Hawo H. Höfer*, Joaquín E. Urrutia Gómez, Anna A. Popova, and Markus Reischl

Bayesian Optimization Driven Cancer Drug Dose-Response Curve Discovery

https://doi.org/10.1515/cdbme-2025-0191

Abstract: Screening of cancer drugs in personalized medicine and cancer treatment research is expensive. Large libraries of compounds must be evaluated, and multiple doses for each compound need to be tested to assess their viability. We introduce a method exploiting cost-aware Gaussian process Bayesian optimization to reduce the number of experiments and amount of compound spent in screening. The method is utilized to iteratively find suitable medication doses, while staying within a predefined budget. Our approach is validated using synthetic data, and subsequently tested using experimental data.

Keywords: drug screening, Gaussian process Bayesian optimization, cost-aware optimization

1 Introduction

In personalized medicine and cancer treatment research, identifying effective drugs and their optimal doses is expensive and time-consuming due to the high costs associated with traditional high-throughput screening (HTS) techniques. These techniques typically require substantial resources, and are accessible only to large pharmaceutical companies and major research centers [1]. Additionally, determining appropriate drug concentrations often involves testing multiple doses on patient-derived cancer cells, further escalating costs and complexity. Consequently, developing methods to substantially lower these costs democratizes access to advanced drug screening technologies, making them more accessible across diverse laboratories and research institutions. Furthermore, patient-derived cells are scarce and must be used effectively.

***Corresponding author: Hawo H. Höfer,** Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein Leopoldshafen, Germany, e-mail: hawo.hoefer@kit.edu
Joaquín E. Urrutia Gómez, Markus Reischl, Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein Leopoldshafen, Germany
Anna A. Popova, Institute of Biological and Chemical Systems, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein Leopoldshafen, Germany

Experimental innovations such as droplet microarrays (DMA) address these challenges by offering high-throughput capabilities at a substantially reduced scale. DMAs utilize a flat, wall-less design combined with extreme wettability contrast, allowing the formation and manipulation of thousands of nanoliter-scale droplets [2]. This approach drastically reduces sample sizes and drug volumes required for testing, enabling the use of volumes 100 to 1000 times smaller than those required by conventional platforms [3]. It is used for comprehensive testing across diverse biological systems including standard cell lines [4], spheroids [5], and patient-derived cancer cells [6, 7] with automation often facilitating subsequent analysis [8]. However, conventional-scale experiments for confirmation of DMA screening results and high-throughput screenings covering large compound libraries are still costly.

To reduce these costs at both small and large experiment scales, we introduce a method utilizing cost-aware Gaussian process Bayesian optimization (GPBO) to efficiently discover cancer medication dose-response curves. Our method iteratively identifies promising concentrations from initial experiments, driving subsequent trials. In short:

- The method is described in detail.
- It is validated using synthetic dose-response curves modeled after experimental data.
- It's viability is shown using experimental data from prior drug screenings on DMA platforms [9].

2 Method

Our cost-aware GPBO-based method exploits a simple, yet powerful concept: only perform experiments expected to provide insight. The initial step is sampling a small amount (2-3) of very low concentrations for a compound, and measuring cell viability after performing the trials. From these experiments at low concentrations, the cost-aware GPBO algorithm is utilized to iteratively minimize cell viability. At each step, it suggests candidate concentrations, and cell viabilities are measured after performing the corresponding experiments. This process repeats until the budget of compound (and trials) is exceeded.

Open Access. © 2025 The Author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

Fundamentally, GPBO is an iterative method consisting of two parts. Each iteration begins with modeling the existing data using a probabilistic regression model (a Gaussian process or GP). Subsequently, a strategy (called acquisition function) is used to query this model and identify the next generation of inputs for the experiment. Commonly used strategies balance exploration of the parameter space and exploitation of regions with desirable values. The algorithm is made *cost-aware* by utilizing an acquisition function biasing the selection of concentrations to those with lower-cost. In each iteration, the bias changes to account for previously spent budget.

