



OPEN ACCESS

EDITED BY

Roland Wohlgemuth,
Lodz University of Technology, Poland

REVIEWED BY

Johannes Felix Buyel,
University of Natural Resources and Life
Sciences Vienna, Austria
Ezhaveni Sathiyamoorthi,
Yeungnam University, Republic of Korea
Suvi Tuulikki Häkkinen,
VTT Technical Research Centre of Finland Ltd.,
Finland

*CORRESPONDENCE

Lars Leyendecker,
✉ lars.leyendecker@ipt.fraunhofer.de

[†]These authors have contributed equally to this work

RECEIVED 24 April 2025

ACCEPTED 10 July 2025

PUBLISHED 19 September 2025

CITATION

Leyendecker L, Nausch H, Wergers C,
Scheffler D and Schmitt RH (2025) Bayesian
experimental design for optimizing medium
composition and biomass formation of tobacco
BY-2 cell suspension cultures in stirred-
tank bioreactors.
Front. Bioeng. Biotechnol. 13:1617319.
doi: 10.3389/fbioe.2025.1617319

COPYRIGHT

© 2025 Leyendecker, Nausch, Wergers,
Scheffler and Schmitt. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Bayesian experimental design for optimizing medium composition and biomass formation of tobacco BY-2 cell suspension cultures in stirred-tank bioreactors

Lars Leyendecker^{1*†}, Henrik Nausch^{2†}, Christian Wergers¹,
Dirk Scheffler² and Robert H. Schmitt^{1,3}

¹Department for Production Quality, Fraunhofer Institute for Production Technology IPT, Aachen, Germany, ²Department Model-Based Product and Bioprocess Engineering, Fraunhofer Institute for Molecular Biology and Applied Ecology IME, Aachen, Germany, ³Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany

Introduction: The plant suspension cell culture Bright Yellow-2 (BY-2) can be an economic platform for producing complex biopharmaceuticals like cytokines and antibodies at small- and medium-scale because of potentially reduced cultivation and purification costs compared to mammalian cells. This is especially relevant for rare diseases. However, the productivity is currently low in terms of biomass formation in a typical batch fermentation. A potential reason might be that the standard BY-2 cultivation medium, as it is used under laboratory shake-flask cultivation conditions, has not yet been comprehensively optimized and tested in industrial bioreactor settings, addressing all four macronutrients relevant for biomass formation (i.e., sucrose, ammonium, nitrate, and phosphate) in parallel. In this article, we therefore propose a multi-variate, multi-objective, and batch-wise Bayesian experimental design (BED) approach for optimally parameterizing macronutrient supply in the cultivation medium, promoting the fresh mass (FM) increase (i.e., growth rate) and final FM (i.e., biomass).

Methods: We performed a sequential and adaptive experimentation utilizing the BED to optimize the cultivation medium in four iterations with four different media each and confirmed the results in two additional experimentation rounds.

Results: Our results show that while nitrate and phosphate can be used to adjust the growth rate (i.e., reaching up to 40 g/L × d FM), it is possible to reduce sucrose and ammonium without impacting the growth rate and only affecting the final biomass yield (i.e., reaching up to 300 g/L FM). Thereby, we improved the overall productivity of biomass formation (i.e., as a ratio between nutrient and FM input and FM output) for this batch fermentation process by 36%.

Discussion: The results demonstrate the advantages of BED to generate new medium compositions (i.e., macronutrient concentrations) by unbiasedly varying the design space. Thereby, the discovery of new process insights (i.e., impact of individual macronutrient concentrations on the growth rate and biomass formation) is facilitated. Moreover, we show that our data-efficient and black-box BED approach represents a promising alternative to traditional design of

experiments (DoE) and mechanistic white-box modeling. Established for the plant suspension cell line BY-2, our BED approach might also apply to yeast and mammalian cells.

KEYWORDS

biopharmaceuticals, upstream production, batch fermentation, process optimization, Bayesian optimization

1 Introduction

The growing prosperity and aging society worldwide have led to a constant rise of non-communicable civilization diseases such as diabetes (affecting 30% of the world population), cancer (affecting 1% of the world population), and numerous rare diseases (affecting in total up to 6% of the world population, whereby each individual disease accounts for less than 0.001%). These diseases cannot be efficiently treated with synthetic drugs (small-molecules). Accordingly, there is an increasing demand for recombinant complex biopharmaceuticals like growth factors, metabolic enzymes, and antibodies (biologics) (World Health Organization, 2024). These biopharmaceuticals are predominantly produced in mammalian cell cultures like human embryonic kidney (HEK) or Chinese hamster ovary (CHO). However, due to the high risk of contamination with human viruses (requiring absolute sterility during production and sophisticated purification processing), mammalian cells are only economic at large-scale (e.g., 2,000-L to 20,000-L bioreactors producing kg amounts) but not at small- or medium-scale (e.g., 5-L to 500-L bioreactors providing g amounts), which is why production costs can exceed 150,000 €/g of biopharmaceutical for rare diseases (Kelley, 2009; Hernandez et al., 2018; Puetz and Wurm, 2019; Walsh and Walsh, 2022). In these cases, plant cell cultures (PCCs) might be a vegan alternative (free of any animal components) because human pathogens do not replicate in plant cells, enabling simplified production and purification processes and making this production platform also economic at small- and mid-scale, potentially lowering the costs to 10,000 €/g (Nausch et al., 2023b). For these reasons, a glucocerebrosidase for the treatment of *Morbus Gaucher* and an α -galactosidase for the treatment of *Morbus Fabry* are already produced in plant cells (Protalix, Israel) as approved drugs (Xu et al., 2011; Xu and Zhang, 2014; Santos et al., 2016; Moon et al., 2019; Nausch et al., 2023b).

Among different PCCs like carrot (*Daucus carota*), rice (*Oryza sativa*), and tobacco (*Nicotiana tabacum*), the tobacco cell line Bright Yellow-2 (BY-2) became the prevalent PCC for the production of biopharmaceuticals (Nagata et al., 1992; Kowalczyk et al., 2022; Nagata, 2023). This is because the BY-2 cell line has the highest growth rate among plant cells ($\mu_{max} = 0.03/h$ corresponding to a doubling time of 25–30 h compared to a $\mu_{max} = 0.001–0.004$ or a doubling time of 250–500 h for other PCCs), being similar to HEK and CHO. Moreover, the BY-2 cell line can be grown to high cell densities of up to 600 g/L fresh mass (FM) (corresponding to 18 g/L dry mass (DM)) in 10,000-L to 100,000-L large-scale stirred-tank reactors (STRs), and several operation modes like batch, fed-batch, and semi-continuous fermentation regimes are already established (Hogue et al., 1990; Gao and Lee, 1992; Ho et al., 1995; Fischer et al., 1999; Ullisch, 2012; Ullisch et al., 2012; Reuter et al., 2014; Raven et al., 2015; Häkkinen et al., 2018). Given these features and due to its higher productivity compared to other PCCs in commercial settings,

the BY-2 cell line is preferred for industrial applications, particularly for the production of biopharmaceuticals for rare diseases. For example, Protalix recently shifted from carrot to BY-2 cells (Hanania et al., 2017). Nevertheless, PCCs still represent a niche because the yields of 10 mg to 1 g/L tend to be lower than those of 100 mg to 6 g/L, which can be achieved in HEK and CHO by batch and fed-batch fermentation (Kunert and Reinhart, 2016).

Similar to CHO cells (Schinn et al., 2021; Yeo et al., 2022), it has been recently demonstrated that the volumetric productivity of PCCs can be substantially increased by model-based optimization of the cultivation in a typical batch fermentation setting (Nausch et al., 2023a). For example, in an iterative design of experiment (DoE)-based experiment-modeling-optimization workflow, a mechanistic (Monod-type) white-box model was established that linked the nutrient consumption to the biomass formation. The model is based on the macronutrients sucrose, ammonium, nitrate, and phosphate plus the initial FM as five controllable input variables, whereby the macronutrients are converted by the initial FM into new biomass, thus defining the growth rate and final FM yield.

Using sucrose concentration in the cultivation medium as a key parameter, a substantial increase in the growth rate and final FM yield of BY-2 cells at the end of the batch phase was achieved through multi-criteria Pareto optimization. Although the mechanistic white-box modeling aims to construct a model precisely representing the collected data, the predictability of conditions not represented in the data is neglected. Thus, the model could not predict cell growth under non-optimal conditions deviating from the parameter space conditions in the experiments used for model setup and refinement (Nausch et al., 2023a). Moreover, such mechanistic white-box models become limited in cases where (i) the relevant process knowledge is missing (e.g., the plant cell metabolism and kinetics); (ii) not all relevant parameters can be measured (e.g., intracellular macronutrient pools); (iii) there is high variation in the data (e.g., between replicates).

