Machine Learning and Derivative-Free Optimization for PID Tuning: Case Study of Improved Black Liquor Concentration Control

Mohamed El Koujok[1], Haitian Zhang[2], Hakim Ghezzaz[1], Mouloud Amazouz [1] and Ali Elkamel[2]

***Abstract*— Proportional-Integral-Derivative (PID) controllers are instrumental in managing industrial processes. Their effectiveness hinges on the precision of their tuned parameters. Consequently, it becomes essential to monitor their performance frequently and re-tune them regularly to ensure optimal operation. However, conventional tuning methods require significant effort and expertise. This research tackles this challenge by using historical data to construct machine learning (ML) models, coupled with optimization algorithms, to streamline PID tuning. Specifically, we propose integrating an Explainable Boosting Machine (EBM) as an ML model and harnessing Bayesian Optimization (BO) within a comprehensive PID tuning framework. EBM stands out for its ease of construction and accuracy. The synergistic combination of EBM and BO yields an effective solution, as demonstrated through a case study involving black liquor concentration control in a multiple-effect evaporator system within kraft pulp manufacturing.**

I. INTRODUCTION

Proportional-Integral-Derivative (PID) controllers, which constitute nearly 90% of all industrial process controllers, are renowned for their effectiveness, straightforward design, and ease of implementation in distributed control systems and programmable logic controller [1]. Consequently, PID controllers are regarded essential instruments for the regulation and control of various process variables, encompassing, but not limited to, temperature, pressure, flow rates and levels [2]. In the contemporary industrial landscape, PID controllers are a fundamental component, playing a pivotal role in optimizing process operation efficiency and reducing environmental impact. The key to achieving this lies in the proper tuning of the PID controllers. PIDs do not adapt to changes in the process over time, and their effectiveness tends to decrease over time due to various factors, such as equipment aging and feedstock variations. Consequently, regular tuning of these PIDs is essential to ensure their continued operation and maintain optimal performance.

Given the high number of PIDs involved in industrial processes, the tuning of PID controllers is a challenge for experts in the field, as the actual and traditional methods are based on several rules [3]. The best means of tuning a PID loop to achieve optimal performance is still an open question. According to [4], only one-third of PID controllers are operated with optimal efficacy, while another third function under suboptimal tuning conditions. The remaining third of these loops are managed manually. Despite the presence of various tuning methods and algorithms such as Ziegler-Nichols and ström-Hägglund [5], the tuning procedure often requires a combination of iterative trial and error in conjunction with the input of specialist process control expertise. Ultimately, tuning is often viewed to rely more on nuanced judgment than strict scientific formulas. This requires significant time and effort from the control engineers [6].

Moreover, a reduction in process control staff due to retirements and other factors has created an acute shortage of personnel proficient in tuning. Consequently, the primary factors contributing to the suboptimal tuning of numerous loops are knowledge and time constraints [7]. Poor PID tuning can undermine plant performance, leading to instability, oscillation, and overshooting [8]. These problems can escalate into excessive control activity, overcorrection, diminished product quality, and inefficient energy use [9]. Thus, suboptimal PID tuning is identified as the fundamental cause of various operational inefficiencies. In response to tuning challenges, developing effective automatic tuning procedures is undeniably crucial to achieving better-tuned controllers in the process industry. In addition, ongoing efforts should improve current methods to align them more closely with the evolving needs of the industry [10].

Recently, the Industrial Internet of Things (IIoT) revolution has opened a new era of process optimization. By providing unprecedented access to data, IIoT empowers the use of powerful machine learning (ML) tools for data-driven decision-making. Moreover, the exponential growth in computational power has made optimization algorithms not only practical but also significantly faster, enhancing their real-world applicability. This synergy of data, machine learning, and optimization algorithms has opened doors to various approaches for auto-tuning the PIDs. For example, the academic literature explores approaches for PID tuning, such as First-Order Plus Dead Time (FOPDT) models [11] and reinforcement learning (RL), an ML subdisciplinary domain [12]. However, FOPDT tuning requires bump tests, which may not be feasible in industrial settings due to cost and off-spec product risks. RL excels at learning, but its need for environmental interaction is challenging in the real world. Directly deploying an RL agent for PID auto-tuning in complex systems can be risky due to safety constraints. Creating a simulator to train an agent for PID parameter determination is possible, but even advanced simulations may fail to fully replicate real-world complexities, affecting the agent's performance in industrial applications.