The method is composed of the following steps:

1. Decide on a budget of compound and setup effort τ.
2. Perform an initial set of (cheap) tests at low doses.
3. Use a Gaussian process to approximate the dose-response curve from the current set of experiments.
4. Optimize the cost-aware acquisition function to produce concentrations leading to lower cell viabilities.
5. Stop iterating if the budget would be exceeded in the next trial. If not, perform the experiment and go to step 2.

Since the last step entails stopping the iteration if the budget is exceeded, the budget may not be fully exploited.

Technical Details

To model the dose-response curve in step 3 of the algorithm, we utilize a GP with a Matérn 5/2-Kernel. Inputs into the model are the normalized $\log$-concentrations of the compounds, and outputs are the standardized cell viability scores. Bayesian optimization is implemented using the SingleTaskGP model, acquisition functions, optimization and fitting routines provided by BoTorch [10]. The cost-cooling expected improvement acquisition function **EI-cool** [11] is utilized to account for costs while identifying concentrations for the next experiment. It attenuates expected improvement using a power α of the cost $\mathcal{C}$ of performing an experiment. α encodes the currently available and total available budget

$$\alpha = \underbrace{(\tau - \tau_{\text{used}})}_{\text{current}} / \underbrace{(\tau - \tau_{\text{init}})}_{\text{total}}. \tag{1}$$

Here, τ is the budget, τ_{init} is the budget used in the initial experiments, and τ_{used} the budget used so far. For numerical stability and noise tolerance, we adapt their definition to use BoTorch's qLogNoisyExpectedImprovement (qLogNEI)

$$\textbf{EI-cool}(c) = \textbf{qLogNEI}(\log c) - \log\left(\mathcal{C}(c)^{\alpha}\right). \tag{2}$$

We utilize a linear cost model $\mathcal{C}$ of the compound concentrations c, with the price of compound p and k related to the concentration-independent effort of experiment setup

$$\mathcal{C}(c) = pc + k. \tag{3}$$

In our experiments, we set $p = 1.0$ and $k = 0.15$. This corresponds to the setup effort k of a single experiment being $15\,\%$ of the cost of a $1\,\mu\text{M}$ dose of compound. While keeping the budget fixed, p and k allow tuning of the method's behavior.

3 Experiments and Discussion

Synthetic Data

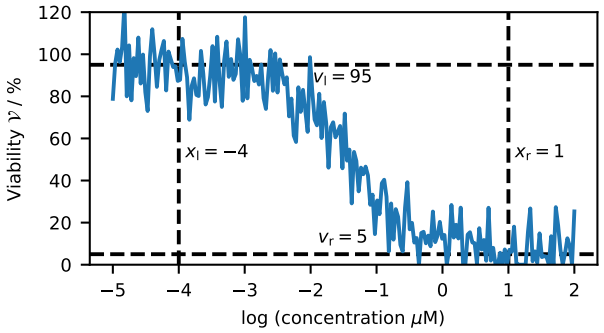

Fig. 1: Example for synthetic dose-response curve generation. The viabilities v_l and v_r are limits of viability at the left and right edge of the curve, and the concentrations x_l and x_r can be used to set the $\log$ concentrations where the viability is at the 1 and 99 $\%$ marks between v_l and v_r.

First, the method is validated using synthetic dose-response curves $\mathcal{V}$. They are generated using shifted and scaled sigmoid functions σ of the concentration c and Gaussian noise $\mathcal{N}$ with standard deviation d

$$\mathcal{V}(c)/\%\text{ viability} = v_\text{l} + \sigma(x_\text{s}(c))(v_\text{r} - v_\text{l}) + \mathcal{N}(0, d) \tag{4}$$

$$\text{with } x_\text{s}(c) = \frac{\log c - x_\text{r}}{x_\text{l} - x_\text{r}}(\sigma_{0.99} - \sigma_{0.01}) + \sigma_{0.01}. \tag{5}$$

We used $d = 10\,\%$ viability to emulate experimental noise.

