stated as follows: "Given an initial set of potentially noisy process measurements of system outputs, determine the optimal operational policy that maximizes a defined operational target. In this context, the process outputs may include stream compositions, such as mole fractions, or other composition metrics. Achieving the specified operational target requires selecting a set of input variables to enforce optimal operation. As evident from this definition, the problem falls within the domain of dynamic optimization. However, it is important to note that dynamic process calculations are not involved. Instead, it is assumed that the application of input variables drives the system to a new steady state, where calculations are repeated iteratively until the target conditions are satisfied. BO emerges as a robust alternative to the reinforcement learning methodologies commonly utilized in modern artificial intelligence systems, offering a structured and efficient approach to address such optimization challenges. This work proposes a data-driven optimization strategy tailored for complex process systems, such as the Petlyuk column, where traditional optimization methods are computationally expensive. The approach utilizes a multi-objective function that aims to minimize cost while maximizing the desirable component purities in each output stream of the Petlyuk column. Furthermore, the impact of random noise on the BO process is investigated to assess the feasibility of the proposed strategy in real-world applications

METHODOLOGY

The use of BO that can be used to evaluate expensive objective functions, as mentioned before the term “expensive” can be used for both economic or computational cost to evaluate a process or an experiment. In the absence of a Petlyuk column to realize the experiments, the commercial simulator Aspen plus was used to mimic the behavior to obtain the results of the expensive objective function. Moreover, in the figure 1 explains the Bayesian optimization process, where the core are two steps [2]: First, the estimation of the black-box function from data through a probabilistic surrogate model, the most common a GP and secondly, maximizing an acquisition function that trade off the parameters of exploration and exploitation. Therefore, the criteria of the convergence test can be in different ways as a stop criterion for a condition in the objective function or a defined number of iterations, the last one was selected for a more balanced comparative of the hyperparameters; by stop the optimization process after 50 iterations in all cases.

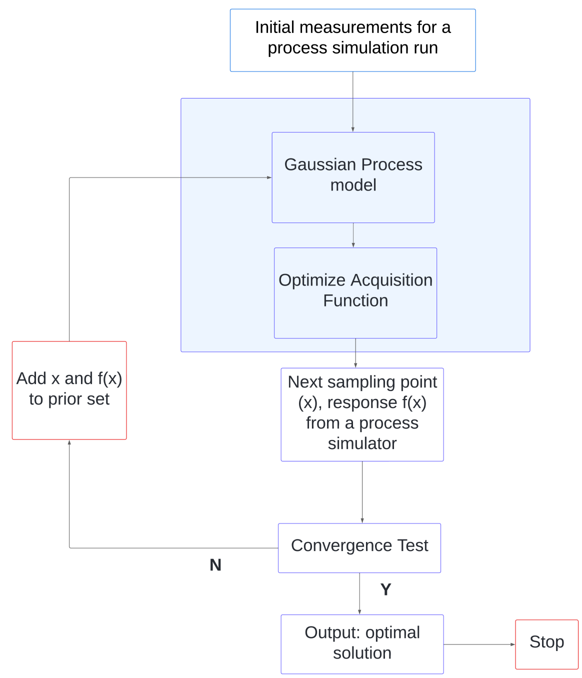

Figure 1. Flow chart for the BO process. Key steps, such as GP and the acquisition function, are denoted in blue. The optimization loop begins with the IP obtained from the simulation and adapted to the GP, and then the next sample is determined by the Acquisition Function. Furthermore, the loop terminates when the user objective function criteria are met for the convergence test.

