12th IFAC International Symposium on
Advanced Control of Chemical Processes
July 14-17, 2024. Toronto, Canada

Human-in-the-loop controller tuning using Preferential Bayesian Optimization

João P. L. Coutinho * Ivan Castillo ** Marco S. Reis *

* University of Coimbra, CERES, Department of Chemical Engineering, 3030-790 Coimbra, Portugal
** Chemometrics, AI and Statistics, Technical Expertise and Support, The Dow Chemical Company, Lake Jackson, Texas 77566, United States

Abstract: The development of human-centric platforms that are able to combine computational resources and advanced analytics with human judgment and qualitative processing ability is a key driver of the Industry 5.0 movement. In this setting, humans are not only active in the loop but also play a key role in the decision-making process. In this work, we propose the use of Preferential Bayesian Optimization (PBO) for human-in-the-loop controller tuning. PBO relies on pairwise comparisons and preference feedback (A is better than B) to search for the optimal trade-off between different performance criteria from the user's perspective. The advantages of PBO are demonstrated in a simulated Proportional Integral (PI) controller tuning example with real user feedback under a reduced number of experiments. The results show that PBO leads to a greater emphasis on closed-loop responses closer to the user's desired behavior when compared with multi-objective alternatives, while being straightforward to implement from the user's perspective.

Keywords: Bayesian Optimization, Preference Learning, Multi-objective optimization, PID tuning

1. INTRODUCTION

The design of high performance controllers for industrial process control applications requires tuning of their parameters. This task is often carried out manually by an operator or plant engineer, which uses system identification and trial-and-error tuning to select suitable controller parameters. The efficiency of this approach largely depends on the user knowledge and may lead to high economic or engineering costs, especially for more complex control structures.

Controller calibration can be automated by using data-driven optimization methods that iteratively optimize the controller parameters based on only closed-loop data. One such method is Bayesian Optimization (BO) (Shahriari et al., 2016; Frazier, 2018), which has been receiving increased interest from the process engineering community for automatic tuning of arbitrary control structures (Piga et al., 2019; Neumann-Brosig et al., 2020; Khosravi et al., 2022a,b; Makrygiorgos et al., 2022; Paulson et al., 2022; Coutinho et al., 2023). BO has several desirable characteristics including experimental efficiency, possibility to account for noisy observations, the flexibility to incorporate constraints, possibility of experiment parallelization and integration of multiple information sources (Paulson et al., 2023). BO relies on the specification of a quantitative objective function calculated based on experimental data, which is assumed to reflect the desired performance. In cases where there are multiple conflicting criteria, Multi-Objective BO (MOBO) methods (Daulton et al., 2020) can be used to approximate the Pareto front that determines the trade-off between objectives. However, in some cases it can be hard to specify different quantitative performance criteria or decide between the importance of each objective, especially when these are qualitative in nature. On the other hand, it is usually rather straightforward for a user to choose the preferred option from different closed-loop responses. This preference feedback can be used to learn the underlying utility function which dictates human preferences, in what is known in machine learning literature as preference learning (Chu and Ghahramani, 2005). In the scope of controller design, such preference-based optimization methods have been recently proposed for calibration of Model Predictive Controllers (MPC) (Zhu et al., 2021, 2022) and human-collaborative robots (Roveda et al., 2023).

Preferential BO (PBO) methods expand the BO framework to account for user preferences (Brochu et al., 2010; González et al., 2017; Nguyen et al., 2021; Lin et al., 2022; Astudillo et al., 2023). In PBO, a Bayesian surrogate model is used to approximate the Decision Maker (DM) utility function based on his responses from a series of queries, each one consisting of two or more different configurations. An auxiliary acquisition function uses the posterior distribution of the utility surrogate model to propose new queries which are expected to be informative with respect to the global optimum by combining exploration and exploitation.

⋆ João Coutinho and Marco Reis acknowledge support from the Fundação para a Ciência e Tecnologia, I.P., through the projects with DOI:10.54499/UIDB/00102/2020 and DOI:10.54499/UIDP/00102/2020.