Values for $\mathcal{V}(c)$ are clipped in the interval $[0, \infty)$. $\sigma_{0.01}$ and $\sigma_{0.99}$ are inputs at which the sigmoid function σ reaches 0.01 and 0.99, respectively. $\mathcal{V}$ interpolates the viabilities v_l and v_r using a sigmoid shape. x_l and x_r mark the $\log$ concentrations where the function reaches the 0.01 and 0.99 interpolation points between v_l and v_r (v_l and v_r are only reached asymptotically) Figure 1 shows an example synthetic curve. To produce all used curves are plotted in Figure 2, $x_\text{l}, x_\text{r}, y_\text{l}$ and y_r are sampled uniformly from the ranges in Table 1.

We chose sampling at the fixed concentrations used for *Dasatinib* in Popova *et al.* [9] as a baseline. The budget consumed by the baseline was computed using (3), and summing across all trials. Our method's budget is restricted to $25\,\%$ of the baseline's budget, since multiple repetitions were performed for each concentration. This means that the method could in theory sample each concentration only once, instead

Parameter	min	max	unit
x_l	-7.5	-0.5	$\log \mu\mathrm{M}$
x_r	-1.5	5.5	$\log \mu\mathrm{M}$
y_l	92.5	97.5	%
y_r	2.5	7.5	%

Tab. 1: Parameter ranges for synthetic data generation.

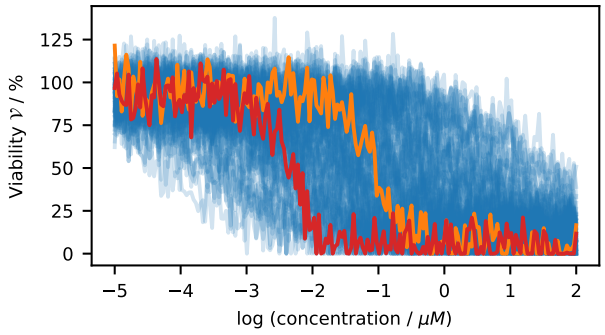

Fig. 2: All synthetic dose-response curves used for validation. Two example curves are highlighted in orange and red.

of the four repetitions in the baseline. For each synthetic dose-response curve, the initial concentrations were set to $\log c = -5$ and $\log c = -4$. From these points, the algorithm is run to iteratively produce sample concentrations. Once the budget is exhausted, $\mathcal{V}(c)$ (see Equation (4)) is fit to the sample points. It's shape is compared to the true dose-response curve by computing the average squared difference at 100 regular log concentrations on the interval $[-5, 2]$

$$\mathcal{E} = \frac{1}{100} \sum_{i=1}^{100} \left(\mathcal{V}^{\text{true}}(c_i) - \mathcal{V}^{\text{approx}}(c_i) \right)^2 . \tag{6}$$

Across all 100 samples, our method achieved a lower mean $\mathcal{E}$ of 87.5, while the baseline method achieved mean $\mathcal{E} = 171.9$. GPBO-assisted sampling on average utilizes $8.6\,\%$ of the budget of the baseline method. Thus, our model produced better fits of the true dose-response curve on average, while utilizing a fraction of the baseline's budget.

An example application of our method and the baseline to synthetic data is shown in Figure 3. The GPBO-method samples fewer locations, preferentially at lower concentrations, and utilizes $1.9\,\%$ of the baseline's budget. The method can be biased to produce more or fewer samples by increasing the setup effort k while keeping the budget constant.

