OXFORD

Optimal design of dynamic experiments for scalar-on-function linear models with application to a biopharmaceutical study

Damianos Michaelides [1,*], Maria Adamou[2], David C. Woods[2], Antony M. Overstall[2,*]

[1]Biostatistics Unit, The Cyprus Institute of Neurology and Genetics, 1683 Nicosia, Cyprus, [2]Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, SO17 1BJ, United Kingdom

*Corresponding authors: Damianos Michaelides, Biostatistics Unit, The Cyprus Institute of Neurology and Genetics, 1683 Nicosia, Cyprus (damianosm@cing.ac.cy) and Antony M. Overstall, Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, SO17 1BJ, United Kingdom (A.M.Overstall@soton.ac.uk).

ABSTRACT

A Bayesian optimal experimental design framework is developed for experiments where settings of one or more variables, referred to as profile variables, can be functions. For this type of experiment, a design consists of combinations of functions for each run of the experiment. Within a scalar-on-function linear model, profile variables are represented through basis expansions. This allows finite-dimensional representation of the profile variables and optimal designs to be found. The approach enables control over the complexity of the profile variables and model. The method is illustrated on a real application involving dynamic feeding strategies in an Ambr250 modular bioreactor system.

KEYWORDS: basis functions; design of experiments; functional linear model; profile variables; dynamic experiments.

1 INTRODUCTION

In a variety of experiments, especially in biological studies, there can be more flexibility over the control of design variables than is the norm in standard experimental designs. In particular, it is becoming increasingly common to run studies where the value taken by one or more *profile variables* can be adjusted during a single run of the experiment. Hence, for these variables, optimal design concerns how to vary their values through the choice of a function for each variable in each run of the experiment. We refer to such examples as *dynamic experiments*.

Previous research on the design and analysis of dynamic experiments can mostly be classified under one of two different frameworks.

(i) Optimal design for dynamic mechanistic models, usually derived as the solution to a set of differential equations, which naturally incorporate time-evolving behavior. Early work was reviewed by Titterington (1980). Further developments have often come from the field of control theory (e.g., Espie and Macchietto, 1989), especially applications in chemical and biological processes where dynamical systems abound (e.g., Braniff and Ingalls, 2018). The methods and designs are often quite closely tailored to the specific experiment or model being studied.

(ii) Response surface methodology (RSM), typically using standard designs. Georgakis (2013) proposed the ex-

tension of standard RSM to include profile variables. Designs were obtained through application of standard RSM or optimal designs to the so-called sub-factors defined by a dimension reduction of the profile variables (see Section 2). Roche (2018) applied related methods to nuclear safety experiments and extended the approach to include data-driven dimension reduction, for example, principal components analysis. These ideas have now seen fairly wide-spread adoption in a variety of pharmaceutical applications, see Georgakis (2024), including extensions to experiments with functional responses.

Our work fits within this second framework, and is motivated by biopharmaceutical process development via experiments using an Ambr250 modular bioreactor system. Such a system allows process conditions to be individually controlled for each reaction. A typical experiment would measure the concentration of the product of interest from each reaction (run of the experiment) after a fixed time period, usually of the order of two weeks.

Design of experiments methods have seen considerable uptake on such systems; for example, Tai et al. (2015) applied a definitive screening design to perform variable screening. However, previous applications of statistically designed experiments in such systems have only employed traditional *static* variables, whose values are set and kept constant through each individual run. Several studies have, though, highlighted the potential benefits of dynamic experimentation. Yoon et al. (2003), Trummer

et al. (2006) and Rameez et al. (2014) all found that varying the temperature during a reaction can increase concentration. Lu et al. (2013) discussed the increase in concentration that could be obtained from a dynamic feeding strategy in place of the usual fixed bolus feeding. However, none of these studies used experiments specifically designed to estimate the effects of profile variables.

To study the applicability of such suggestions to their own studies, our collaborators at GlaxoSmithKline (GSK) wanted to design and run a series of dynamic experiments. We use one of these experiments to demonstrate our novel approach to designing such studies. The practicalities of the experiment involved varying three static variables: initial viable cell concentration (IVCC), pH, and temperature, and one profile variable: feed volume. The study aimed to investigate the titer content of the product (concentration), with the eventual aim of optimizing cell growth.

In this paper, we present a novel Bayesian approach to designing experiments with profile variables, assuming a scalar-on-function linear model (Section 2). Optimal designs are found exploiting the connections between this model and standard linear model theory (Section 3). The impact of the model on the resulting designs is explored via illustrative examples (Section 4) and methods demonstrated via application to the Ambr250 experiment (Section 5).