In principle, small-scale bioreactors or even microtiter plates can be used to obtain more data. However, in the case of PCC, the data obtained in these systems are hardly transferable to industrial large-scale bioreactors. While suitable scale-down models were already developed for mammalian cells (i.e., CHO), this has not yet been achieved for PCCs due to several reasons: (i) Plant cells (with an average diameter of 100–500 μm) have a volume that is 100 to 1,000 times larger than that of mammalian cells (which typically measure 10–100 μm in diameter) and sediment up to 25 times faster. Therefore, they cannot be regarded as two-phase systems consisting of only a gas and liquid phase, with in the liquid phase homogeneously distributed cells. Instead, PCCs represent a three-phase system comprising a gas, liquid and solid phase, with the plant cells as a solid phase. This introduces an additional dimension of heterogeneity. (ii) PCCs are associated with rheological constraints as the PCC medium represents a viscous, non-Newtonian broth, leading to imperfect mixing. (iii) PCCs are shear sensitive, so the

stirrer and mixing speed must be controlled during scale-up to avoid cell damage (Doran, 1999; Kieran et al., 1997; Cheung et al., 2018). This limits the applicability of small-scale bioreactors for process optimization so that they are rather used for process establishment, i.e., cell line generation and selection. For the actual process optimization, larger bioreactors are needed to get representative data. However, the larger the scale, the fewer conditions can be tested (Cheung et al., 2018).

As an alternative to mechanistic white-box modeling, Bayesian experimental design (BED) is a global black-box optimization strategy that does not make any assumptions about the properties of the process. BED represents an adaptive experimental design approach that utilizes Bayesian optimization (BO) to iteratively perform experiments to identify optimal process parameters (Duris et al., 2020; Garnett, 2023; Rainforth et al., 2024; Zhang et al., 2021; Figure 2). A key difference between BED and mechanistic (theory-based) modeling lies in the aim of BED not to create a model that is as accurate as possible but to efficiently find the optimal process settings. In contrast to traditional, non-sequential DoE methods (e.g., one-factor-at-a-time, response surface methodology, full factorial, or partial factorial DoE), BED represents a sequential DoE approach, builds a surrogate model of the objective function, and utilizes an acquisition function for maximizing the informational content of the experiment series. Thus, BED iteratively performs experimental design, execution, and evaluation and adapts the experimental design based on previous experiment results in each iteration. Thereby, BED aims to minimize the total number of experiments for the model setup and process optimization. Moreover, whereas DoE approaches are limited to low-dimensional design spaces with a limited number of discrete parameter levels and manually defined experimental designs (e.g., Bayer et al., 2021), BED allows the exploration of higher-dimensional continuous design spaces guided by the acquisition function that optimizes the informational gain of each experiment. Consequently, BED promises to be a data-driven, exploratory, bias-free, and efficient alternative to traditional DoE that is particularly suited for complex-to-evaluate bioreactor-based fermentation processes with multiple parameters (i.e., batch, fed-batch, semi-continuous, and continuous) at an industrial scale.

In this article, we investigate the applicability of BED to PCC fermentation processes of BY-2 and maximize the productivity of the system by means of biomass formation per time while minimizing the resource input and additionally identifying the best trade-off between both. In order to search for an optimal cultivation medium composition that promotes biomass formation (i.e., as a ratio between nutrient and FM input and FM output) during typical batch fermentation of BY-2 cells in STRs (Patent WO2015165583A1, Figure 1), we propose a multi-variate, multi-objective, and batch-wise BED algorithm. With the long-term intention to establish a mechanism for adaptive, dynamic real-time process monitoring and control (e.g., for a semi-continuous long-term fermentation), we focused on FM and sucrose as parameters that can be measured in- or online.

Accordingly, the contributions of our article can be summarized as follows:

1. We propose a novel and data-driven multi-variate, multi-objective, and batch-wise BED approach for yield-optimal parameterization of biomass formation.
2. We apply our proposed BED algorithm for analyzing and finding optimal cultivation medium compositions in the batch fermentation of BY-2 cells in STRs and report the experiment results.
3. We analyze, interpret, and discuss the results in relation to the findings of previous studies (i.e., parameter-parameter and parameter-objective relationships in the BY-2 batch fermentation).

This proof-of-concept study will demonstrate that BED can be used to efficiently optimize a typical batch fermentation process of BY-2 by utilizing parameters measurable in- or online, without aiming for a complete optimization of the fermentation process itself.

2 Materials and methods

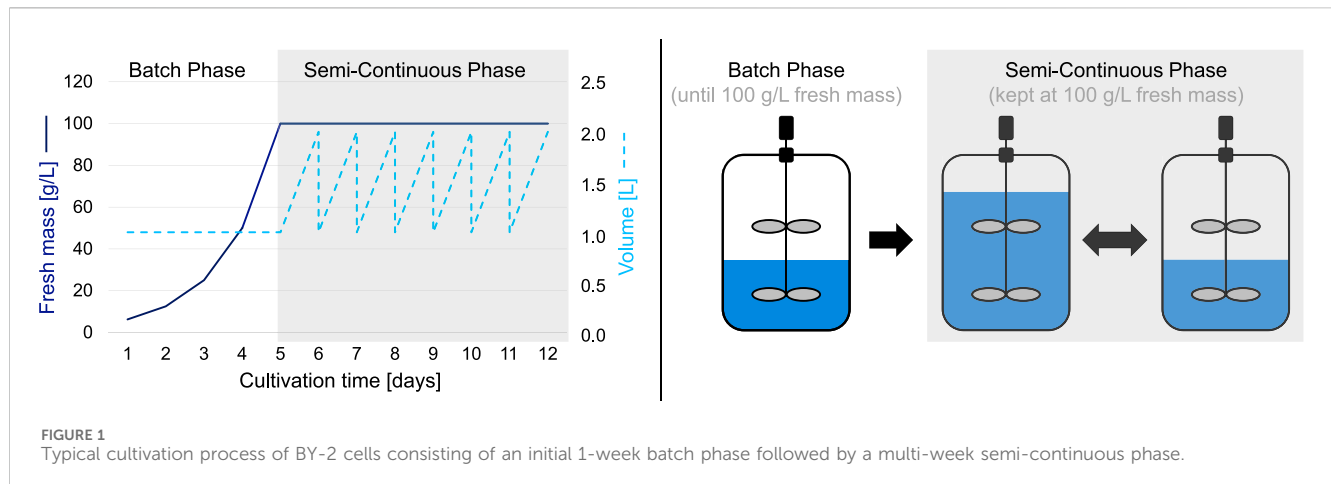
2.1 Cultivation of tobacco cells

As a representative batch fermentation for BY-2 cells, we aimed to optimize the patented process (Patent WO2015165583A1, Figure 1) that employs the Murashige & Skoog (MS) medium, as this is typically used for the cultivation of BY-2 cells (Nausch et al., 2023b). The BY-2 cells were grown as previously described by Nausch et al. (2023a). In summary, the cultivation of BY-2 routine cells was done in 200 mL of MS medium in 1,000-mL Erlenmeyer shake flasks in a Climo-Shaker ISF1-X orbital shaker (Kuhner Shaker, Herzogenrath, Germany) at 26°C and 160 rpm, and the BY-2 cells were passaged every 7 days. For cultivation by fermentation in STRs (Patent WO2015165583A1), these 7-day-old BY-2 cells from the routine culture were used to inoculate 1,000 mL of a defined medium (Supplementary Table S1) in a 2,000-mL STR (Geringe Deutschland, Rastatt, Germany).

The patented cultivation in STRs is divided into two phases (Figure 1). In phase 1, the batch phase, the BY-2 cells are cultivated over several days in a constant medium volume without any additional medium feed until an FM of 100 g/L is reached. The batch phase is followed by the semi-continuous phase, in which new cultivation medium is continuously fed into the bioreactor to keep the FM constant at 100 g/L, the medium is harvested daily, and the volume is reduced to the initial one (in this example, 1,000 mL). Due to significantly different optimization conditions between a batch phase (quasi-discrete experiments with static parameter settings that are independent of a previous state) and a semi-continuous phase (long-term sequence of several interdependent parameter settings with intermediate measurements), a two-stage optimization system is required. Consequently, in this article, we focus solely on the optimization of the batch phase (which is why the semi-continuous phase is grayed out in Figure 1).

2.2 Cell biomass and macro nutrient analysis

While FM was determined through vacuum infiltration, the DM was determined by subsequent drying in an oven at 60°C for 3 days (Nausch et al., 2023a). Nutrients were determined by specific offline assays (sucrose: abcam ab83387, ammonium: Sigma-Aldrich



MAK310, nitrate: Cayman Cay780001, phosphate: abcam ab65622) according to the manufacturer's instructions.