To overcome the shortcomings of previous studies outlined above, this research presents a practical and applicable approach for the auto-tuning of PID controllers in industrial

[1] CanmetENERGY, EETS, Natural Resources Canada, Varennes, Canada `mohamed.elkoujok@nrcan-rncan.gc.ca`

[2]Department of Chemical Engineering, University of Waterloo, Waterloo, Canada `haitian.zhang@uwaterloo.ca`

processes, with a strong emphasis on real-world applicability. Our methodology adeptly avoids the requirement for both bump tests and the development of complex simulators that are laborious and time-consuming to construct. Our strategy utilizes readily available historical data from industrial plants, employing ML techniques and optimization algorithms to easily determine optimal PID parameters and simplify control engineers' tasks. In essence, our methodology offers a streamlined solution that not only simplifies the tuning process but also enhances the precision of PID controllers, thereby contributing significantly to the field of industrial process control. The following summarizes our key contributions:

- **Conceptual Framework**: A novel structure for PID controller auto-tuning, grounded on ML and optimization techniques, is introduced.
- **Algorithm Selection**: Explainable Boosting Machines (EBM) and Bayesian Optimization (BO) are adopted as the M model and optimization algorithm, respectively, within the proposed auto-tuning framework.
- **Performance evaluation**: A multiple effect evaporator simulator is used to generate data in a closed-loop mode, simulating the behavior of real-world processes. This data is used to train and assess the EBM and BO within the proposed framework. The capability of the methodology to exceed existing first-principle-based methods is demonstrated.

The following sections detail the inner workings of our proposed methodology for PID controller auto-tuning (Section 2). Section 3 then demonstrates its effectiveness through a practical case study. Finally, Section 4 concludes the paper by summarizing our key findings and exploring potential avenues for future research.

II. PROPOSED METHODOLOGY TO AUTOMATE THE TUNING OF PID

PID performance can deteriorate over time due to dynamic changes in the process. This deterioration requires frequent adjustments to update the PID parameters. Fortunately, historical data collected during the PID operation at various setpoints provides valuable information. Therefore, the methodology proposed here seeks to use these data to simplify and enhance the process of tuning and updating the PID parameters. Importantly, in this research, the parameters newly identified by our methodology are offered as recommendations to experts for the validation phase before being transmitted to the control system, as shown in Fig. 1.

Equation (1) illustrates the output of the PID controller ($\mathbf{u}(t)$) along with the parameters that influence the control function

$$\mathbf{u}(t) = K_c\mathbf{e}(t) + K_i\int \mathbf{e}(t)\mathrm{d}t + K_d\frac{\mathrm{d}\mathbf{e}(t)}{\mathrm{d}t} \tag{1}$$

where $\mathbf{e}(t)$ signifies the error value, mathematically expressed as $(S_p - y_p)$, with S_p denoting the setpoint and y_p representing the process output. K_c, $K_i = K_c/\tau_i$, and $K_d = K_c\tau_d$ denote the gain coefficients related to the proportional,

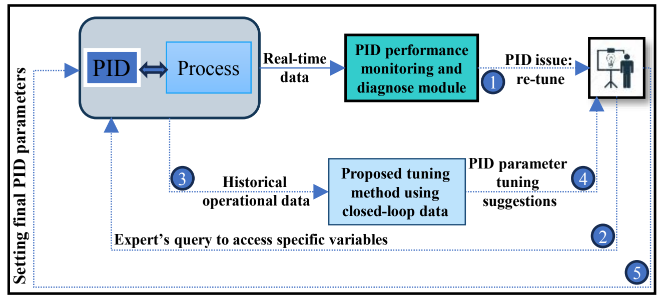

Fig. 1. PID parameter tuning utilizing historical data from closed-loop mode

integral, and derivative terms, respectively. The parameters denoted as τ_i and τ_d refer to the integral and derivative time, respectively. An optimally configured PID controller is proficient in regulating the process by providing $\mathbf{u}(t)$ to mitigate the error ($\mathbf{e}(t)$) and effectively reducing it to a value approaching zero.