To analyses the Petlyuk column, the involve mixture consists of feed flow (F=800 kmol/h) of benzene(A), toluene(B), and xylene(C) (figure 2) also known as a BTX mixture [3]. Therefore, the selected manipulated variables are the Distillate flow rate of the main column(D), reflux ratio (R), liquid flow rate (LR) and vapor flow rate(VR) to the pre-fractionator column, the interval of the manipulated variables in the same order are: D ∊ [5–100] kmol/h, R ∊ [2–35], LR ∊ [20–150] kmol/h and VR ∊ [20–150] kmol/h. The BO process is computing using the

Scikit-Optimize library [6], in Python and establish a connection with the ASPEN Plus simulator through Windows Component Object Model (COM) access. Additionally, all manipulable variables are normalized in the GP and denormalized for use in the Aspen Plus simulator.

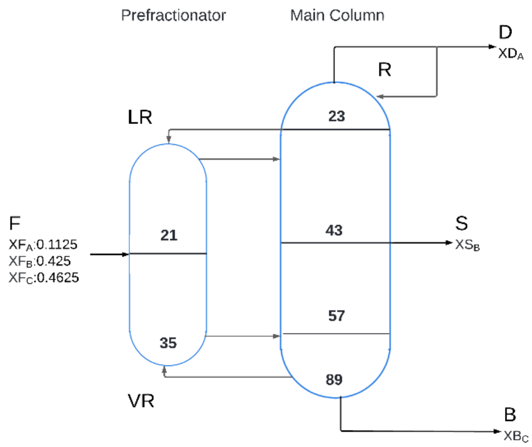

Figure 2. Flowsheet of a Petlyuk column system, consisting of a main column and a prefractionator, the locations of streams and stages are illustrated.

The multi-objective function, defined in Equation 1, is formulated as a minimization function, where we aim to maximize the desired composition of the component of interest in each stream: the distillate (XD_A), the lateral side (XS_B), and the bottoms (XB_C). In this case study, the function also seeks to improve economic performance by maximizing the desirable components in each stream while minimizing the total operational cost (TOC), which is calculated based on the costs of cooling water and steam incorporated into the ASPEN simulation. Moreover, the composition variables obtained from the Aspen Plus simulation are already normalized to values between 0 and 1. However, in the case of TOC, two preliminary BO runs are conducted, using default values Matern for the kernel and LCB for the acquisition function. These runs define the upper (UB) and lower (LB) bounds of TOC values by minimizing and maximizing TOC as an objective function. The resulting bounds are then used to perform a min-max normalization, as defined in Equation 2 [8]. This normalization ensures that the multi-objective function maintains an equitable trade-off among all decision variables.

$$\Omega_i = \left(XD_{B,i} + XD_{C,i} - XD_{A,i}\right) + \left(XS_{A,i} - XS_{C,i} - XS_{B,i}\right) + \left(XB_{A,i} + XB_{B,i} - XB_{C,i}\right) + \phi_{TOC,i} \quad (1)$$

$$\phi_{cost,i} = \frac{cost_i - LB_{cost}}{UB_{cost} - LB_{cost}} \quad (2)$$

Some aspects are analyzed: First, the number of initial points is varied. Instead of providing random information, these initial points are designed to generate insights into the possible optimal zone using an initial BO run with default values. This initial run starts with a single point at the center of the search space for all manipulated variables. The optimal result obtained from this run is then used to generate initial points (IPs) by varying the optimal manipulated variables within a random range of ±5%, creating new initial points. This approach aims to provide high-quality information to the Gaussian Process (GP) rather than relying on arbitrary random points in a non-optimal region. Second, this process is executed using different acquisition functions, including Expected Improvement (EI), Lower Confidence Bound (LCB), and Probability of Improvement (PI). Additionally, uncertainty is analyzed by incorporating noise into the results. This is done by introducing cumulative noise (ε) into the objective function value (Ω), adding random variations at a level of ±5% of the value of Ω, as shown in Equation 3.