In this work, we adopted PBO to develop a human-in-the-loop alternative approach to automatic controller calibration methods, which entirely bypass human intervention during the closed-loop optimization. The advantages of the proposed methodology are demonstrated in a simulated Proportional Integral (PI) controller tuning example with conflicting performance specifications. We consider that PBO offers great potential to join human-centered decision making with automated machine learning methods, a key direction in the pursuit of Artificial Intelligence for operation of chemical processes and Industry 5.0 (Chiang et al., 2022).

The remainder of this paper is structured as follows. Section 2 presents the problem description. In Section 3, we introduce the use of PBO for preference-based tuning. Section 4 presents the results considering a simulated example and a real user. Finally, Section 5 summarizes the conclusions of this work and indicates opportunities for future work.

2. PROBLEM DESCRIPTION

Lets us consider the following dynamical system [1]

$$y_k = g(y_{k-1}, u_{k-1}), \; k = 1, 2, ..., K, \tag{1}$$

where y is the state variable, u is the control input and g is the, possibly non-linear, state transition function. For the sake of simplicity, we do not consider at this point external disturbances or noise. At each sampling time, k, the control input is given by a Proportional Integral (PI) controller in the velocity form:

$$u_k = u_{k-1} + K_c \left[e_k - e_{k-1} + \frac{\Delta t}{\tau_I} e_k \right], \tag{2}$$

where Δt is the sampling time and $e = y - y_{sp}$ is the error, which quantifies the deviation from the setpoint, y_{sp}. In optimization-based tuning, the goal is to find the controller parameters $x = [K_c, \tau_I]$, that maximize an objective function related to closed-loop performance. Controller tuning generally involves balancing conflicting criteria such as setpoint tracking and disturbance rejection or a fast closed-loop response and smooth control action. Thus, we may have a n-dimensional vector, $\mathbf{J}(x)$, of individual performance objectives, J_i:

$$\max_{x \in \mathcal{X}} \mathbf{J}(x) = [J_1(x), J_2(x), ..., J_n(x)], \tag{3}$$

where $\mathcal{X} \in \mathbb{R}^D$ is a bounded design space with dimensionality D. We assume that each performance objective, J_i, can generally be seen as a black-box, with unknown structural form or mathematical relationship with the controller parameters. Moreover, when considering a physical experiment or complex simulation, each evaluation is expensive or time-consuming, motivating the need for data-efficient optimization methods such as BO.

A common approach to solve optimization problem (3) is to use standard BO and solve a single objective optimization problem by considering a weighted sum of the individual objectives, properly scaled, assigning higher weights to those that are considered more important. However, it is not trivial to specify weights that lead to the desired performance beforehand. This can lead to further rounds of experiments if the results are not satisfactory.

Alternatively, one could use methods that aim to discover the set of Pareto optimal parameters, i.e. Pareto front, using multi-objective BO (MOBO) (Daulton et al., 2020). The user then selects the parameters from the observed or approximated Pareto set that best represents the desired trade-off between the different performance criteria. However, since each user is usually only interested in a specific region of the Pareto front, a portion of experiments may lead to undesirable results from the user's perspective. Moreover, it can be hard to choose a specific option from this set, especially when there are more than 3 objectives, making the Pareto front difficult to visualize.

In most situations, users ultimately rely on visual inspection of the closed-loop dynamic behavior to select the optimal controllers parameters. While it can be hard to make this decision solely based on quantitative performance metrics, the user can very easily decide if controller A leads to a better dynamic response than controller B. Such an example is presented in Fig. 1. In the next section, we explain how such comparisons can be used to optimize controller parameters through PBO.