Experimental Data

We utilize drug screening data of *Dasatinib* and *Vorinostat* on HeLa cells, using *Resazurin* as an indicator, as presented in Popova *et al.* [9], to test our method. The precise dose-response curves for the drugs is unknown. While the data obtained from the screenings approximates it, it cannot be used as ground truth. Acquiring a precise estimate would require many more experiments, which would incur large costs at this early stage of the method's development. Therefore, we verify our method using a different scheme compared to the synthetic data. The acquisition function is evaluated discretely for all possible inputs in the dataset, and promising concentrations are selected from these discrete values. The two smallest concentrations for each compound are used as initial concentrations, and the algorithm is run with of $75\,\%$ of the reference experiment's budget. Sigmoid functions are fit to all sample points and those selected by our method separately. These fits and the corresponding concentrations with $50\,\%$ cell viability (IC_{50}-values) are compared. To increase fitting stability, v_r was fixed at $0.0\,\%$ viability for this experiment, encoding the assumption that all compounds kill all cells at sufficiently high concentrations. Even when using all points for curve fitting, we are not able to reproduce the values obtained by Popova *et al.* [9]. The optimization algorithm used by OriginPro in their work may differ from our scipy-based [12] solution.

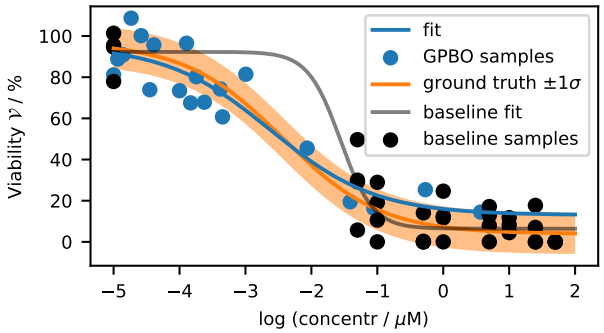

Fig. 3: Example fit produced by the GPBO-method compared to the baseline fit. It produces a better fit, while utilizing $1.9\,\%$ of the baseline's budget.

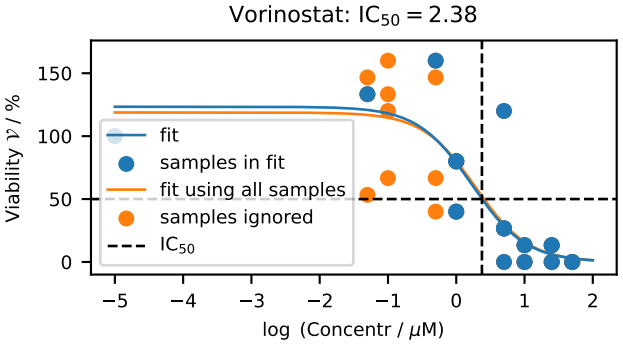

Fig. 4: Our method applied to the experimental data for *Vorinostat*. Our method's fit is close to the full fit, and the difference in IC_{50} between the two is small.

Figure 4 shows the model applied to the data for *Vorinostat*. Here, a good fit to the original function was achieved, while utilizing approximately $73\,\%$ of the budget. The produced IC_{50}-value is close to the full fit ($2.38\,\mu\mathrm{M}$ vs $2.55\,\mu\mathrm{M}$). Figure 5 shows the fit for our method compared to the fit using all datapoints for *Dasatinib*. Our method used approximately $72\,\%$ of the budget. Here, differences to the full fit can be observed, and arise from to the reduction in budget and restriction of the method's sampling points. These experiments show that even in a reduced context, our method produces sensible sampling points. It achieves results similar to the full samples while reducing budget. However, high noise and strong restrictions on the sampling points impact model performance.

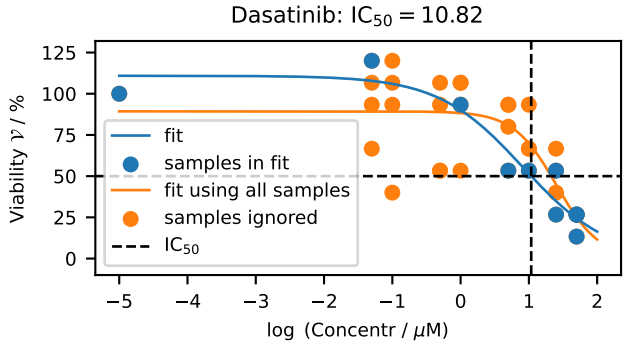

Fig. 5: Our method applied to the experimental data for *Dasatinib*. It used approximately $72\,\%$ of the budget, and found $\mathrm{IC}_{50} = 10.82$ compared to the full fit at $\mathrm{IC}_{50} = 21.24$.