$$\Gamma_i = \Omega_i + \varepsilon \quad (3)$$

RESULTS

As is denoted in tables (1-6), each of the different acquisition functions with the variation of the initial points was computed, with the iteration at which the optimal value is reached also counting the initial points added. First, in the case of EI (tables 1-2), there is no notable difference in the objective value of the optimal solution achieved with an increase in the number of initial points. Even with 30 initial points, the objective value does not show a significant increase compared to using 20 initial points; instead, it decreases to the level observed with 5 initial points. Moreover, this minimal change in the objective function indicates that when the purity product fractions are analyzed across all optimal solutions, there is no sharp difference. The purities achieved are 0.99, 0.96, and 0.96 for XD_A, XS_B, and XB_C, respectively, for all cases. On the other hand, only the flow of VR is the manipulable variable that shows a notable change from 150 kmol/h (1 initial point) to 110.02 kmol/h (5 initial points). Therefore, regarding the economic aspect, the solutions related to the R value fall within the same interval, which is directly related to operational costs. In this manner, the trend of no notable difference continues.

Secondly, for the LCB acquisition function (tables 3-4), there is a slight increase in the optimal value of the objective function. This results in a slight improvement in the achieved purity of the products, allowing the purity of the intermediate component X_{SB} to reach up to 0.96 and the bottom component X_{BC} to reach 0.98. However, in the trade-off with TOC, there is an increase of more than 0.47x 10^6 USD/year, and this occurs only when using more than 10 initial points. Only the case with 30 initial points can achieve a trade-off with similar economic performance to EI, while also achieving better purity values for the components (XS_B = 0.964; XB_C = 0.986).

Table 1: Manipulable variables Petlyuk column EI.

IP	Manipulable variables			
	D	R	VR	LR
1	79.55	16.75	150.00	150.00
5	79.0	16.26	110.02	148.91
10	78.59	16.53	115.76	150.00
20	78.59	16.30	121.28	150.00
30	78.66	16.30	132.32	150.00

Table 2: Results BO EI.

IP	Description			TOC x 10^6	Ω	Iter.
	XD_A	XS_B	XB_C			
1	0.999	0.961	0.967	5.11	-2.480	48
5	0.999	0.963	0.962	4.95	-2.487	39
10	0.999	0.964	0.965	5.00	-2.491	51
20	0.999	0.962	0.965	4.93	-2.494	48
30	0.999	0.960	0.965	4.93	-2.488	44

Table 3: Manipulable variables Petlyuk column LCB.

IP	Manipulable variables			
	D	R	VR	LR
1	77.18	16.74	126.07	150.00
5	79.37	16.47	131.18	150.00
10	80.95	17.64	136.92	150.00
20	80.95	17.64	123.92	129.88
30	78.64	16.82	132.23	150.00

Table 4: Results BO LCB.

IP	Description			TOC x 10^6	Ω	Iter.
	XD_A	XS_B	XB_C			
1	0.999	0.957	0.962	4.96	-2.475	48
5	0.999	0.965	0.970	5.03	-2.500	45
10	0.999	0.968	0.987	5.47	-2.509	8
20	0.999	0.968	0.987	5.47	-2.509	8
30	0.999	0.964	0.986	5.06	-2.520	78

Table 5: Manipulable variables Petlyuk column PI.

IP	Manipulable variables			
	D	R	VR	LR
1	78.82	17.38	126.00	150.00
5	79.65	16.52	131.18	150.00
10	78.04	17.59	136.92	150.00
20	77.45	17.03	123.92	129.88
30	78.78	16.27	132.33	150.00

Table 6: Results BO PI.

IP	Description			TOC x 10^6	Ω	Iter.
	XD_A	XS_B	XB_C			
1	0.999	0.964	0.968	5.26	-2.479	42
5	0.999	0.964	0.968	5.06	-2.493	36
10	0.999	0.965	0.980	5.26	-2.505	52
20	0.999	0.962	0.962	5.06	-2.476	42
30	0.999	0.961	0.965	4.93	-2.489	45

Finally, in the case of PI (tables 5-6), the performance shows the trend of EI, with achieved purities of 0.99 (XD_A), 0.96 (XS_B), and 0.96 (XB_C). The results are very similar, with the difference in the objective function between the cases of 1 and 10 initial points being just 0.001. In this manner, the economic aspect also shows only slight differences. Additionally, when comparing the different acquisition functions, the differences are minimal.