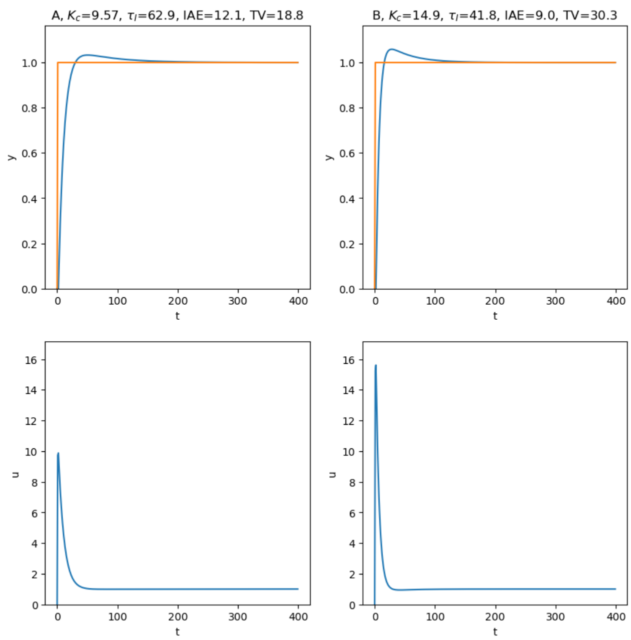

Fig. 1. Example of a query in preference-based tuning. In this comparison, a user concerned with controller smoothness, prefers option A (on the left) over B (on the right) due to lower controller action with similar deviation from the setpoint. Two metrics, IAE and TV, which are later introduced, are also shown, but not required, to illustrate the differences in quantitative performance.

3. CONTROLLER TUNING USING PREFERENTIAL BAYESIAN OPTIMIZATION

In preference-based tuning the goal is to maximize a utility function, $f(x)$, that reflects the DM's desired performance as a function of controller parameters, x:

[1] We note that the proposed methodology is applicable for arbitrary types of processes and controllers.

$$\max_{x \in \mathcal{X}} f(x). \tag{4}$$

It is assumed that the DM possess sufficient domain expertise to decide on the best closed-loop responses. However, since human feedback is not always consistent with their underlying preferences, we consider noise in the DM's responses by assuming $f(x)$ is corrupted by zero mean Gaussian noise, $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$. Following the work of Chu and Ghahramani (2005), we consider a Gaussian Process (GP) prior (Rasmussen and Williams, 2006), with mean $\mu(x)$ and covariance function $K(x, x')$, on the utility function,

$$f(x) \sim \mathcal{GP}(\mu(x), K(x, x')). \tag{5}$$

Here, we consider a constant mean and a squared exponential covariance function with Automatic Relevance Determination (ARD):

$$K(x, x') = \exp\left(-\frac{1}{2}(x - x')^T L (x - x')\right), \tag{6}$$

where $L = \mathrm{diag}(l^{-2})$ is a diagonal matrix with separate lengthscales, $l_1, .., l_D$, for each input dimension.

In equation 5, $f(x)$ is a latent unobservable function, which means that we do not have access to direct evaluations. Instead, we can learn the utility function by proposing different configurations and querying the DM to give his feedback in terms of the preferred option. Although this framework can be extended to multiple options per query (Nguyen et al., 2021; Astudillo et al., 2023), we limit the discussion to pairwise comparisons.

Consider that we have previously carried out a set of n experiments $X_n = \{x_i : 1, 2, ..., n\}$ and m pairwise comparisons, such as:

$$D_m = \{r_k(x_1^k, x_2^k) : k = 1, ..., m\}, \tag{7}$$

where x_1 and x_2 are two distinct elements of X_n and $r(x_1, x_2)$ is as a binary variable that takes the value of 1 when x_1 is preferred over x_2 and 0 otherwise. This can be interpreted as $f(x_1) \geq f(x_2)$, or in other terms, the DM considers the utility of x_1 to be larger than x_2. This preference response can be captured in the form of a probit likelihood function (Chu and Ghahramani, 2005):

$$\mathbb{P}(r(x_1, x_2) = 1 \mid f(x_1), f(x_2)) = \Phi\left(\frac{f(x_1) - f(x_2)}{\sqrt{2}\sigma_n}\right), \tag{8}$$

where Φ is the Gaussian CDF. Because the likelihood is non-Gaussian, the posterior distribution is analytically intractable. Therefore, we use Laplace approximation to approximate the posterior distribution as a Gaussian. Under this approximation, the marginal log likelihood has a closed form and can be maximized to find the GP model hyper-parameters, i.e., the mean, kernel lengthscales and the noise level (Chu and Ghahramani, 2005).