2.3 Bayesian experimental design (BED)

BO is a sample-efficient statistical method for surrogate-based global optimization of expensive-to-evaluate and noisy objective functions $\mathbf{y} = f(\mathbf{x}) + \varepsilon$ (Greenhill et al., 2020; Kumar et al., 2022). Based on the BO framework, BED represents a sequential DoE method for scientific experimentation in engineering systems. BED is a global optimization approach that does not make any assumptions about the properties of the to-be-optimized system. As a black-box optimization algorithm, BED can be applied (but limited) to black-box systems that (1) do not have a closed-form representation; (2) do not provide functional derivatives; (3) only allow for point-wise evaluation (Greenhill et al., 2020). Under the influence of a noise term ε , the objective function $f(\mathbf{x})$ transforms a vector of n independent input variables $\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathcal{X} \subseteq \mathbb{R}^n$ into m dependent objective variables $\mathbf{y} = [y_1, y_2, \dots, y_m]^T \in \mathcal{Y} \subseteq \mathbb{R}^m$. The multi-variate and multi-objective global black-box optimization problem can be expressed as finding the solution to Equation 1:

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) \quad (1)$$

by optimizing the analytically unknown objective function yielding Equation 2:

$$\mathbf{y}^* = \max_{\mathbf{x} \in \mathcal{X}} f(\mathbf{x}) = f(\mathbf{x}^*) \quad (2)$$

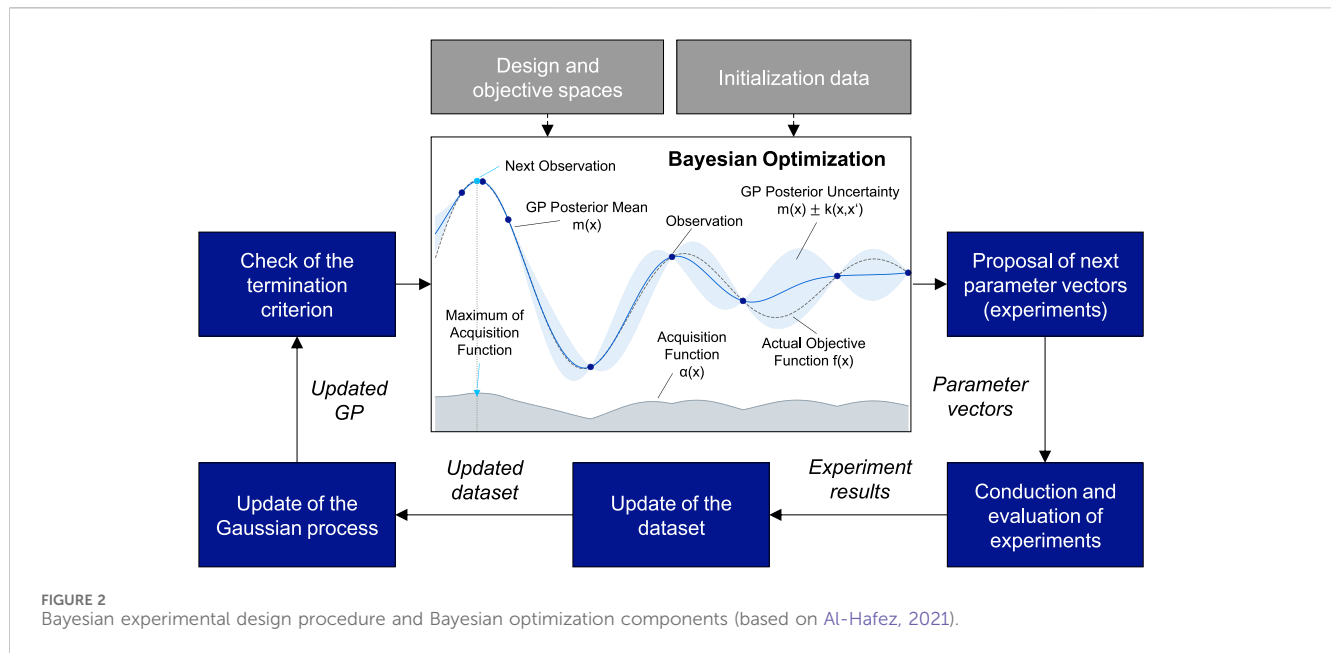
Therefore, the aim of BED is to find the optimal parameterization $\mathbf{x}^*$ within the n -dimensional design space $\mathcal{X} \in \mathbb{R}^n$ that yields an optimal process output $\mathbf{y}^*$ in the m -dimensional objective space $\mathcal{Y} \in \mathbb{R}^m$. For sequential optimization, BED comprises two core components: (1) a surrogate model to model $f(\mathbf{x})$ and (2) an acquisition function $\alpha(\mathbf{x})$ to determine the next experiments. While extra trees, random forests, tree parzen estimators, support vector machines, and Bayesian neural networks are suitable surrogate model options, Gaussian processes (GPs) (also known as Gaussian process regressors) (Rasmussen, 2004) are the preferred surrogate model

choice because of their data efficiency, flexibility, simplicity, and built-in quantification of uncertainty (Garnett, 2023; Johnson et al., 2025; Kariminejad et al., 2024; Harris et al., 2024; Kumar et al., 2022; Rodemann et al., 2024). Consequently, GPs are the most commonly used surrogate models in data-sparse, noisy, and low-dimensional optimization problems. GPs are non-parametric and probabilistic models that are constructed directly from a dataset, allowing the complexity of the model to grow with the number of data elements (Duris et al., 2020). GPs fulfill the requirement of the acquisition function to provide uncertainty estimates of the objective function for given parameter sets (Duris et al., 2020). A GP comprises a mean $m(\mathbf{x})$ and a covariance function (also denoted kernel) $k(\mathbf{x}, \mathbf{x}')$ (Equation 3). The kernel describes the similarities between two sets of input parameters $\mathbf{x}$ and $\mathbf{x}'$ (Duris et al., 2020).

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \quad (3)$$

BED substitutes the optimization of the expensive-to-evaluate objective function with the optimization of an inexpensive, analytically differentiable, and therefore, more tractable acquisition function $\alpha: \mathcal{X} \rightarrow \mathbb{R}$ (Garnett, 2023). The acquisition function $\alpha(\mathbf{x})$ uses the expected mean $m(\mathbf{x})$ and uncertainty $k(\mathbf{x}, \mathbf{x}')$ of the GP to make an informed decision about the next set of parameter values by assigning a score to each parameter location to maximize the information gain (Gabler and Wollherr, 2022; Garnett, 2023).

According to the BED algorithm (Figure 2), starting with a predefined design space $\mathcal{X}$ comprising one or multiple parameters x_i and based on an initial dataset $\mathcal{D}(\mathbf{x}, \mathbf{y})$, BED iteratively proposes a set of parameter values (e.g., nutrient concentrations) that defines the next set of experiments (step 1). The initial dataset $\mathcal{D}(\mathbf{x}, \mathbf{y})$ can contain historical process data or can be systematically collected using quasi-random sampling approaches (e.g., Sobol algorithm (Sobol', 1967) or space-filling designs) that aim to equidistantly cover and explore $\mathcal{X}$. After BED has proposed a parameter vector $\mathbf{x}_t$ at iteration t , the experiments are conducted and evaluated (step 2). The evaluation (measuring and quantifying the process outcomes for each objective) yields the experiment results $\mathbf{y}_t$ for each parameter set $\mathbf{x}$ (step 3). The results of the experiments are added to the dataset $\mathcal{D}(\mathbf{x}, \mathbf{y})$ (step 4). The GP surrogate model is then updated (step 5), and afterward, the fulfillment of the termination criterion is checked (step 6). This constitutes the end



of the iteration. The termination criterion can be arbitrarily defined by the expert and can include acceptance levels for each objective, experimental budgets, and convergence measures. As long as the termination criterion is not fulfilled and the optimization has not been manually terminated by the expert, experimentation continues.

2.4 Multi-variate, multi-objective, and batch-wise Bayesian experimental design for optimization of the BY-2 batch fermentation

To establish BED for the optimization of biomass formation and nutrient consumption of BY-2 in batch cultivation mode, we define a five-dimensional design space comprising the five individual macronutrients of the cultivation medium, that is, sucrose, ammonium, nitrate, and phosphate, and the initial FM concentration. The lower and upper limits for each medium component were determined in previous experiments (Nausch et al., 2023a). Macronutrient concentrations were chosen so that reducing or increasing them further did not enhance the BY-2 growth rate or final biomass yield or even reduced them. Furthermore, we define a two-dimensional target space that comprises the final FM concentration (i.e., biomass) and the corresponding FM concentration increase per hour (i.e., growth rate). The full specifications of the to-be-optimized biomass formation process, including parameters and objectives, are given in Figure 3. Details on the parameters (units and valid value ranges) and objectives (units, minimum and maximum values, and direction of optimization) are contained in Tables 1, 2. As stated above, minimum and maximum values are based on process experience and previous experiments (Nausch et al., 2023a). Accordingly, we define the biomass formation as a multi-variate and multi-objective optimization problem $y = f(x)$, assuming a black-box system comprising five controllable input parameters and two fully observable objectives (FM concentration (g/L) and FM

concentration increase ($g/(L \cdot h)$) with x_i being the set of input parameters

with $i \in \{\text{sucrose, phosphate, ammonium, nitrate, start FM}\}$ and y_j being the two objective variables with $j \in \{\text{final FM, FM increase}\}$.