As mentioned earlier, our methodology focuses on fine-tuning outdated PID parameters: K_c, τ_i, τ_d, where oscillations around the setpoint can indicate the condition of inadequate tuning. Subsequently, in order to identify the appropriate updated parameters K_c, τ_i, τ_d, our methodology comprises a three-step process (Fig. 2):

1) Collect historical process data, assuming it reflects suboptimal PID control and varied setpoint conditions. This dataset, rich in dynamic variations, enables the ML model to capture process behavior in a comprehensive way and effectively generalize across diverse operating scenarios.
2) Develop an ML surrogate model from the historical data. This model serves as a simulator for optimizing PID parameters, using the control variable $u(t)$ and the process variables $X = [X_1, \ldots, X_p]$ as inputs to predict the output $y_p(t)$.
3) Integrate the ML model with the PID controller in a closed-loop system. Simulate this ML-PID framework and apply a robust optimizer to iteratively compute PID parameters that minimize the optimization problem defined in the subsequent equation:

$$\begin{aligned} \min_{K_c,\tau_i,\tau_d} \ & (S_p - y_p)^2 \\ \text{s.t.} \ \ & y_p = f(u(t), X); f: ML\ model, \\ & L_1 < K_c < H_1, \\ & L_2 < \tau_i < H_2, \\ & L_3 < \tau_d < H_3, \\ & L_4 < u < H_4 \end{aligned} \tag{2}$$

In Equation 2, L_j and H_j $(j = 1,2,3)$ are the lower and higher bounds of K_c, τ_i and τ_d. The terms L_4 and H_4 are used to verify the lower and higher limits of the control signal u. If the solution u from the solver exceeds these limits, our methodology includes an anti-reset wind-up strategy to prevent integral wind-up. In line with our