Several acquisition functions have been proposed for preference learning using Bayesian Optimization (Brochu et al., 2010; González et al., 2017; Nguyen et al., 2021). In this work we consider the recently proposed acquisition function Expected Utility of the Best Option (EUBO) (Lin et al., 2022; Astudillo et al., 2023), due to its computational simplicity and superior empirical performance. EUBO considers the expected maximum utility of a query, under the predicted posterior distribution, with the current data, D_m:

$$\alpha(x_1, x_2; D_m) = \mathbb{E}_m[\max\{f(x_1), f(x_2)\}]. \tag{9}$$

For pairwise comparisons, EUBO can be computed analytically (Lin et al., 2022) as:

$$\begin{aligned}\alpha(x_1, x_2; D) =& \Delta(x_1, x_2)\Phi\left(\frac{\Delta(x_1, x_2)}{\sigma(x_1, x_2)}\right) + \\ & \sigma(x_1, x_2)\phi\left(\frac{\Delta(x_1, x_2)}{\sigma(x_1, x_2)}\right) + \mu_m(x_2),\end{aligned} \tag{10}$$

where ϕ is the Gaussian PDF. Considering the GP mean and covariance functions with the current data as μ_m and K_m, respectively, the mean, $\Delta(x_1, x_2)$, and variance, $\sigma^2(x_1, x_2)$, of $f(x_1) - f(x_2)$, are given by:

$$\Delta(x_1, x_2) = \mathbb{E}[f(x_1) - f(x_2)] = \mu_m(x_1) - \mu_m(x_2) \tag{11}$$

$$\begin{aligned}\sigma(x_1, x_2) &= \mathbb{V}[f(x_1) - f(x_2)] \\ &= K_m(x_1, x_1) + K_m(x_2, x_2) - 2K_m(x_1, x_2)\end{aligned} \tag{12}$$

Algorithm 1 Preferential Bayesian Optimization algorithm

Require: Domain $\mathcal{X}$, initial experiments $X_0 = \{x_i\}_{i=1}^{N_0}$ and comparisons $D_0 = \{r_k(x_1^k, x_2^k)\}_{k=1}^{m_0}$
for $n = 1, ..., m_{max}$ **do**
Use current dataset D_{n-1} to estimate the GP model hyper-parameters
Find the parameters that maximize the acquisition function: $(x_1^n, x_2^n) = \arg\max_{x_1, x_2 \in \mathcal{X}} \alpha(x_1, x_2; D_{n-1})$
Perform an experiment at x_1^n and x_2^n and query the user to obtain preference feedback, $r_n(x_1^n, x_2^n)$
Augment the dataset: $D_n = \{D_{n-1}, ((x_1^n, x_2^n), r_n)\}$
end for
Return: Optimal parameters x^*, selected by the user or based on the utility model posterior distribution.

The PBO algorithm is presented in Algorithm 1. Initially, a set of N_0 experiments are performed based on a Sobol space-filling design. From this set, m_0 pairwise comparisons, randomly selected from all possible combinations, are presented to the DM to obtain preference feedback. Afterwards, in each iteration of the PBO loop, the GP hyper-parameters are estimated using the current dataset and based on the available posterior, the acquisition function is maximized to find a pair of controller parameters to evaluate for the next query. A comparison of the two closed-loop responses is presented to the DM, which then decides on the best response. The results of the experiments and comparison are added to the current dataset and the procedure continues until the allowed budget of comparisons, m_{max}, is exhausted. In the preference learning setting, the best controller parameters are the ones that maximize the DM's expected utility based on all the collected data. These can be either selected manually by having the DM choose his preferred option from a subset of all the observed closed-loop responses, or by relying on the predictions of the final GP model.

4. RESULTS

Consider the following first order process:

$$y_k = 0.99y_{k-1} + 0.00995u_{k-1} \tag{13}$$

where a delay of 1 sampling time, taken as 1 s, is added to the process output. The controller bounds are selected as $K_c \in [1, 20]$ and $\tau_I \in [10, 300]$. If a process model is available, a more informed design space can be defined using controller tuning relations (Coutinho et al., 2023).

We compare PBO with two other approaches: i) random query selection by drawing two samples from a Sobol design and ii) Expected Hyper Volume Improvement (EHVI) (Daulton et al., 2020), a MOBO algorithm. The BO algorithms were implemented using BoTorch (Balandat et al., 2020). For EHVI, we select the simultaneous minimization of two objectives: Integrated Absolute Error (IAE) and Total Variation (TV), which quantify controller performance and smoothness of the controller output, respectively:

$$\mathrm{IAE} = \sum_{k=0}^{K} \mid y_{sp,k} - y_k \mid \Delta t, \ \mathrm{TV} = \sum_{k=0}^{K-1} \mid u_{k+1} - u_k \mid . \tag{14}$$

Validation of the proposed method is carried out considering an external subject matter expert as the DM, which provided the following statement about his desired closed-loop response: "A fast and stable response, with minimal overshoot and moderate controller action". Due to possibly noisy responses in terms of preference feedback, 5 separate trials were performed, each one using a different initial set of experiments. Due to space limitations, we will only present the detailed results for one trial.