2 SCALAR-ON-FUNCTION LINEAR MODELS AND BASIS REPRESENTATIONS

2.1 Scalar-on-function linear models

Suppose there are J profile variables that can be controlled over a time interval $\mathcal{T} = [0, T]$, where $x_j(t)$ denotes the value of the jth profile factor at time $t \in \mathcal{T}$. For $j = 1, \ldots, J$, it is assumed that $x_j(\cdot) \in \mathcal{X}_j \subset \mathcal{L}^2(\mathcal{T})$, where $\mathcal{L}^2(\mathcal{T})$ is the set of all square integrable functions on $\mathcal{T}$. The sets $\mathcal{X}_1, \ldots, \mathcal{X}_J$ are determined by the nature of the experiment and the profile variables. For example, if the jth profile variable cannot be changed over $\mathcal{T}$, then $\mathcal{X}_j$ will be the set of constant functions. Therefore, a standard, static experiment with fixed controllable variables is a special case.

The experiment consists of n runs. For $i = 1, \ldots, n$, the ith run consists of specifying the J controllable profile variables, $\boldsymbol{x}_i(t) = \left[x_{i1}(t), \ldots, x_{iJ}(t) \right]^{\mathrm{T}} \in \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_J$, and measuring the scalar response, y_i, at time $t = T$.

To investigate the effect of the J profile variables on the response, we assume a scalar-on-function linear model (e.g., Ramsay and Silverman, 2005, pages 261-277). Specifically, it is assumed

$$y_i = \int_{\mathcal{T}} \boldsymbol{f} \left[\boldsymbol{x}_i(t) \right]^{\mathrm{T}} \boldsymbol{\beta}(t) \, \mathrm{d}t + \epsilon_i, \qquad (1)$$

for $i = 1, \ldots, n$. In (1), $\boldsymbol{f}[\boldsymbol{x}(t)] = \left\{ f_1[\boldsymbol{x}(t)], \ldots, f_Q[\boldsymbol{x}(t)] \right\}^{\mathrm{T}}$ is a vector of regression functions of the profile variables, controlling the complexity of the model, where it is assumed that $f_q : \mathcal{X} \to \mathcal{G}_q \subset \mathcal{L}^2(\mathcal{T})$. For example, a first-order model with an intercept has $\boldsymbol{f}[\boldsymbol{x}(t)] = [1, \boldsymbol{x}(t)]^{\mathrm{T}}$. Furthermore, $\boldsymbol{\beta}(t) = \left[\beta_1(t), \ldots, \beta_Q(t) \right]^{\mathrm{T}}$ is a $Q \times 1$ vector of unknown functional parameters. It is assumed that

$\beta_q(t) \in \mathcal{B}_q \subset \mathcal{L}^2(\mathcal{T})$. Lastly, $\epsilon_1, \ldots, \epsilon_n$ are random errors, which are assumed independent, with expectation $\mathrm{E}(\epsilon_i) = 0$ and variance $\mathrm{var}(\epsilon_i) = \sigma^2 < \infty$, for $i = 1, \ldots, n$. The interpretation of the functional parameter $\beta_q(t)$, for $q = 1, \ldots, Q$, are that the times with large absolute value of $\beta_q(t)$ have highest influence on the response (Reiss et al., 2017).

The model given by (1) with univariate functional parameters will often provide a suficiently flexible representation of the data-generating process. It is a special case of the multivariate functional parameter model of Yao and Müller (2010); see Supplementary Material SM1. However, multivariate functional parameters are not straightforward to interpret, hence we focus on the simplification to univariate functional parameters.

2.2 Basis expansions

2.2.1 Functional parameter basis expansion

The functional parameters $\beta_1(t), \ldots, \beta_Q(t)$ in model (1) are infinite-dimensional. Estimation from a finite number, n, of scalar responses, $y_1, \ldots, y_n$, can be achieved by assuming a parametric form via a finite basis expansion (Ramsay and Silverman, 2005, page 44).

For $q = 1, \ldots, Q$, the basis expansion of $\beta_q(t)$ is

$$\beta_q(t) = \sum_{l=1}^{m_{\beta,q}} \theta_{ql} b_{ql}(t) = \boldsymbol{b}_q(t)^{\mathrm{T}} \boldsymbol{\theta}_q,$$

where the functions $b_{q1}(t), \ldots, b_{qm_{\beta,q}}(t)$ are known basis functions and $\boldsymbol{\theta}_q$ is an $m_{\beta,q} \times 1$ vector of unknown coefficients. Consequently, the problem is reduced to estimating $p = \sum_{q=1}^{Q} m_{\beta,q}$ unknown coefficients given by $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_Q)^{\mathrm{T}}$. A special case of a scalar parameter, θ_q, is represented by the single ($m_{\beta,q} = 1$) basis function $b_q(t) = 1$. We can write $\boldsymbol{\beta}(t) = B(t)^{\mathrm{T}} \boldsymbol{\theta}$, where $B(t)$ is a $p \times Q$ block diagonal matrix with qth block given by $\boldsymbol{b}_q(t)$, that is

$$B(t) = \begin{bmatrix} \boldsymbol{b}_1(t) & 0 & \cdots & \cdots \\ 0 & \boldsymbol{b}_2(t) & 0 & \cdots \\ \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \cdots & \boldsymbol{b}_Q(t) \end{bmatrix}.$$