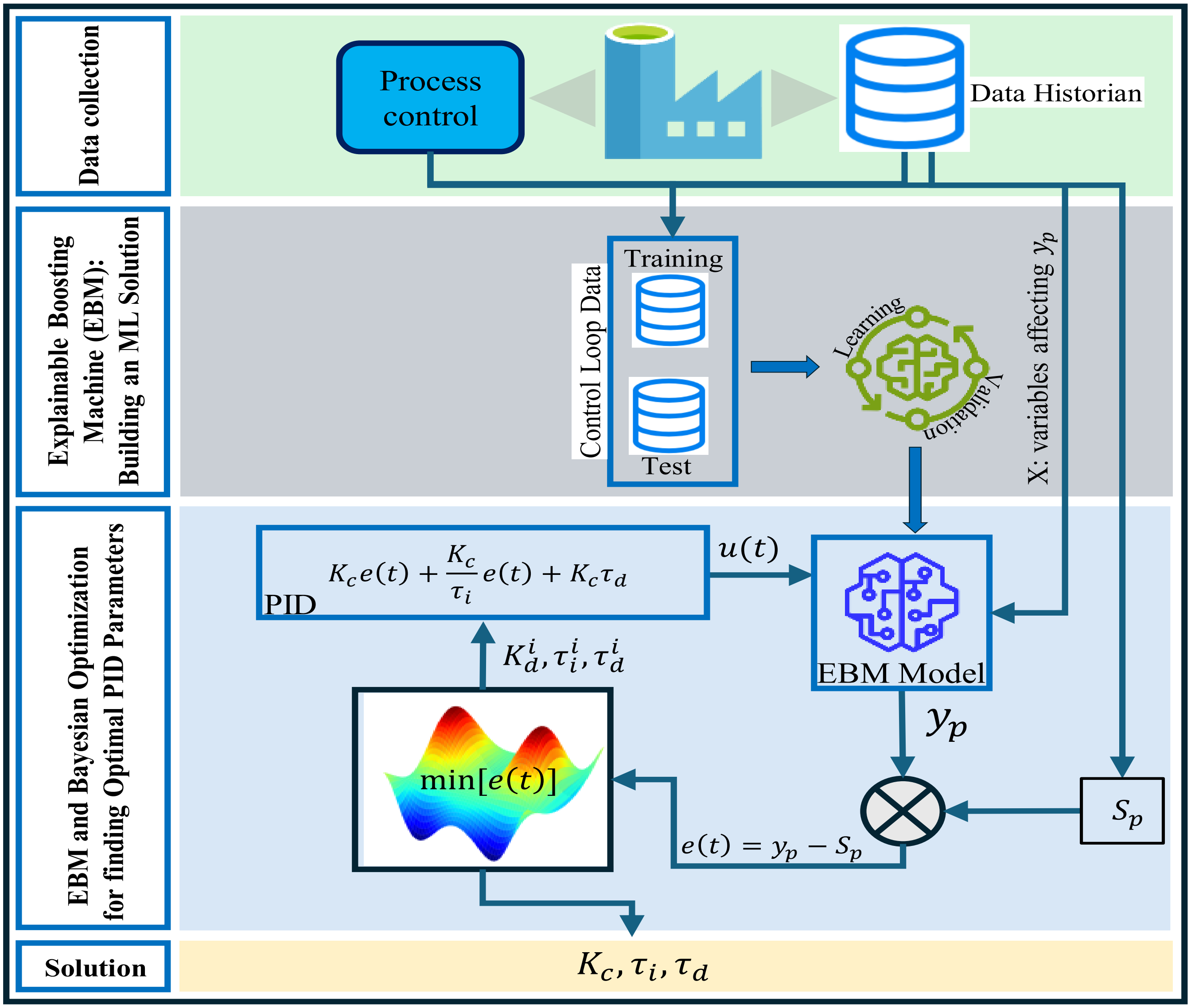

Fig. 2. Proposed methodology with its three steps

established methodology, it becomes important to select an appropriate ML model to act as a surrogate model. We have opted to adapt and implement the EBM within this framework. The EBM is a generalized additive model (GAM) based on trees with cyclic gradient boosting [13]. The EBM often shows accuracy levels similar to those of leading-edge black-box models, such as Random Forests (RFs) and Boosted Trees. Furthermore, the EBM's capability for rapid prediction renders it particularly appealing for deployment in real-world applications where swift decision-making is essential. In addition, its precision and competitiveness have been validated in various studies, particularly in metrics related to reliable prediction, stability, and fairness [14]. This makes the EBM a compelling choice as a surrogate model. Ultimately, this study selects the EBM as the surrogate model to be incorporated into the objective function, owing to its classification as a GAM, which guarantees that the associations between the independent predictors and the dependent variable display smooth patterns [13]. Consequently, this results in a smoother objective function for optimization purposes. This facilitates the optimizer's task of progressing towards the global optimum [15]. In contrast, neural networks, while effective as universal approximators, can generate complex response surfaces that risk trapping optimization algorithms in local optima [16].

Upon selecting the EBM as the surrogate model and in accordance with our proposed methodology, the next step involves identifying an optimizer to explore and determine the optimal PID parameters. The literature presents a variety of optimization algorithms, including conjugate gradient [17], Powell's conjugate direction method [18], bound optimization by quadratic approximation [19], particle swarm optimization [20], genetic algorithm [21] and BO [22]. In this study, we have chosen to utilize BO, given its sample efficiency and gradient-free nature, which makes it suitable for optimizing black-box functions such as the EBM to fine-tune the parameters of PIDs. Furthermore, BO is a robust and computationally efficient optimization technique that requires fewer function evaluations. It is capable of addressing a broad spectrum of optimization problems, managing noisy objective functions [22], and with a high probability of finding the global optimum of functions [23]. This makes it well-suited for complex real-world processes.

In summary, the PID tuning framework in this study, shown in Fig. 2, consists of three main steps: gathering closed-loop historical data, constructing an EBM model to represent the industrial process (in this research context, the EBM signifies the function f- see Equation 2), and incorporating this model into the optimization problem established in Equation 2. BO is then applied to resolve this structure and determine the optimal PID parameters (K_c, τ_i, τ_d).