Similarly to the previous results, when noise is added to the objective function, as shown in Tables 7–12, the trend of achieving a high-purity distillate product (XD_A = 0.999) continues, except for an outlier in the LCB acquisition function with 30 IPs. Additionally, both XS_B and XB_C maintain purities of ≥0.95, showing no significant difference compared to the noise-free case.

On the economic side, TOC results remain in the range of 4.63 to 6.85 × 10^6 USD, with an exception for an outlier in the EI acquisition function with 1 IP. Thus, no improvements are observed in relation to any acquisition function, and increasing the number of IPs does not lead to better performance in the objective functions. It is important to highlight, as in the previous case, that adding noise does not negatively impact the BO process, and good results can still be achieved even with just one IP. Moreover, when comparing the different acquisition functions with and without uncertainty, the differences between the results are minimal. Therefore, the number of IPs providing quality information is not necessarily associated with better objective function performance, regardless of the acquisition function. Additionally, the multi-objective function ensures that the trade-off between purities and TOC remains balanced. However, as mentioned before, although TOC stays within a narrow range, its scale (10^6 USD) means that even a slight change, especially in the reflux ratio (R), can significantly impact TOC, either increasing or decreasing it by a considerable amount.

Table 7: Manipulable variables with noise Petlyuk column EI.

IP	Manipulable variables			
	D	R	VR	LR
1	80.74	22.32	150.00	150.00
5	77.65	19.63	150.00	150.00
10	75.79	19.83	110.25	115.46
20	79.40	18.28	150.00	150.00
30	79.62	18.63	150.00	150.00

Table 8: Results BO EI with noise.

IP	Description			TOC x 10^6	Γ	iteration
	XD_A	XS_B	XB_C			
1	0.999	0.972	0.975	6.85	-2.503	34
5	0.999	0.962	0.966	5.82	-2.521	51
10	0.999	0.957	0.962	5.73	-2.519	21
20	0.999	0.966	0.970	5.56	-2.583	59
30	0.999	0.967	0.971	5.68	-2.550	73

Table 9: Manipulable variables with noise Petlyuk column LCB.

IP	Manipulable variables			
	D	R	VR	LR
1	81.75	18.75	150.00	150.00
5	77.29	19.61	150.00	150.00
10	79.12	16.51	114.92	150.00
20	73.25	21.17	90.45	20.00
30	79.99	16.28	118.43	150.00

Table 10: Results BO LCB with noise.

IP	Description			TOC x 10^6	Γ	Iter.
	XD_A	XS_B	XB_C			
1	0.999	0.969	0.976	5.90	-2.570	34
5	0.999	0.961	0.965	5.78	-2.535	50
10	0.999	0.965	0.964	5.03	-2.612	46
20	0.999	0.949	0.955	5.90	-2.490	20
30	0.999	0.964	0.969	5.02	-2.579	51

Table 11: Manipulable variables with noise Petlyuk column PI.

IP	Manipulable variables			
	D	R	VR	LR
1	79.41	19.41	150.00	150.00
5	76.25	16.41	149.13	150.00
10	76.83	18.07	150.00	150.00
20	76.70	15.67	130.34	150.00
30	77.82	17.27	132.73	150.00

Table 12: Results BO PI with noise.