2.2.2 Profile variable basis expansion

In the design of experiments setting, we have the freedom to specify the profile variables $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ to optimize gain in information in estimating the functional parameters $\boldsymbol{\beta}(t)$, or equivalently coefficients $\boldsymbol{\theta}$, from the scalar responses $y_1, \ldots, y_n$. Similar to Section 2.2.1, $\boldsymbol{x}_i(t)$, for $i = 1, \ldots, n$, are infinite-dimensional. To make the optimization problem tractable, we again use a finite basis expansion. For $i = 1, \ldots, n$ and $j = 1, \ldots, J$,

$$x_{ij}(t) = \sum_{l=1}^{m_{x,j}} \gamma_{ijl} c_{jl}(t) = \boldsymbol{\gamma}_{ij}^{\mathrm{T}} \boldsymbol{c}_j(t), \qquad (2)$$

where $c_{j1}(t), \ldots, c_{jm_{x,j}}(t)$ are known basis functions and $\boldsymbol{\gamma}_{ij} = (\gamma_{ij1}, \ldots, \gamma_{ijm_{x,j}})$ are coefficients. Let $\boldsymbol{\Gamma} = (\boldsymbol{\gamma}_{11}, \ldots, \boldsymbol{\gamma}_{nJ})^{\mathrm{T}}$ be the $n \sum_{j=1}^{J} m_{x,j} \times 1$ vector of coefficients.

For observational data, Γ are estimated and treated as fixed in the estimation of θ. Conversely, in design of experiments, Γ are specified to optimize gain in information in estimating θ (see Section 3). A sufficiently complex basis expansion should be chosen so that an optimal choice of Γ leads to the profile variables reconstructed via (2) providing a good approximation to the optimal infinite-dimensional profile variables.

2.3 Standard linear model representation of the scalar-on-function linear model

The key to design of experiments for scalar-on-function linear models is the observation that the model given by (1), under the basis expansions of Section 2.2, can be written in the form of a standard linear model. This observation has been made previously (e.g., Reiss et al., 2017) for the case where $J = 1$ and the regression functional is of the form $f[x_i(t)] = [1, x_i(t)]^{\mathrm{T}}$. Below we extend this representation to $J > 1$, and more general forms of the regression functional.

The qth element, $f_q[\boldsymbol{x}_i(t)]$, of the regression functional can be written as $f_q[\boldsymbol{x}_i(t)] = \prod_{a \in \mathcal{F}_q} x_{ia}(t)$, where $\mathcal{F}_q$ is a set defining the qth term. The members of $\mathcal{F}_q$ belong to the set $\{1, \ldots, J\}$ and can be repeated. For example, the intercept has $\mathcal{F}_q = \varnothing$, the main effect of the jth profile factor has $\mathcal{F}_q = \{j\}$, the quadratic effect of the jth profile factor has $\mathcal{F}_q = \{j, j\}$ and the two-way interaction between the j_1th and j_2th profile factors has $\mathcal{F}_q = \{j_1, j_2\}$.

The model given by (1) can now be written as

$$y_i = \int_{\mathcal{T}} \sum_{q=1}^{Q} \left[\prod_{a \in \mathcal{F}_q} \boldsymbol{\gamma}_{ia}^{\mathrm{T}} \boldsymbol{c}_a(t) \right] \boldsymbol{b}_q(t)^{\mathrm{T}} \boldsymbol{\theta}_q \mathrm{d}t + \epsilon_i. \quad (3)$$

Using properties of Kronecker products (denoted by $\otimes$)

$$\prod_{a \in \mathcal{F}_q} \boldsymbol{\gamma}_{ia}^{\mathrm{T}} \boldsymbol{c}_a(t) = \left[\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{\gamma}_{ia} \right]^{\mathrm{T}} \left[\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{c}_a(t) \right], \quad (4)$$

where $\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{\gamma}_{ia}$ and $\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{c}_a(t)$ are both $\prod_{a \in \mathcal{F}_q} m_{x,a} \times 1$ vectors. Note, we use the convention that a Kronecker product over the empty set, $\varnothing$, is one.