III. Implementing the Methodology and Results

A. Process Description

The proposed methodology has been tested and validated using data generated with a dynamic simulation of a multiple-effect evaporator (MEEV) system. This system represents a major part of the chemical recovery cycle in kraft pulp mills [24]. Its role is to increase the content of black liquor dissolved solids produced in the pulping process of wood chips from approximately 13%-17% to about 50%. The system uses steam for the evaporation of black liquor water. This energy-intensive system is characterized by non-linear behaviors and changes in its dynamics. In fact, various factors can influence the dynamics of the system, such as the gradual buildup of fouling on the heat transfer surfaces, changes in system operating conditions, and the high variability of the physical properties of black liquor due to changes in wood species and pulping conditions. The parameters of the local PID controllers of this system should be tuned regularly to adapt to changes in system dynamics, and therefore maintain efficient control of the system.

The dynamic simulation of the MEEV system (see Fig. 3) was developed using the CADSIM Plus process simulation software. The model closely mimics the behavior of the real system and can simulate different operational scenarios, add noise, incorporate PIDs to control specific variables, and generate datasets.

The MEEV system selected for this study comprises six effects, as depicted in Fig. 3. To illustrate our proposed method, the adjustment of the parameters was demonstrated in the controller PID#1 (Fig. 3), the controller that regulates the concentration of black liquor solids (SCBL) after the first evaporator (Effect#1&Body#1, Fig. 3). It is imperative to control the SCBL, which constitutes the process variable (y_p), at a specified level prior to its arrival at the recovery boiler for steam generation. It should be noted that the same principle of our methodology can be applied to other PID controllers within the MEEV system.

B. Results and Discussion

A comprehensive dataset of 5300 samples was generated by performing a precise simulation of the MEEV model using the advanced capabilities of CADSIM. This dataset functions as historical data within the context of our case study. A rigorous one-minute sampling interval was implemented, allowing a high-resolution capture of dynamic process operations. Throughout this simulation, the process was governed by outdated PID#1 parameters ($K_{c0} = 150$, $\tau_{i0} = 2$ and $\tau_{d0} = 0$), resulting in significant oscillations around setpoints (refer to Fig. 4). This dataset encompasses multiple operational modes achieved by varying the value of the setpoint of the PID#1. Alongside the setpoint, this dataset includes four key variables: fresh steam flow (manipulated variable, u), SCBL (process variable, y_p), solid concentration (X_1) and black liquor flow (X_2). In particular, the measurements for the latter two variables (X_1 and X_2) are taken at the inlet of Effect1&Body1. The variables indicated as $\{[u, X_1, X_2]$ are chosen to serve as input, while y_p is designated as output to build the EBM-based surrogate model to effectively capture the dynamics of Concentrator#1 (see Fig. 3). The accuracy of the EBM was quantified using the recognized metric, the coefficient of determination (R^2). Following completion of the training phase, a strong R^2 value of 0.99 was achieved in the unseen validation data. This outcome demonstrates the ability of the EBM to capture the dynamics of the process and predict accurately the y_p. Following this, our suggested methodology utilized BO in conjunction with the EBM&PID#1 system to determine the optimal PID#1 parameters (K_c, τ_i, and τ_d). Adjusting these parameters are crucial in regulating the system by eradicating oscillations and adhering to the desired setpoints with minimal error. Within the scope of this optimization task, the constraints were defined as: $100 \leq K_c \leq 250$, $2 \leq \tau_i \leq 6$, and $0 \leq \tau_d \leq 2$. The determination of the constraint boundaries was informed by historical insights regarding previous PID#1 parameters in conjunction with a basic understanding of the dynamics of the process.

Table I presents the results of our methodology aimed at identifying the optimal parameters for PID#1, alongside the optimization iterations. The algorithm was terminated at the 7th iteration due to the absence of further improvements in the minimization of the cost function. The algorithm exhibited swift convergence toward the value that minimizes the cost function by the second iteration. Consequently, as delineated in Table I, the optimal parameters found for PID#1 are $K_c = 236.80$, $\tau_i = 3.53$, and $\tau_d = 1.96$. These parameters effectively minimize the cost function as defined in Equation 2. It is crucial to underscore the algorithm's expeditious convergence towards minimizing the cost function in just two iterations, which signifies its practical efficacy in real-world applications. As a result, the values of the PID#1 parameter obtained should replace the existing ones ($K_{c0} = 150$, $\tau_{i0} = 2$, and $\tau_{d0} = 0$). After substituting the previous values of PID#1 and running the controlled process, the result is illustrated in Fig. 4. As depicted in Fig. 4, the success of