We utilize BED as a statistical sequential optimization method that aims to efficiently identify optimal parameter values x^* that yield an optimal process outcome y^* of the objective function $f(x)$ (Greenhill et al., 2020). Due to the high complexity of fermentation processes in STRs at industrial scales, only a limited number of experiments can be carried out. Given the small dataset size, we decided to use GPs as surrogate models because of their data efficiency. For multi-objective optimization, a multi-objective GP is used to model the two objectives of FM concentration and FM concentration increase (i.e., growth rate). Based on the five input variables, the GP performs a regression task predicting the objective value for each one of the objectives within the corresponding value range. The GP is initialized using a constant mean as the mean function and the Matérn-5/2 kernel as the covariance function. We initialize the GPs on historical data from Nausch et al. (2023a), comprising a total of ten experiments (Batch 0 in Supplementary Tables S2, S3). Due to the duration of individual experiments of 7 days (164–172 h), we decided to run four optimization iterations (termination criterion), followed by two iterations for confirmation of the results. The GP models the input noise as homoscedastic noise. Automatic relevance determination (ARD) is used to iteratively determine the relevance of each parameter and adjust the GP's length scale parameters iteratively during optimization (Duvenaud, 2014).

Because the fermentation setup in this study allows the simultaneous fermentation of four separate medium compositions, we propose batch-wise BED allowing the simultaneous provision of four experiment sets (batch-size $q = 4$). Accordingly, in each iteration (batch), BED proposes a set of four different medium compositions (experiments). In contrast to single-point optimization ($q = 1$), where the parameter set with the highest acquisition function value is selected, a set of q experiments with the q

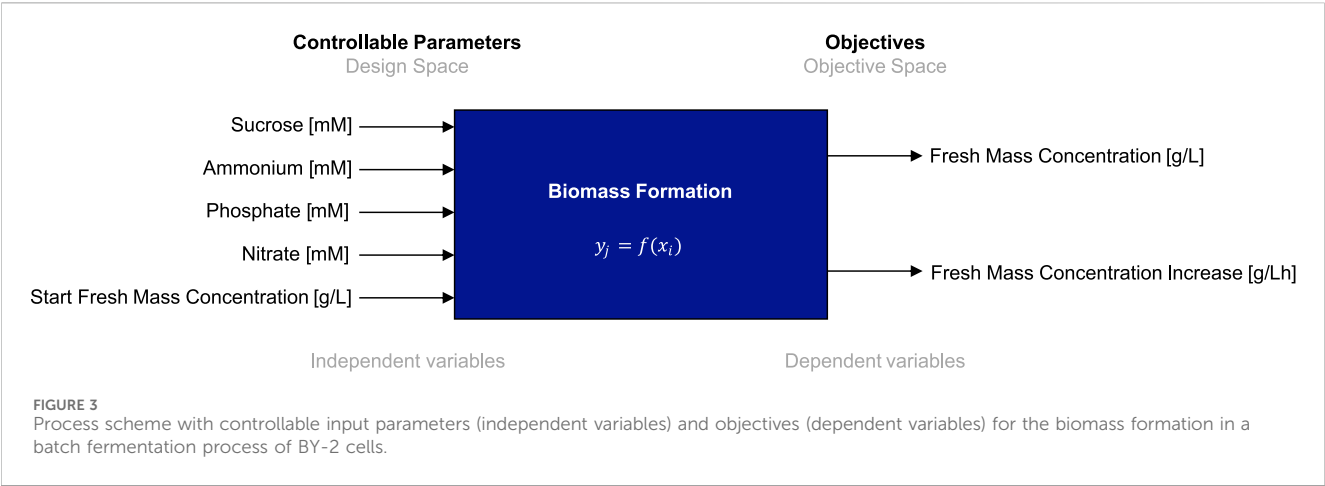


TABLE 1 Overview of controllable input parameters (design space) for biomass formation in a batch fermentation process of BY-2 cells.

Design parameter cultivation medium component	Unit	Minimum value	Maximum value
Sucrose concentration (C ₁₂ H ₂₂ O ₁₁)	mM	20	175
Ammonium concentration (NH ₄ ⁺)	mM	5	60
Phosphate concentration (PO ₄ ³⁻)	mM	2	15
Nitrate concentration (NO ₃ ⁻)	mM	5	150
Start fresh mass concentration	g/L	5	22

TABLE 2 Overview of objectives (objective space) for biomass formation in a batch fermentation process of BY-2 cells.

Target variables	Unit	Minimum value	Maximum value	Optimization type
Fresh mass concentration	g/L	0	200	Maximization
Fresh mass concentration increase	g/(L×h)	0	25	Maximization

highest acquisition function values is proposed. Although, the batch size positively correlates with the required number of experiments because parameter sets with suboptimal informational gain are selected (Zhan et al., 2024), we decided to parallelize experimentation across all four STRs to reduce experimentation time by a factor of four. As a multi-objective, noisy, and batch-wise acquisition function $\alpha(x)$, q Noisy Expected Hypervolume Improvement (qNEHVI) is used (Daulton et al., 2021). $\alpha(x)$ uses the hypervolume (HV) metric as a multi-objective performance metric and determines next experiment locations based on the approximated improvement of the HV (Audet et al., 2021).

2.5 Software implementation, computational resources, and source code

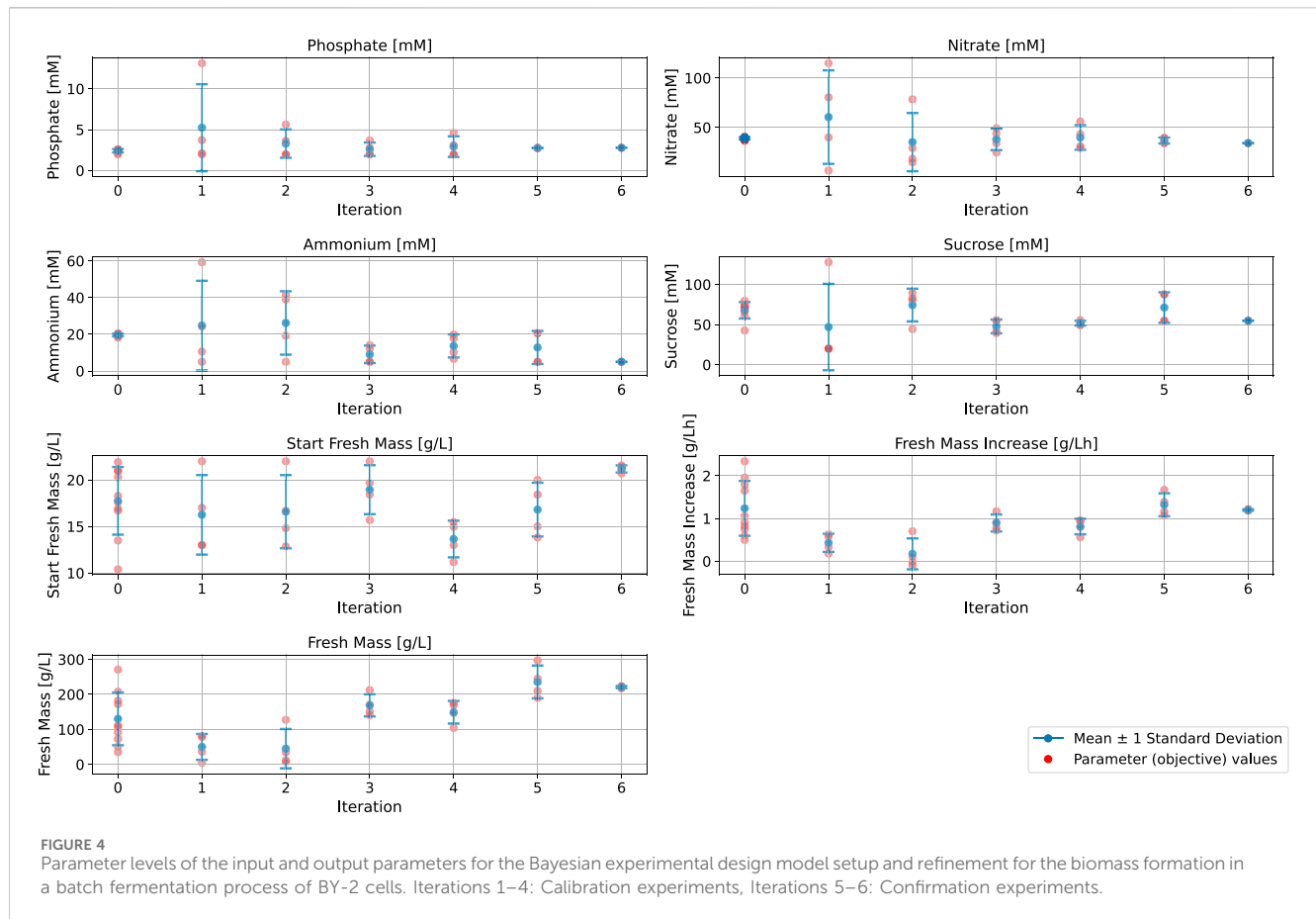
Our approach for multi-variate, multi-objective, and batch-wise BED was implemented using the Python programming language and the Python library BoTorch (Balandat et al., 2019). BoTorch is a BO programming framework that is built on top of PyTorch (Paszke

et al., 2019). The source code is available on GitLab (<https://gitlab.cc-asp.fraunhofer.de/fraunhofer-ipt/biodapt>). Calculations were performed on hardware running Linux Ubuntu 22.04 equipped with Dual Intel Xeon Gold CPUs with 16 cores, Dual Nvidia Quadro RTX5000 with 16 GB, and 384 GB of DDR4 RAM.