IP	Description			TOC x 10^6	Γ	Iter.
	XD_A	XS_B	XB_C			
1	0.999	0.967	0.971	5.89	-2.550	9
5	0.999	0.946	0.952	4.81	-2.534	17
10	0.999	0.958	0.963	5.32	-2.550	21
20	0.999	0.945	0.951	4.63	-2.570	51
30	0.999	0.965	0.963	5.16	-2.572	51

CONCLUSIONS

In this work, we present a practical methodology for solving multi-objective problems. The results obtained demonstrate a favorable trade-off between objectives for the Petlyuk distillation column, which serves as an energy-saving alternative to traditional sequences. This approach allows for analyzing the feasibility of the process without specifying a particular conversion fraction, as is typically required in distillation optimization problems. The solutions indicate that the achieved trade-off demonstrates the possibility of equilibrium between cost and product purity, with all cases achieving high purity fractions (≥0.96). Furthermore, as mentioned earlier, in low-dimensional Bayesian optimization problems like the one presented, there is no significant difference in performance based on the number of initial points or the acquisition function used. Additionally, despite variations in the number of initial points (IPs), the acquisition function, and the influence of noise in the process, no notable differences were observed, and all results remained within the same range of manipulable variable settings. However, even a slight change in the normalized cost represents a relatively significant change in the total operational cost (TOC) due to its magnitude (10^6 USD/year). While these differences are noticeable, increasing the number of initial points does not necessarily lead to better performance in the objective function. This work highlights the potential of Bayesian Optimization (BO) for process simulation, demonstrating that, when combined with modern computing power, it can achieve an optimal design in just a few fast iterations unlike traditional techniques that rely on explicit mathematical models, which often require simplifications. Thus, regardless of the number of initial points or the acquisition function selected, in cases with a low number of manipulated variables and system uncertainties, a multi-objective function can still provide a good trade-off for the Petlyuk column. Moreover, the procedure is easily extendable to other problems, not only those that are complex due to model structure but also large-scale or stochastic problems where computational costs are significant. These types of approaches will be explored in future work.

REFERENCES

1. Gutiérrez-Antonio C, Briones-Ramírez A. Pareto front of ideal Petlyuk sequences using a multiobjective genetic algorithm with constraints. *Computers & Chemical Engineering* 33(2):454–464 (2009). https://doi.org/10.1016/j.compchemeng.2008.11.004
2. Moriconi R, Deisenroth MP, Sesh Kumar K. High-dimensional Bayesian optimization using low-dimensional feature spaces. *Machine Learning* 109:1925–1943 (2020).
3. Premkumar R, Rangaiah G. Retrofitting conventional column systems to dividing-wall columns. *Chemical Engineering Research and Design* 87(1):47–60 (2009). https://doi.org/10.1016/j.cherd.2008.06.013
4. Pérez-Ones O, Flores-Tlacuahuac A. A stochastic data-driven Bayesian optimization approach for intensified ethanol–water separation systems. *Chemical Engineering and Processing - Process Intensification* 197:109708 (2024). https://doi.org/10.1016/j.cep.2024.109708
5. Ramirez-Corona N, Jiménez-Gutiérrez A, Castro-Agüero A, Rico-Ramírez V. Optimum design of Petlyuk and divided-wall distillation systems using a shortcut model. *Chemical Engineering Research and Design* 88(10):1405–1418 (2010).
6. Scikit-optimize. Sequential model-based optimization. scikit-optimize/scikit-optimize. https://github.com/scikit-optimize/scikit-optimize (2020).
7. Siivola E, Paleyes A, Gonzalez J, Vehtari A. Good practices for Bayesian optimization of high-dimensional structured spaces. *Applied AI Letters* (2021). https://doi.org/10.1002/ail2.24
8. Singh D, Singh B. Investigating the impact of data normalization on classification performance. *Applied Soft Computing* 97:105524 (2020). https://doi.org/10.1016/j.asoc.2019.105524
9. Velasco L, Guerrero H, Hospitaler A. A literature review and critical analysis of metaheuristics recently developed. *Archives of Computational Methods in Engineering* 31(1):125–146 (2024).
10. Ye H, Gao W, Guang Dong H, Bi M. An inherently safer development approach for thermally coupled distillation sequences: Application in hazardous chemical separation. *Process Safety and Environmental Protection* 160:786–802 (2022). https://doi.org/10.1016/j.psep.2022.02.044