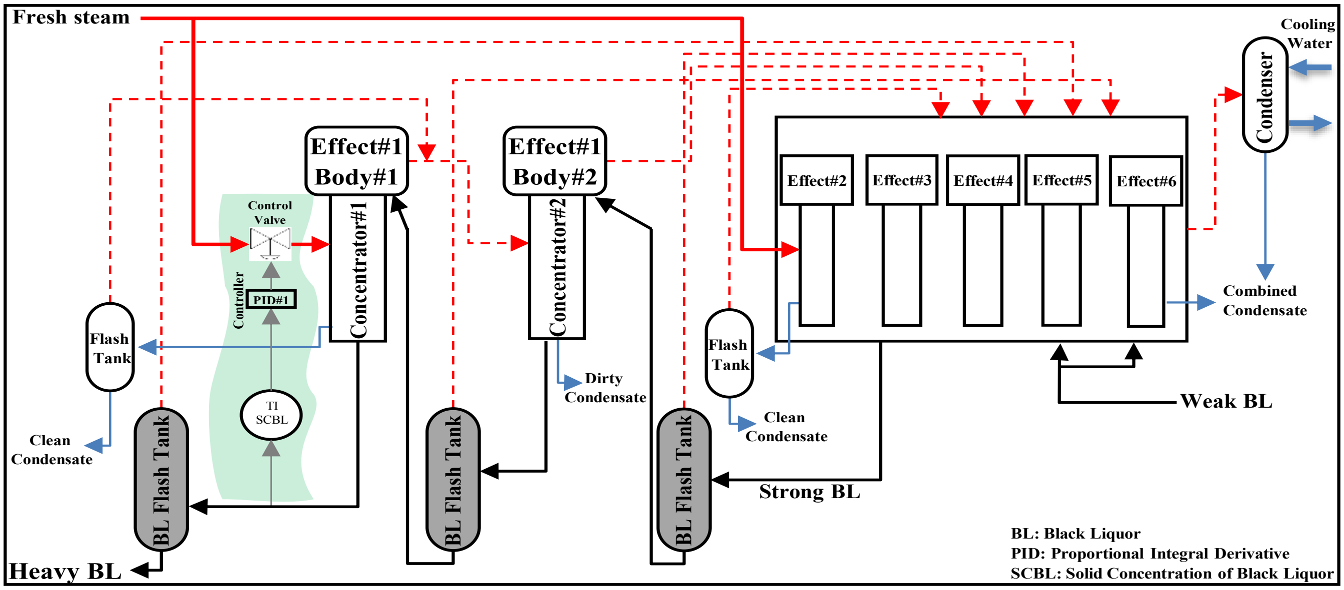

Fig. 3. Flowsheet of black liquor multitple-effect evaporators (MEEV)

our proposed method in recommending new parameter values for PID#1 is evidenced by the elimination of oscillations and the reduction of error between the process values and the setpoint. To further validate our methodology, a comparative analysis was performed between our results and those derived from a first-principles-based automated method. For this purpose, the Integral Square Error (ISE) metric was used, recognized as a standard indicator of the performance of the control system. Table II presents the comparative outcomes, illustrating that our data-driven methodology outperforms the conventional first-principles approach implemented within CADSIM by achieving a significantly reduced ISE value, thus underscoring its superior performance. It is especially significant that our methodology derives its power from an archive of historical data, which serves as a compelling proof of its remarkable adaptability. This is clearly illustrated in the complex context of the MEEV case study, where its success is undoubtedly evident and notable. Thus, with the availability of data collected from process control systems, our methodology reveals promising potential for broad and extensive applicability, indicating transformative advancements in PID controller tuning across a diverse and expansive range of industrial processes.