3 Results and discussion

3.1 The BY-2 growth rate can be regulated by nitrate and phosphate, while it is possible to reduce sucrose and ammonium without impacting the growth rate and only affecting the final biomass yield

The batch fermentation lasted for 7 days (164–172 h) with daily FM measurements (intervals of 22–26 h). The FM increase was determined for these 7 days (at the end of fermentation) and for 4 days (at the time point at which 100 g/L FM is usually reached and the semi-continuous phase is typically started). The BED algorithm



for the optimization of macronutrients and biomass formation (Figure 3; Table 1; Table 2) showed convergence behavior (i.e., stabilization of parameters and objectives over the number of iterations and rapprochement toward consistent values) within the four iterations comprising a total of 16 different media (Figure 4; Supplementary Table S1). The medium composition that led to the highest growth rate and final biomass yield was confirmed in two further iterations. Accordingly, Supplementary Table S1 contains the experimental results, and Figure 4 shows the evolution of the distributions of parameters and objectives along the iterations. In detail, the distributions of all five parameters after all iterations are shown in Figure 5. In analogy, Figure 6 shows the distribution of both FM and FM increase after all iterations. Please refer to the appendix for a detailed analysis of the progressions of parameters and objectives along all iterations (Supplementary Figure S1 (sucrose), Supplementary Figure S2 (ammonium), Supplementary Figure S3 (nitrate), Supplementary Figure S4 (phosphate), Supplementary Figure S5 (start FM), Supplementary Figure S6 (final FM), and Supplementary Figure S7 (FM increase)).

Iteration 0 contains historical data from previous experiments (Nausch et al., 2023a). Iterations 1–4 are the optimization runs, and iterations 5 and 6 are the confirmation runs. The convergence behavior within the first four iterations (iteration 0 not included) is indicated by the narrowing distributions of the values for all four macronutrients (sucrose, ammonium, nitrate, phosphate) and the initial FM (Figure 4). Based on the results of iterations 1 to 4, we

selected the set of parameter values that led to the highest final biomass yield (FM) and growth rate (FM increase) while requiring comparatively less concentration of the individual media components (highlighted in blue in Supplementary Table S1). The selected medium was then confirmed in a side-by-side comparison with the standard MS medium used for the cultivation of BY-2 (iteration 5), and we confirmed its reproducibility by parallel fermentations (iteration 6).

During early iterations of the BED algorithm, the starting FM showed the widest spread of the input parameters, and sucrose and ammonium displayed a broader distribution than nitrate and phosphate, which exhibited a strong concentration around specific parameter values (Figure 4). The greater spread of the initial FM, sucrose, and ammonium during the optimization process, compared to nitrate and phosphate, can be interpreted as the BED algorithm is attributing a higher relevance to these parameters, leading to a greater variation compared to those considered less relevant. However, in the case of sucrose and ammonium, the data also indicated that these two parameters can be altered without a substantial impact on the growth rate, although there was a pronounced effect on the final biomass (Figure 6). Vice versa, because of the strong impact of nitrate and phosphate on FM increase (growth rate), these two parameters seem to converge toward specific values and thus can be used to regulate the growth rate.

In analogy to the distributions of the media components (Figure 5), Figure 6 shows the distribution of final FM and

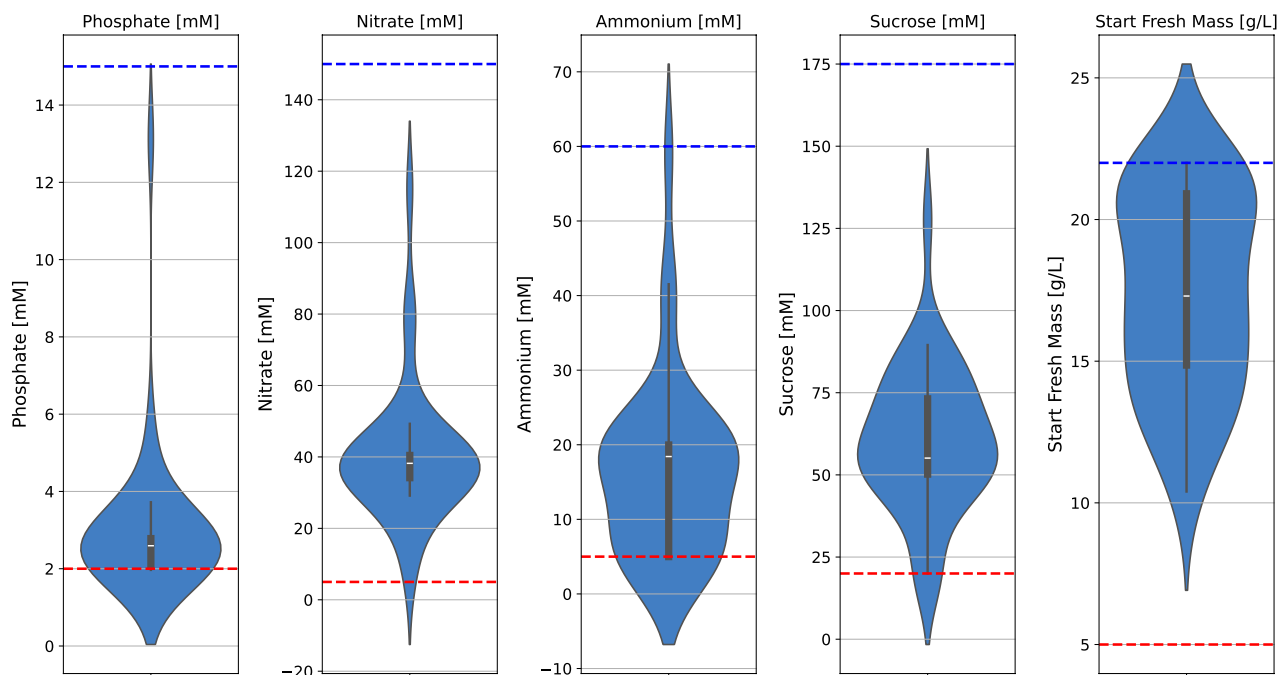


FIGURE 5
Distribution of all five input parameters after all iterations, including upper (blue) and lower (red) design space limits (dashed lines) in a batch fermentation process of BY-2 cells.

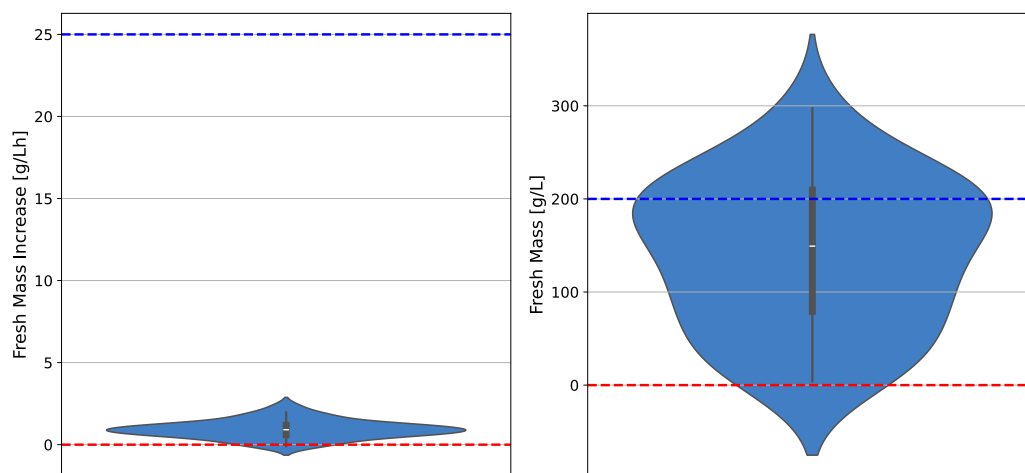


FIGURE 6
Distribution of all two objective values after all iterations, including upper (blue) and lower (red) objective value limit ranges.

growth rate. The final FM exhibits a broader distribution than the growth rate. The wider spreading of final FM might result from the fact that this objective is affected by all five parameters and their corresponding changes, while the growth rate is mainly affected by nitrate and phosphate.

Compared to the MS standard medium, the final medium M31 had a substantially reduced sucrose concentration as carbon source of -38% , which complies with the previous study, in which the cultivation medium was optimized via a mechanistic model for the nutrient consumption, indicating that sucrose is in excess in the

MS medium and can be reduced without impacting growth rate and only affecting the final biomass (Nausch et al., 2023a).

Similarly to sucrose, the nitrogen source ammonium was also drastically lowered in the M31 medium by -76% , which in turn complies with the fact that alternative cultivation media such as the Gamborg B5 (Gamborg et al., 1968), Schenk and Hildebrandt (SH) (Schenk and Hildebrandt, 1972), and Chu N6 Chih-ching (Chu et al., 1975) contain only low levels of ammonium at 2.02 mM, 2.61 mM and 7.00 mM, respectively. This might be explained by the fact that BY-2 cannot grow in a medium without ammonium

(Holland et al., 2010; Ullisch, 2012), but higher ammonium levels are potentially cytotoxic (Behrend and Mateles, 1975; Behrend and Mateles, 1976). However, as already discussed by Nausch et al. (2023a), another study in which the MS medium was optimized using DoE showed that a reduction of ammonium to 6.87 mM led to a decreased final biomass after a 7-day batch cultivation in shake flasks (Häkkinen et al., 2018). In addition, in two further studies, it has been observed for a similar shake-flask setting that an increase of ammonium up to 51.54 mM promotes the BY-2 final biomass (Holland et al., 2010; Ullisch, 2012; Ullisch et al., 2012; Holland, 2013). Hence, even though lower ammonium concentrations did not seem to affect the growth rate, they did limit the final biomass yield.