Initially, each trial began with 30 comparisons, randomly selected from all possible combinations from a set of 10 initial experiments. After this stage, 10 iterations of the PBO algorithm were carried out, leading to a total of 40 comparisons and 30 experiments. To focus the methodology comparison on the selection of queries, the initial set of experiments is kept the same for all three methods in each trial. For both PBO and Sobol, the final optimal controller parameters are selected as the ones that maximize the predicted posterior mean of the final GP model based on all the observed data. For EHVI, three different points from the observed Pareto front at the final iteration, selected as displayed in Fig. 2, are presented to the DM.

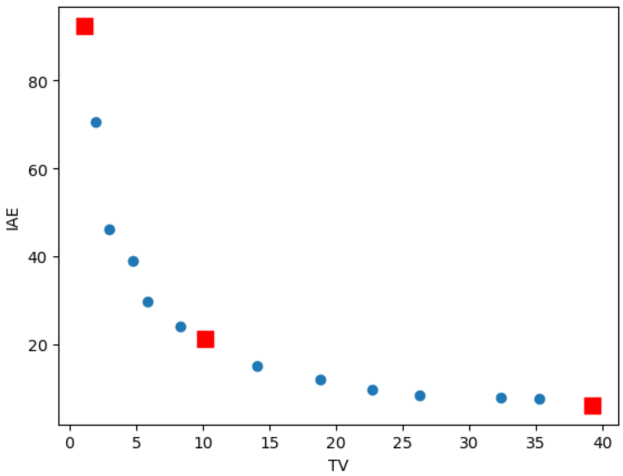

Fig. 2. Observed Pareto front for the two objectives. The squares represent the parameters used for the closed-loop experiments that are shown to the DM.

To perform a quantitative comparison of the three methods, the observed values of the two quantitative objectives considered for multi-objective optimization are presented in Fig. 3.

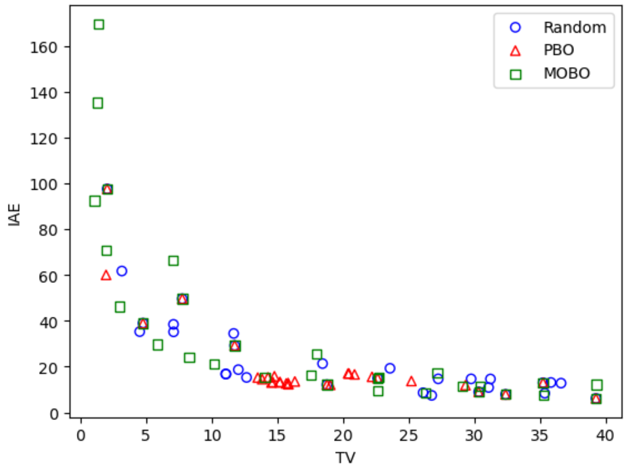

Fig. 3. Observed values of the two objectives, IAE and TV, during the experiments.

It can be observed that while MOBO leads to values that are spread across the entire range of the objective space, the majority of PBO experiments lead to values clustered in a relatively narrow region. This is expected, as the goal in standard MOBO is to discover the entire Pareto front. On the other hand, PBO concentrates experiments in regions that are of high expected utility to the DM. For instance, the region of high IAE is avoided, because it represents responses with large deviations from the setpoint, which is obviously undesired from the DM's perspective. This can be further visualized in the parameter space, shown in Fig. 4, where the algorithm avoids sampling in the region of large integral time and low controller gain.