4 Conclusion

We introduce a method to efficiently determine dose-response curves using cost-aware GPBO. It outperforms the sampling at fixed points on 100 synthetic dose-response curves while on average utilizing approximately $8.6\,\%$ of the budget. It works regardless of experiment scale, and can therefore be used to save costs in DMA-based screenings and their repetitions at larger scales. Our verification using experimental data shows that even when restricting the method to selection from predetermined sampling points, it can produce results comparable to full exploitation of the budget at significantly lower cost.

While our method is sequential in nature, parallelization can be achieved across compounds, or by modification of the algorithm to produce multiple candidate concentrations in each iteration. Cost-savings and sample efficiency gains made through our method can reduce the cost of screenings using DMA technology itself. Additionally, our method offers substantial savings at larger scales where larger doses and amounts of patient-derived cells are used, for example when confirming DMA screenings in conventional experiments.

Future work should verify the method in experimental settings. Additionally, prior knowledge on the sigmoid shape of the dose-response curve should be embedded into the method, by relying on specialized probabilistic models instead of a GP. This may further improve sample efficiency.

Acknowledgment: This work was funded by the program Material Systems Engineering of the Helmholtz Association (43.31.02). Our group also receives funding from the Post Lithium Storage Cluster of Excellence (PoLiS). The experimental data for testing was kindly provided by Anna A. Popova and Joaquín E. Urrutia Gómez.

References

[1] S Ghadermarzi *et al.* "Sequence-derived markers of drug targets and potentially druggable human proteins." In: *Frontiers in Genetics* 10 (2019), p. 1075.

[2] D Kartsev *et al.* "Droplet Microarrays for Miniaturized and High-Throughput Experiments: Progress and Prospectives." In: *Advanced Materials Interfaces* 12.4 (2025), p. 2400905.

[3] JE Urrutia Gómez *et al.* "Highly Parallel and High-Throughput Nanoliter-Scale Liquid, Cell, and Spheroid Manipulation on Droplet Microarray." In: *Advanced Functional Materials* 35.1 (2025), p. 2410355.

[4] H Cui *et al.* "High-throughput formation of miniaturized cocultures of 2D cell monolayers and 3D cell spheroids using droplet microarray." In: *Droplet* 2.1 (2023), e39.

[5] AA Popova *et al.* "Facile one step formation and screening of tumor spheroids using droplet-microarray platform." In: *Small* 15.25 (2019), p. 1901299.

[6] AA Popova *et al.* "Miniaturized drug sensitivity and resistance test on patient-derived cells using droplet-microarray." In: *SLAS TECHNOLOGY: Translating Life Sciences Innovation* 26.3 (2021), pp. 274–286.

[7] R El Khaled El Faraj *et al.* "Drug-Induced Differential Gene Expression Analysis on Nanoliter Droplet Microarrays: Enabling Tool for Functional Precision Oncology." In: *Advanced Healthcare Materials* 14.1 (2025), p. 2401820.

[8] JE Urrutia Gómez *et al.* "ANDeS: An automated nanoliter droplet selection and collection device." In: *SLAS technology* 29.1 (2024), p. 100118.

[9] AA Popova *et al.* "Simple assessment of viability in 2D and 3D cell microarrays using single step digital imaging." In: *SLAS technology* 27.1 (2022), pp. 44–53.

[10] M Balandat *et al.* "BoTorch: A framework for efficient Monte-Carlo Bayesian optimization." In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 21524–21538.

[11] EH Lee *et al.* "Cost-aware Bayesian optimization." In: *arXiv preprint arXiv:2003.10870* (2020).

[12] P Virtanen *et al.* "SciPy 1.0: fundamental algorithms for scientific computing in Python." In: *Nature Methods* 17.3 (2020), pp. 261–272.