In contrast to ammonium, the nitrate level as the second nitrogen source was only slightly altered in the M31 medium at −13%. The reasons for that might be that, on the one hand, nitrate is metabolized to ammonium and thus must compensate for the reduction in ammonium in the medium. On the other hand, an increase in the nitrate concentration negatively impacts the BY-2 growth, as previously shown (Holland et al., 2010; Ullisch, 2012; Ullisch et al., 2012; Holland, 2013; Vasilev et al., 2013). Accordingly, nitrate might be used to adjust the growth rate if needed.

Phosphate remained nearly constant with +2% in the M31 medium, even though it has already been found that increasing phosphate concentration in the MS medium from 2.72 mM to 10.00 mM boosted the biomass accumulation five-fold in a 7-day batch fermentation in shake flasks (Holland, 2013). Hence, adjusting the phosphate might be another option to alter the growth rate in the batch fermentation process.

With respect to the initial FM, the results show that lowering the inoculation FM did not lead to a higher growth rate but rather resulted in a lower final biomass, which contradicts the observation in the previous study (Nausch et al., 2023a). The reason for that must be elucidated. Nevertheless, while the initial conditions yielded a final biomass of up to 81.41 g/L FM and a growth rate of 0.63 g/(L × h), this was increased to 212.41 g/L FM and 1.17 g/(L × h) in the medium M31 in iteration 3.

3.2 The BY-2 biomass yield was compromised by the lower sucrose concentration

To confirm the observations from the BED experiments, a side-by-side comparison of both media in a 7-day batch fermentation was conducted (iteration 5), and its reproducibility was confirmed by parallel fermentations (iteration 6).

Notably, there was a high variation between identical runs with both the MS and M31 medium. Such variations are normal and possibly result from the natural variability of the BY-2 cells as a biological component. Nevertheless, the BY-2 growth rates and biomass yields were similar for the MS and M31 media until day 4. Afterward, the BY-2 growth rate declined in the M31 and yielded a substantially lower final biomass at day seven because the sucrose was depleted in the M31 medium (Supplementary Figure S1; Supplementary Table S1). This confirms that a lower sucrose concentration does not impact the growth rate but limits the final FM yield. However, at day 4 to day 5, the FM concentration reached 100 g/L, where typically the semi-continuous phase starts,

such as in the historical data used to initialize the BED algorithm, which might explain this mismatch in the optimization. Ammonium was consumed completely within 2 days of fermentation but led to an increased degradation of nitrate without any obvious negative impact on growth rate, which was also true for phosphate, which was consumed within 2 days as well. The latter confirms the previously discussed assumption that BY-2 cells seem to have an intracellular pool for phosphate that is used in case this macronutrient is depleted in the medium (Nausch et al., 2023a). Hence, the final M31 medium yielded similar growth rates to the standard MS medium, while the final biomass yield was lower due to the lower sucrose concentration.

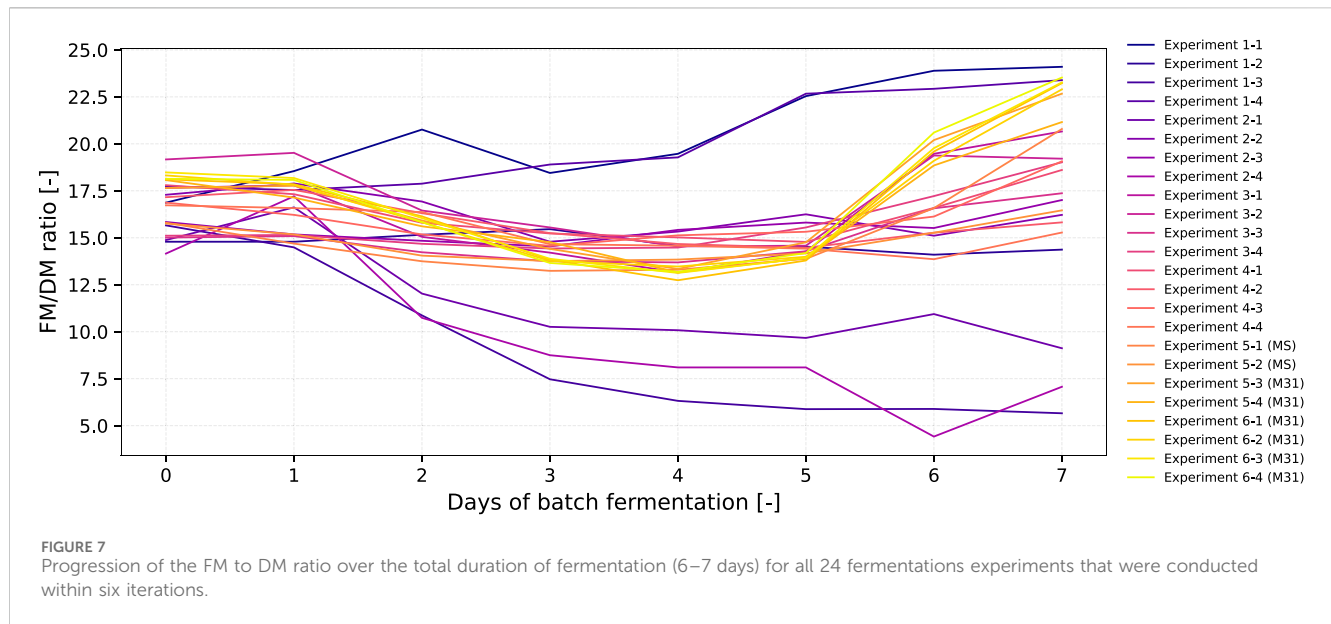
3.3 The fresh mass might be used as a surrogate parameter for the biomass for real-time monitoring and optimization of a BY-2 fermentation process

With the long-term intention of establishing a mechanism for adaptive, dynamic real-time process monitoring and control (e.g., for a semi-continuous long-term fermentation), we focused on FM and sucrose as modeling parameters as these can be measured in- or online, allowing adaptive real-time process monitoring and optimization. However, the FM might be an error-prone parameter as it can be impacted by excessive water uptake or loss, while DM represents the actual biomass more accurately. For example, it is well known that when BY-2 cells enter the stationary phase, they stop growing and the DM remains constant, whereas they still consume water so that FM further increases. Thus, we compared the FM/DM ratio in the various runs as shown in Figure 7 and Supplementary Table S2 and Supplementary Table S3.

For most of the iterations in which BY-2 grew (including medium M31), the FM/DM ratio slightly decreased when the cells shifted from the lag to the exponential phase at days 1–3. Starting from days 5–6, this trend changed, and the FM/DM ratio tended to increase, indicating that the cells then entered the stationary phase from the exponential phase. In the case of the MS reference medium, the increase started at days 6–7. This difference might be explained by the fact that the MS medium was designed to enable a typical 7-day shake-flask batch cultivation, while the BED algorithm was trained on historical data, in which the batch fermentation ended at day 4–5 and the semi-continuous phase with feed supplementation started, as mentioned previously. Nevertheless, the FM/DM ratio seems to be rather constant when the BY-2 cells are in the exponential phase (i.e., between the lag and stationary phase), as it is the case in the semi-continuous long-term fermentation of BY-2 cells, in which the FM concentration is kept at 100 g/L FM by feeding (Figure 1). Thus, the FM might be used as a surrogate parameter for biomass for real-time monitoring and optimization of a BY-2 fermentation process.

4 Conclusion

The biomass formation of the plant suspension cell line BY-2 in a typical batch fermentation process was successfully analyzed and



significantly improved by adapting the nutrient composition in the cultivation medium with only a few experiments and iterations, using a sequential and adaptive BED algorithm, and only using parameters that can be measured in- or online. Notably, the performance equaled that of a previously developed mechanistic white-box model, demonstrating that BED represents an alternative approach for the optimization of biomass formation, particularly in the case when experimental data are limited, which is usually the case for industrial fermentation settings. Our data-driven, black-box BED algorithm allows a bias-free investigation of the experimental design space and thereby facilitates the discovery of novel process insights.

Specifically, our results indicate that nitrate and phosphate can be utilized to regulate the growth rate (up to 40 g/L-d), whereas sucrose and ammonium can be reduced without affecting the growth rate, though this may influence the final biomass yield (up to 300 g/L FM). Notably, when comparing the concentration of the four macronutrients in the MS medium plus the standard initial FM to the one used in the final M31 medium with the reduced initial FM, the optimized setting led to a 36% increase in the biomass yield by means of a ratio between macronutrient and FM input to FM output.