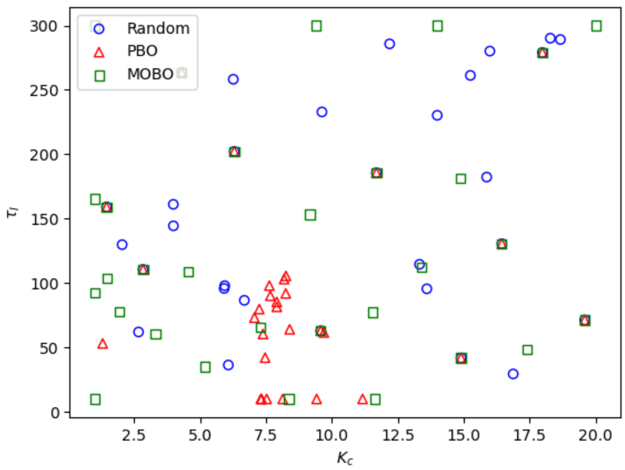

Fig. 4. Controller parameters evaluated during the experiments.

The clustering observed in Figs. 3 and 4 may also indicate that the PBO algorithm has converged to a good solution in terms of the DM's preferences before exhausting the experimental budget. Determining appropriate metrics for convergence in PBO remains to be explored in future work.

The final results of the five trials are presented in Table 1. As a final comparison in each trial, the closed loop responses corresponding to the optimal parameters for the three methods, as shown in Fig. 5, were presented to the DM, which stated his preferred option. Although this will obviously not be done in a practical scenario, this final query is used to compare the best results from the different methods.

Table 1. Optimal controller parameters for each trial. The DM's preferred option is shown in bold.

	PBO		MOBO		Random	
Trial	K_c	τ_I	K_c	τ_I	K_c	τ_I
1	**6.3**	**84.1**	7.7	100.0	3.3	85.8
2	6.0	58.4	**4.8**	**63.5**	3.2	44.5
3	**14.7**	**122.2**	6.5	94.8	14.1	68.9
4	5.7	101.6	4.6	68.2	**12.0**	**103.5**
5	**8.2**	**91.9**	5.2	35.2	9.6	62.9

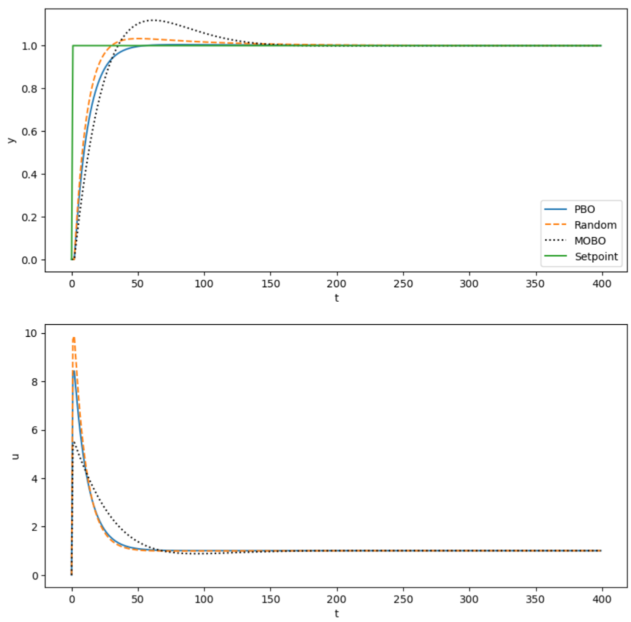

Fig. 5. Comparison of the closed-loop response corresponding to the optimal parameters of each method. The DM's decides which of the three is optimal.

The DM chose the PBO solution in three of five trials, indicating that good controller performance, according to specific user preferences and criteria, can be achieved through pairwise comparisons and PBO, without having the need to specify quantitative performance metrics.

4.1 Sensitivity Analysis

The obtainable performance of the controllers depends on the number of initial experiments, comparisons and total amount of PBO iterations. To determine the impact of having a reduced dataset, we repeat the analysis using half the previous experimental budget, with 10 initial comparisons based on 5 space-filling experiments and using only 5 PBO iterations, for a total of 15 experiments. Table 2 shows the resulting controller parameters for the three methods, with the corresponding closed-loop responses in Fig. 6.

Table 2. Optimal controller parameters for each trial considering 15 total experiments. The DM's preferred option is shown in bold.