TABLE I
PARAMETERS OBTAINED BY BO

Iteration Number	K_c	τ_i	τ_d	Cost Function
0	192.57	5.26	1.47	8.85
1	**236.80**	**3.53**	**1.96**	**8.21**
2	239.32	2.84	1.48	8.44
3	216.31	5.55	1.72	8.46
4	224.93	5.48	0.37	8.40
5	182.56	3.49	1.59	8.41
6	165.11	2.68	0.16	8.63
7	180.52	5.13	0.33	8.93

While our methodology shows considerable promise in tuning the parameters of a PID controller in the MEEV simulated case study, it is not without its limitations, particularly regarding the data upon which it relies. The success of this approach depends heavily on the availability of collected data that adhere to specific conditions: It must encompass a variety of different setpoints and include variations stemming from poorly tuned PIDs. These requirements are critical because they ensure that the data reflect the diverse operating conditions and dynamic behaviors necessary to train accurate ML models. Consequently, the performance of the proposed method may be compromised if the historical data do not naturally encompass these variations. In addition, future investigations are required to validate the methodology on real-world data, ensuring that such datasets similarly reflect diverse setpoints and oscillatory behaviors. This further validation is critical to determine the applicability of the method in practical scenarios and to improve its generalizability across various industrial processes.

IV. CONCLUSION

This research introduces a methodology that harnesses data, machine learning (ML), and optimization techniques to automate the tuning of PID controllers, eliminating the reliance on labor-intensive manual analysis.Through the integration of an Explainable Boosting Machine (EBM) as the machine learning model with Bayesian Optimization, this methodology achieves notable efficiency in refining the PID parameters, thereby substantially alleviating the engineer's workload while simultaneously augmenting the performance of the control system. Demonstrated through simulated energy-intensive processes, this AI-driven strategy not only advances PID control, but also empowers data-driven decision-making, offering a scalable solution for industrial applications. The potential benefits—improved

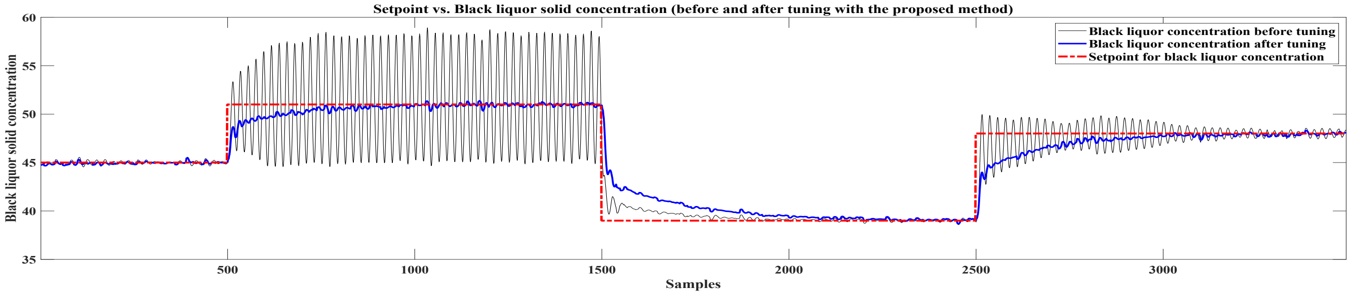

Fig. 4. Tuning performance of the proposed method

TABLE II
COMPARISON BETWEEN EXISTING AND PROPOSED PID TUNING METHODS

Performance index	Automated PID tuning methods	
	First Principal	**ML&BO (proposed method)**
ISE	4576	2961

operational efficiency, reduced energy consumption, and simplified tuning processes—underscore its value as a transformative tool for control strategies. However, real-world testing remains a critical next step to validate and refine this approach. Future studies should focus on applying this methodology to real industrial systems, testing its robustness with various PID controllers, and exploring additional ML techniques or optimization frameworks. This work paves the way for smarter, adaptive control systems that could reshape efficiency standards in energy-intensive industries.

ACKNOWLEDGMENT

This research was financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Program of Energy Research and Development (PERD) of Natural Resources Canada, and the University of Waterloo.