Importantly, this study aimed to provide the proof-of-concept that BED can be used to optimize typical fermentation processes without aiming to fully optimize the fermentation process itself. Hence, we did not fully analyze the fermentation process to resolve the trade-off between growth rate and final biomass yield. For example, the impact of reduced sucrose and ammonium concentrations on the final biomass yield and growth rate has to be elaborated in the future, as the final biomass yield was compromised through the optimization rounds. Moreover, we will conduct further experiment series using our BED algorithm to analyze its optimization behavior and robustness.

Beyond, we focused on the quasi-discrete optimization of the batch phase. Given the state-dependent and long-term characteristics of the subsequent semi-continuous

fermentation phase, regular BED does not represent a suitable approach. In future work, we will therefore investigate contextual BO (Char et al., 2019; Fiducioso et al., 2019) for the optimization of the semi-continuous phase. Moreover, to make BED more efficient and to strengthen acceptance of BED among biopharmaceutical experts, our research will focus on intensifying the collaboration between algorithm and domain experts to investigate how domain knowledge can be integrated into the BED workflow before, during, and after optimization (Kanarik et al., 2023; Hvarfner et al., 2022; Savage and Del Chanona, 2023). To further enhance the overall experimentation, the relationship between batch size and cost-to-target must be analyzed as well.

Nevertheless, we demonstrated the applicability of data-driven BED for data-efficient non-sequential experimental design for the plant cell line BY-2. BED might also be applicable to microbial and mammalian cell lines in batch fermentation mode.

Data availability statement

The data presented in this study is available on GitLab: <https://gitlab.cc-asp.fraunhofer.de/fraunhofer-ipt/biodapt>.

Author contributions

LL: Writing – original draft, Formal Analysis, Project administration, Methodology, Conceptualization, Investigation, Supervision, Writing – review and editing, Software. HN: Writing – original draft, Formal Analysis, Project administration, Methodology, Conceptualization, Investigation, Supervision, Writing – review and editing, Validation, Data curation. CW: Software, Writing – review and editing, Data curation, Visualization. DS: Formal Analysis, Writing – review and editing, Data curation. RS: Resources, Writing – review and editing, Funding acquisition.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Funded by the Fraunhofer Leistungszentrum “ICNAP - International Center for Networked, Adaptive Production.” A Fraunhofer Initiative (Grant no. 423).

Acknowledgments

We would like to thank Mario Pothen from Fraunhofer Leistungszentrum “ICNAP - International Center for Networked, Adaptive Production” for his support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Al-Hafez, F. (2021). Finding the optimal learning rate using bayesian optimization. Available online at: <https://firasalhafez.com/2021/05/12/finding-the-optimal-learning-rate-using-bayesian-optimization/>.
- Audet, C., Bigeon, J., Cartier, D., Le Digabel, S., and Salomon, L. (2021). Performance indicators in multiobjective optimization. *Eur. J. Operational Res.* 292, 397–422. doi:10.1016/j.ejor.2020.11.016
- Balandat, M., Karrer, B., Jiang, D. R., Daulton, S., Letham, B., Wilson, A. G., et al. (2019). Botorch: a framework for efficient Monte-Carlo bayesian optimization. *arXiv*. doi:10.48550/arXiv.1910.06403
- Bayer, B., Duerkop, M., Striedner, G., and Sissolak, B. (2021). Model transferability and reduced experimental burden in cell culture process development facilitated by hybrid modeling and intensified design of experiments. *Front. Bioeng. Biotechnol.* 9, 740215. doi:10.3389/fbioe.2021.740215
- Behrend, J., and Mateles, R. I. (1975). Nitrogen metabolism in plant cell suspension cultures: I. effect of amino acids on growth. *Plant physiol.* 56, 584–589. doi:10.1104/pp.56.5.584
- Behrend, J., and Mateles, R. I. (1976). Nitrogen metabolism in plant cell suspension cultures: II. role of organic acids during growth on ammonia. *Plant physiol.* 58, 510–512. doi:10.1104/pp.58.4.510
- Char, I., Chung, Y., Neiswanger, W., Kandasamy, K., Nelson, A. O., Boyer, M., et al. (2019). “Offline contextual bayesian optimization,” in *Advances in neural information processing systems*. Editors H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Amalho, A. Alché-Buc, E. Fox, and R. Garnett (Red Hook, United States: Curran Associates, Inc), 32.
- Cheung, C. K. L., Leksawadi, N., and Doran, P. M. (2018). Bioreactor scale-down studies of suspended plant cell cultures. *AIChE J.* 64, 4281–4288. doi:10.1002/aic.16415
- Chu, C.-C., Wang, C.-C., Sun, C. S., Hsu, C. S., Yin, K. C., Chu, C.-Y., et al. (1975). Establishment of an efficient medium for anther culture of rice through comparative experiments on the nitrogen sources. *Sci. Sinter.* 18, 659–668. Available online at: <https://www.scienceopen.com/document?vid=7e4d50a5-ee4a-452e-9b74-e2bbe2d47fb4r>.
- Daulton, S., Balandat, M., and Bakshy, E. (2021). Parallel bayesian optimization of multiple noisy objectives with expected hypervolume improvement. *arXiv*. doi:10.48550/arXiv.2105.08195
- Doran, P. M. (1999). Design of mixing systems for plant cell suspensions in stirred reactors. *Biotechnol. Prog.* 15, 319–335. doi:10.1021/bp990042v
- Duris, J., Kennedy, D., Hanuka, A., Shtalenkova, J., Edelen, A., Baxevanis, P., et al. (2020). Bayesian optimization of a free-electron laser. *Phys. Rev. Lett.* 124, 124801. doi:10.1103/PhysRevLett.124.124801
- Duvenaud, D. (2014). “Automatic model construction with gaussian processes,”. Ph.D. thesis (Cambridge, United Kingdom: Apollo - University of Cambridge Repository).
- Fiducioso, M., Curi, S., Schumacher, B., Gwerder, M., and Krause, A. (2019). Safe contextual bayesian optimization for sustainable room temperature PID control tuning. *arXiv*, 5850–5856. doi:10.24963/ijcai.2019/811
- Fischer, R., Liao, Y. C., and Drossard, J. (1999). Affinity-purification of a TMV-specific recombinant full-size antibody from a transgenic tobacco suspension culture. *J. Immunol. methods* 226, 1–10. doi:10.1016/s0022-1759(99)00058-7
- Gabler, V., and Wollherr, D. (2022). Bayesian optimization with unknown constraints in graphical skill models for compliant manipulation tasks using an industrial robot. *Front. robotics AI* 9, 993359. doi:10.3389/frobt.2022.993359
- Gamborg, O. L., Miller, R. A., and Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Exp. cell Res.* 50, 151–158. doi:10.1016/0014-4827(68)90403-5
- Gao, J., and Lee, J. M. (1992). Effect of oxygen supply on the suspension culture of genetically modified tobacco cells. *Biotechnol. Prog.* 8, 285–290. doi:10.1021/bp00016a004
- Garnett, R. (2023). *Bayesian optimization*. Cambridge, United Kingdom: Cambridge University Press.
- Greenhill, S., Rana, S., Gupta, S., Vellanki, P., and Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: a review. *IEEE Access* 8, 13937–13948. doi:10.1109/ACCESS.2020.2966228
- Häkkinen, S. T., Reuter, L., Nuorti, N., Joensuu, J. J., Rischer, H., and Ritala, A. (2018). Tobacco BY-2 media component optimization for a cost-efficient recombinant protein production. *Front. plant Sci.* 9, 45. doi:10.3389/fpls.2018.00045
- Hanania, U., Ariel, T., Tekoah, Y., Fux, L., Sheva, M., Gubbay, Y., et al. (2017). Establishment of a tobacco BY2 cell line devoid of plant-specific xylose and fucose as a platform for the production of biotherapeutic proteins. *Plant Biotechnol. J.* 15, 1120–1129. doi:10.1111/pbi.12702
- Harris, S. B., Vasudevan, R., and Liu, Y. (2024). Active oversight and quality control in standard bayesian optimization for autonomous experiments. *arXiv*. doi:10.48550/arXiv.2405.16230
- Hernandez, I., Bott, S. W., Patel, A. S., Wolf, C. G., Hospodar, A. R., Sampathkumar, S., et al. (2018). Pricing of monoclonal antibody therapies: higher if used for cancer? *Am. J. Manag. care* 24, 109–112. Available online at: <https://www.ajmc.com/view/pricing-of-monoclonal-antibody-therapies-higher-if-used-for-cancer>.
- Ho, C. H., Henderson, K. A., and Rorrer, G. L. (1995). Cell damage and oxygen mass transfer during cultivation of *Nicotiana tabacum* in a stirred-tank bioreactor. *Biotechnol. Prog.* 11, 140–145. doi:10.1021/bp00032a004
- Hogue, R. S., Lee, J. M., and An, G. (1990). Production of a foreign protein product with genetically modified plant cells. *Enzyme Microb. Technol.* 12, 533–538. doi:10.1016/0141-0229(90)90071-W
- Holland, T. (2013). *Development of plant cell cultures with respect to the industrial production of biopharmaceutical ingredients*. PhD thesis. Aachen, Germany: RWTH Aachen University.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2025.1617319/full#supplementary-material>