	PBO		MOBO		Random	
Trial	K_c	τ_I	K_c	τ_I	K_c	τ_I
1	**11.1**	**67.5**	6.3	90.9	12.0	20.8
2	**9.3**	**85.1**	4.7	68.0	19.0	79.5
3	**9.9**	**74.3**	4.4	87.9	6.8	39.9
4	**9.6**	**93.4**	3.9	73.3	9.1	22.4
5	**9.4**	**114.0**	9.6	62.9	9.6	62.9

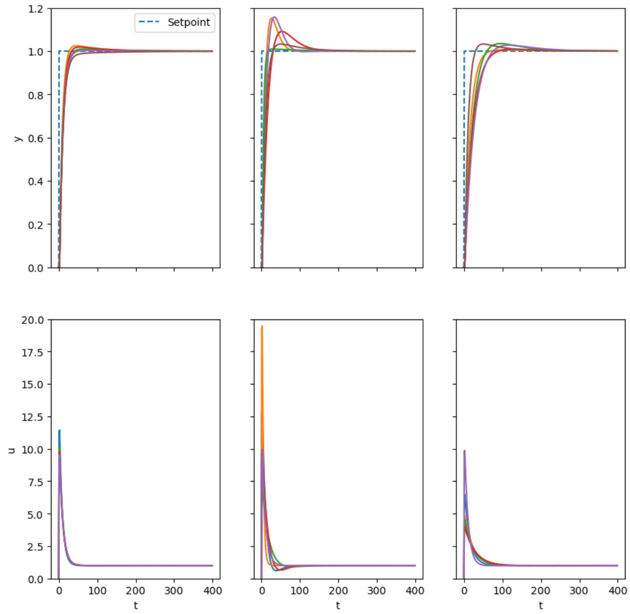

Fig. 6. Best closed-loop responses using a reduced number of experiments across the 5 trials for PBO (left), random sampling (center) and MOBO (right).

It can be observed that the best closed-loop responses obtained using PBO are more consistent across the different trials when compared with the other two methods. In this case, the DM chose the PBO solution as the best one in all five trials. Overall, this indicates that PBO shows advantages compared to the other methods even when using a reduced number of experiments.

5. CONCLUSION

We presented the use of Preferential Bayesian Optimization (PBO) for human-in-the-loop controller tuning. Instead of optimizing a quantitative objective function, PBO suggests pairwise comparisons of different closed-responses and asks a human calibrator which is the preferred option. As demonstrated in a simulated controller tuning example, this leads to a larger focus of experiments on the desired trade-offs between different objectives, when compared with standard multi-objective BO.

Future work will consider further validation of the approach by considering multiple users with different preferences and also incorporate ties when there is no preference towards either option (Nguyen et al., 2021). This would enable a more thorough comparison with other active preference learning approaches (Zhu et al., 2021). It is also straightforward to apply PBO for other more complex controller calibration problems, including MPC tuning, although this may involve a larger number of decision variables.

ACKNOWLEDGEMENTS

The authors acknowledge the contribution of Professor Lino Santos as the decision-maker during the experiments.

REFERENCES

Astudillo, R., Lin, Z.J., Bakshy, E., and Frazier, P. (2023). qeubo: A decision-theoretic acquisition function for preferential bayesian optimization. In F. Ruiz, J. Dy, and J.W. van de Meent (eds.), *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, 1093–1114. PMLR.

Balandat, M., Karrer, B., Jiang, D., Daulton, S., Letham, B., Wilson, A.G., and Bakshy, E. (2020). BoTorch: A framework for efficient monte-carlo bayesian optimization. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, 21524–21538. Curran Associates, Inc.

Brochu, E., Brochu, T., and de Freitas, N. (2010). A bayesian interactive optimization approach to procedural animation design. In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '10, 103–112. Eurographics Association, Goslar, DEU.

Chiang, L.H., Braun, B., Wang, Z., and Castillo, I. (2022). Towards artificial intelligence at scale in the chemical industry. *AIChE Journal*, 68(6), e17644. doi: 10.1002/aic.17644.

Chu, W. and Ghahramani, Z. (2005). Preference learning with Gaussian processes. In *Proceedings of the 22nd International Conference on Machine Learning - ICML '05*, 137–144. ACM Press, Bonn, Germany. doi: 10.1145/1102351.1102369.