REFERENCES

[1] S. N. Al-Bargothi, G. M. Qaryouti, and Q. M. Jaber, “Speed control of dc motor using conventional and adaptive pid controllers,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 16, no. 3, pp. 1221–1228, 2019.
[2] K. H. Ang, G. Chong, and Y. Li, “Pid control system analysis, design, and technology,” *IEEE transactions on control systems technology*, vol. 13, no. 4, pp. 559–576, 2005.
[3] A. O'dwyer, *Handbook of PI and PID controller tuning rules*. World Scientific, 2009.
[4] L. Desborough and R. Miller, “Increasing customer value of industrial control performance monitoring-honeywell's experience,” in *AIChE symposium series*, no. 326. New York; American Institute of Chemical Engineers; 1998, 2002, pp. 169–189.
[5] R. P. Borase, D. Maghade, S. Sondkar, and S. Pawar, “A review of pid control, tuning methods and applications,” *International Journal of Dynamics and Control*, vol. 9, pp. 818–827, 2021.
[6] G. K. McMillan, *Tuning and control loop performance*. Momentum Press, 2014.
[7] T. Hägglund, “The one-third rule for pi controller tuning,” *Computers & Chemical Engineering*, vol. 127, pp. 25–30, 2019.
[8] M. Shamsuzzoha, *PID control for Industrial processes*. BoD–Books on Demand, 2018.
[9] A. K. Shah, A. Markana, and N. Parikh, “Performance analysis of imc based pid controller tuning on approximated process model,” *Nirma University Journal of Engineering and Technology*, vol. 1, no. 2, pp. 51–54, 2010.
[10] J. Berner, K. Soltesz, T. Hägglund, and K. J. Åström, “An experimental comparison of pid autotuners,” *Control Engineering Practice*, vol. 73, pp. 124–133, 2018.
[11] H. Zhang, M. E. Koujok, Q. Zhu, H. Ghezzaz, M. Amazouz, and A. ElKamel, “A data-driven approach for pids tuning in energy-intensive industry: Application to evaporator in kraft pulp mill,” in *6th European International Conference on Industrial Engineering and Operations Management*, Lisbon, Portugal, 2023.
[12] N. P. Lawrence, M. G. Forbes, P. D. Loewen, D. G. McClement, J. U. Backström, and R. B. Gopaluni, “Deep reinforcement learning with shallow controllers: An experimental application to pid tuning,” *Control Engineering Practice*, vol. 121, p. 105046, 2022.
[13] Y. Lou, R. Caruana, and J. Gehrke, “Intelligible models for classification and regression,” in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, Aug. 2012.
[14] V. Dsilva, J. Schleiss, and S. Stober, “Trustworthy academic risk prediction with explainable boosting machines,” in *International Conference on Artificial Intelligence in Education*. Springer, 2023, pp. 463–475.
[15] J. Calvin and A. Žilinskas, “One-dimensional p-algorithm with convergence rate o(n3+) for smooth functions,” *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 297–307, August 2000. [Online]. Available: https://doi.org/10.1023/A:1004699313526
[16] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 4th ed., ser. Wiley Series in Probability and Statistics. Hoboken, New Jersey: John Wiley Sons, Inc., 2016.
[17] J. L. Nazareth, “Conjugate gradient method,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 1, no. 3, pp. 348–353, 2009.
[18] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, “Numerical recipes: The art of scientific computing. rd ed,” 1992.
[19] M. J. Powell *et al.*, “The bobyqa algorithm for bound constrained optimization without derivatives,” *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, vol. 26, 2009.
[20] Y. Shi, “Particle swarm optimization,” *IEEE connections*, vol. 2, no. 1, pp. 8–13, 2004.
[21] A. Lambora, K. Gupta, and K. Chopra, “Genetic algorithm-a literature review,” in *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*. IEEE, 2019, pp. 380–384.
[22] P. I. Frazier, “A tutorial on bayesian optimization,” *arXiv preprint arXiv:1807.02811*, 2018.
[23] J. Wu, X.-Y. Chen, H. Zhang, L.-D. Xiong, H. Lei, and S.-H. Deng, “Hyperparameter optimization for machine learning models based on bayesian optimization,” *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, 2019.
[24] D. Clay, “Evaporation principles and black liquor properties,” *TAPPI Kraft Recovery Short Course*, pp. 1–6, 2011.