- Holland, T., Sack, M., Rademacher, T., Schmale, K., Altmann, F., Stadlmann, J., et al. (2010). Optimal nitrogen supply as a key to increased and sustained production of a monoclonal full-size antibody in BY-2 suspension culture. *Biotechnol. Bioeng.* 107, 278–289. doi:10.1002/bit.22800
- Hvarfner, C., Stoll, D., Souza, A., Lindauer, M., Hutter, F., and Nardi, L. (2022). Pbo: augmenting acquisition functions with user beliefs for bayesian optimization. *arXiv*. doi:10.48550/arXiv.2204.11051
- Johnson, J. E., Jamil, I. R., Pan, L., Lin, G., and Xu, X. (2025). Bayesian optimization with gaussian-process-based active machine learning for improvement of geometric accuracy in projection multi-photon 3d printing. *Light, Sci. and Appl.* 14, 56. doi:10.1038/s41377-024-01707-8
- Kanarik, K. J., Osowiecki, W. T., Lu, Y. J., Talukder, D., Roschewsky, N., Park, S. N., et al. (2023). Human-machine collaboration for improving semiconductor process development. *Nature* 616, 707–711. doi:10.1038/s41586-023-05773-7
- Kariminejad, M., Tormey, D., Ryan, C., O'Hara, C., Weinert, A., and McAfee, M. (2024). Single and multi-objective real-time optimisation of an industrial injection moulding process via a bayesian adaptive design of experiment approach. *Sci. Rep.* 14, 29799. doi:10.1038/s41598-024-80405-2
- Kelley, B. (2009). Industrialization of mAb production technology: the bioprocessing industry at a crossroads. *mAbs* 1, 443–452. doi:10.4161/mabs.1.5.9448
- Kieran, P. M., MacLoughlin, P. F., and Malone, D. M. (1997). Plant cell suspension cultures: some engineering considerations. *J. Biotechnol.* 59, 39–52. doi:10.1016/S0168-1656(97)00163-6
- Kowalczyk, T., Mercz-Sadowska, A., Picot, L., Brčić Karačonji, I., Wiecefinska, J., Śliwiński, T., et al. (2022). Genetic manipulation and bioreactor culture of plants as a tool for industry and its applications. *Mol. Basel, Switz.* 27, 795. doi:10.3390/molecules27030795
- Kumar, A., Rana, S., Shilton, A., and Venkatesh, S. (2022). “Human-ai collaborative bayesian optimisation,” in *Advances in neural information processing systems*. Editors S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (New Orleans, United States: Curran Associates, Inc), 35, 16233–16245.
- Kunert, R., and Reinhart, D. (2016). Advances in recombinant antibody manufacturing. *Appl. Microbiol. Biotechnol.* 100, 3451–3461. doi:10.1007/s00253-016-7388-9
- Moon, K.-B., Park, J.-S., Park, Y.-I., Song, I.-J., Lee, H.-J., Cho, H. S., et al. (2019). Development of systems for the production of plant-derived biopharmaceuticals. *Plants Basel, Switz.* 9, 30. doi:10.3390/plants9010030
- Nagata, T. (2023). Hidden history of the tobacco BY-2 cell line. *J. plant Res.* 136, 781–786. doi:10.1007/s10265-023-01490-4
- Nagata, T., Nemoto, Y., and Hasegawa, S. (1992). Tobacco BY-2 cell line as the “HeLa” cell in the cell biology of higher plants. *Int. Rev. Cytol.* 132, 1–30. doi:10.1016/S0074-7696(08)62452-3
- Nausch, H., Baldan, M., Teichert, K., Lutz, J., Claussen, C., Bortz, M., et al. (2023a). Simulation and optimization of nutrient uptake and biomass formation using a multi-parameter monod-type model of tobacco BY-2 cell suspension cultures in a stirred-tank bioreactor. *Front. plant Sci.* 14, 1183254. doi:10.3389/fpls.2023.1183254
- Nausch, H., Knödler, M., and Buyel, J. F. (2023b). “Production of complex proteins in plants: from farming to manufacturing,” in *Biopharmaceutical manufacturing*. Editor R. Pörtner (Cham: Springer International Publishing), Vol. 11, 241–278. doi:10.1007/978-3-031-45669-5\textunderscore{8}
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. *arXiv*. doi:10.48550/arXiv.1912.01703
- Puetz, J., and Wurm, F. M. (2019). Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. *Processes* 7, 476. doi:10.3390/pr7080476
- Rainforth, T., Foster, A., Ivanova, D. R., and Bickford Smith, F. (2024). Modern bayesian experimental design. *Stat. Sci.* 39. doi:10.1214/23-ST5915
- Rasmussen, C. E. (2004). “Gaussian processes in machine learning,” in *Advanced lectures on machine learning*. Editors O. Bousquet, U. Von Luxburg, and G. Rätsch (Berlin, Heidelberg: Springer Berlin Heidelberg), 3176, 63–71. doi:10.1007/978-3-540-28650-9_4
- Raven, N., Rasche, S., Kuehn, C., Anderlei, T., Klöckner, W., Schuster, F., et al. (2015). Scaled-up manufacturing of recombinant antibodies produced by plant cells in a 200-L orbitally-shaken disposable bioreactor. *Biotechnol. Bioeng.* 112, 308–321. doi:10.1002/bit.25352
- Reuter, L. J., Bailey, M. J., Joensuu, J. J., and Ritala, A. (2014). Scale-up of hydrophobin-assisted recombinant protein production in tobacco by-2 suspension cells. *Plant Biotechnol. J.* 12, 402–410. doi:10.1111/pbi.12147
- Rodemann, J., Croppi, F., Arens, P., Sale, Y., Herberinger, J., Bischl, B., et al. (2024). Explaining bayesian optimization by shapley values facilitates human-ai collaboration. *arXiv*. doi:10.48550/arXiv.2403.04629
- Santos, R. B., Abranches, R., Fischer, R., Sack, M., and Holland, T. (2016). Putting the spotlight back on plant suspension cultures. *Front. plant Sci.* 7, 297. doi:10.3389/fpls.2016.00297
- Savage, T., and Del Chanona, E. A. R. (2023). Expert-guided bayesian optimisation for human-in-the-loop experimental design of known systems.
- Schenk, R. U., and Hildebrandt, A. C. (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can. J. Bot.* 50, 199–204. doi:10.1139/b72-026
- Schinn, S.-M., Morrison, C., Wei, W., Zhang, L., and Lewis, N. E. (2021). Systematic evaluation of parameters for genome-scale metabolic models of cultured mammalian cells. *Metab. Eng.* 66, 21–30. doi:10.1016/j.ymben.2021.03.013
- Sobol', I. (1967). On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Comput. Math. Math. Phys.* 7, 86–112. doi:10.1016/0041-5553(67)90144-9
- Ullisch, D. (2012). *A fundamental research of growth, metabolism and product formation of tobacco suspension cells at different scales*. PhD thesis. Aachen, Germany: RWTH Aachen University.
- Ullisch, D. A., Müller, C. A., Maibaum, S., Kirchhoff, J., Schiermeyer, A., Schillberg, S., et al. (2012). Comprehensive characterization of two different nicotiana tabacum cell lines leads to doubled gfp and ha protein production by media optimization. *J. Biosci. Bioeng.* 113, 242–248. doi:10.1016/j.jbiosc.2011.09.022
- Vasilev, N., Grömping, U., Lipperts, A., Raven, N., Fischer, R., and Schillberg, S. (2013). Optimization of BY-2 cell suspension culture medium for the production of a human antibody using a combination of fractional factorial designs and the response surface method. *Plant Biotechnol. J.* 11, 867–874. doi:10.1111/pbi.12079
- Walsh, G., and Walsh, E. (2022). Biopharmaceutical benchmarks 2022. *Nat. Biotechnol.* 40, 1722–1760. doi:10.1038/s41587-022-01582-x
- World Health Organization (2024). Noncommunicable diseases.
- Xu, J., and Zhang, N. (2014). On the way to commercializing plant cell culture platform for biopharmaceuticals: present status and prospect. *Pharm. Bioprocess.* 2, 499–518. doi:10.4155/pbp.14.32
- Xu, J., Ge, X., and Dolan, M. C. (2011). Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures. *Biotechnol. Adv.* 29, 278–299. doi:10.1016/j.biotechadv.2011.01.002
- Yeo, H. C., Park, S.-Y., Tan, T., Ng, S. K., Lakshmanan, M., and Lee, D.-Y. (2022). Combined multivariate statistical and flux balance analyses uncover media bottlenecks to the growth and productivity of chinese hamster ovary cell cultures. *Biotechnol. Bioeng.* 119, 1740–1754. doi:10.1002/bit.28104
- Zhan, D., Zeng, Z., Wei, S., and Wu, P. (2024). A simple and efficient approach to batch bayesian optimization. *arXiv*. doi:10.48550/arXiv.2411.16206
- Zhang, M., Parnell, A., Brabazon, D., and Benavoli, A. (2021). Bayesian optimisation for sequential experimental design with applications in additive manufacturing. *arXiv*. doi:10.48550/arXiv.2107.12809