Coutinho, J.P., Santos, L.O., and Reis, M.S. (2023). Bayesian Optimization for automatic tuning of digital multi-loop PID controllers. *Computers & Chemical Engineering*, 173, 108211. doi: 10.1016/j.compchemeng.2023.108211.

Daulton, S., Balandat, M., and Bakshy, E. (2020). Differentiable expected hypervolume improvement for parallel multi-objective bayesian optimization. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, 9851–9864. Curran Associates, Inc.

Frazier, P.I. (2018). A tutorial on bayesian optimization. *arXiv e-prints*, arXiv–1807.

González, J., Dai, Z., Damianou, A., and Lawrence, N.D. (2017). Preferential Bayesian Optimization. In *International Conference on Machine Learning*, 1282–1291. PLMR.

Khosravi, M., Behrunani, V., Myszkorowski, P., Smith, R.S., Rupenyan, A., and Lygeros, J. (2022a). Performance-Driven Cascade Controller Tuning with Bayesian Optimization. *IEEE Transactions on Industrial Electronics*, 69(1), 1032–1042. doi: 10.1109/TIE.2021.3050356.

Khosravi, M., Koenig, C., Maier, M., Smith, R.S., Lygeros, J., and Rupenyan, A. (2022b). Safety-Aware Cascade Controller Tuning Using Constrained Bayesian Optimization. *IEEE Transactions on Industrial Electronics*, 1–1. doi:10.1109/TIE.2022.3158007.

Lin, Z.J., Astudillo, R., Frazier, P., and Bakshy, E. (2022). Preference exploration for efficient bayesian optimization with multiple outcomes. In *International Conference on Artificial Intelligence and Statistics*, 4235–4258. PMLR.

Makrygiorgos, G., Bonzanini, A.D., Miller, V., and Mesbah, A. (2022). Performance-oriented model learning for control via multi-objective Bayesian optimization. *Computers & Chemical Engineering*, 162, 107770. doi: 10.1016/j.compchemeng.2022.107770.

Neumann-Brosig, M., Marco, A., Schwarzmann, D., and Trimpe, S. (2020). Data-Efficient Autotuning With Bayesian Optimization: An Industrial Control Study. *IEEE Transactions on Control Systems Technology*, 28(3), 730–740. doi:10.1109/TCST.2018.2886159.

Nguyen, Q.P., Tay, S., Low, B.K.H., and Jaillet, P. (2021). Top-k ranking bayesian optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 9135–9143.

Paulson, J.A., Makrygiorgos, G., and Mesbah, A. (2022). Adversarially robust Bayesian optimization for efficient AUTO-TUNING of generic control structures under uncertainty. *AIChE Journal*, 68(6), e17591. doi: 10.1002/aic.17591.

Paulson, J.A., Sorourifar, F., and Mesbah, A. (2023). A Tutorial on Derivative-Free Policy Learning Methods for Interpretable Controller Representations. In *2023 American Control Conference (ACC)*, 1295–1306. IEEE, San Diego, CA, USA. doi: 10.23919/ACC55779.2023.10156412.

Piga, D., Forgione, M., Formentin, S., and Bemporad, A. (2019). Performance-Oriented Model Learning for Data-Driven MPC Design. *IEEE Control Systems Letters*, 3(3), 577–582. doi:10.1109/LCSYS.2019.2913347.

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, Cambridge.

Roveda, L., Mantovani, L., Maccarini, M., Braghin, F., and Piga, D. (2023). Optimal physical Human–Robot collaborative controller with user-centric tuning. *Control Engineering Practice*, 139, 105621. doi: 10.1016/j.conengprac.2023.105621.

Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., and de Freitas, N. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1), 148–175. doi: 10.1109/JPROC.2015.2494218.

Zhu, M., Bemporad, A., and Piga, D. (2021). Preference-based MPC calibration. In *2021 European Control Conference (ECC)*, 638–645. IEEE, Delft, Netherlands. doi:10.23919/ECC54610.2021.9654900.

Zhu, M., Piga, D., and Bemporad, A. (2022). C-GLISp: Preference-Based Global Optimization Under Unknown Constraints With Applications to Controller Calibration. *IEEE Transactions on Control Systems Technology*, 30(5), 2176–2187. doi:10.1109/TCST.2021.3136711.