Substituting (4) into (3) gives $y_i = \sum_{q=1}^{Q} \boldsymbol{z}_{iq}^{\mathrm{T}} \boldsymbol{\theta}_q + \epsilon_i$, where

$$\boldsymbol{z}_{iq} = \int_{\mathcal{T}} \boldsymbol{b}_q(t) \left[\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{c}_a(t)^{\mathrm{T}} \right] \mathrm{d}t \left[\bigotimes_{a \in \mathcal{F}_q} \boldsymbol{\gamma}_{ia} \right], \quad (5)$$

is an $m_{\beta,q} \times 1$ vector. Therefore, the model can be written in the form of a standard linear model as $\boldsymbol{y} = Z\boldsymbol{\theta} + \boldsymbol{\epsilon}$, where $\boldsymbol{y} = (y_1, \ldots, y_n)$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$, and Z is an $n \times p$ model matrix given by

$$Z = \begin{pmatrix} \boldsymbol{z}_{11}^{\mathrm{T}} & \cdots & \boldsymbol{z}_{1Q}^{\mathrm{T}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{z}_{n1}^{\mathrm{T}} & \cdots & \boldsymbol{z}_{nQ}^{\mathrm{T}} \end{pmatrix}.$$

Note that Z is a function of the coefficients, $\Gamma = (\boldsymbol{\gamma}_{11}, \ldots, \boldsymbol{\gamma}_{nJ})^{\mathrm{T}}$, of the basis expansion of the profile variables. However, this dependence is suppressed in the notation for clarity.

From (5), $\boldsymbol{z}_{iq} = R_q \bigotimes_{a \in F_q} \boldsymbol{\gamma}_{ia}$ where $R_q = \int_{\mathcal{T}} \boldsymbol{b}_q(t) \bigotimes_{a \in \mathcal{F}_q} \boldsymbol{c}_a(t)^{\mathrm{T}} \mathrm{d}t$ is an $m_{\beta,q} \times \prod_{a \in \mathcal{F}_q} m_{x,a}$ matrix. The elements of R_q are integrals of products of basis functions. The exact form of the integrands will depend on the choice of basis functions for both the $\boldsymbol{x}_i(t)$'s and $\boldsymbol{\beta}_q(t)$'s. However, for any combination of monomial and B-spline basis functions, we show, in Section SM2 of the Supplementary Material, that the integration can be performed in closed form. We also provide a closed form expression for the elements of R_q.

3 OPTIMAL DESIGN FOR SCALAR-ON-FUNCTION LINEAR MODELS

3.1 Basis expansions and constraints

The aim is to specify the vector of profile variable coefficients Γ (subsequently referred to as the design) to optimize gain in information in estimating the unknown functional parameter coefficients θ. This will be achieved by minimising an objective function $\Psi(\Gamma)$. The choice of objective function will be addressed in Section 3.2.

However, beforehand, there are several issues that need discussion, in regards to the basis expansion of both the functional parameters and profile variables, and the complexity of the model specified by the regression functional, $\boldsymbol{f}(\cdot)$. These matters will need addressing before the experiment, and therefore can be considered as part of the design process.

The basis functions, $\boldsymbol{b}_1(t), \ldots, \boldsymbol{b}_Q(t)$ of the functional parameter expansion need specification. For these functions, we consider monomials and B-splines. Any combination of these, combined with B-splines for the profile variable basis expansion, leads to the integrals in the model matrix Z being available in closed form (as discussed in Section 2.3). Monomial basis functions have the advantage of higher interpretability but B-splines are more computationally efficient.

The complexity of the scalar-on-function linear model needs to be specified via the regression functional, $\boldsymbol{f}(\cdot) = [f_1(\cdot), \ldots, f_Q(\cdot)]^{\mathrm{T}}$, and the number of basis functions, $m_{\beta,1}, \ldots, m_{\beta,Q}$. Specifying the complexity of a statistical model before observing the responses is a challenge shared by all applications of optimal design of experiments (e.g., Atkinson et al., 2007, page 329). It is a subjective decision which should be made in consultation with the researchers, considering the number of runs, n, in the experiment and, potentially, the estimability of $\boldsymbol{\theta}$. On this last point, for $Z^{\mathrm{T}}Z$ to be non-singular, and the least squares estimates of $\boldsymbol{\theta}$ to exist, the $n \times p$ model matrix Z requires full column-rank. A necessary, but not sufficient, condition for this to occur is that $p = \sum_{q=1}^{Q} m_{\beta,q} \leq n$. This places an upper bound on the complexity $(m_{\beta,1}, \ldots, m_{\beta,Q})$ of the basis expansion of the functional parameters, based on the available experimental resources (n).

Similarly, the functions, $\boldsymbol{c}_1(t), \ldots, \boldsymbol{c}_J(t)$, and numbers, $m_{x,1}, \ldots, m_{x,J}$, of the profile variable basis expansion need specification. Ultimately, there will be a compromise between specifying profile variables $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_n(t)$ that are simple enough to be implemented in a physical experiment but

complex enough to provide a good approximation to the actual optimal infinite-dimensional profile variables. The former can be determined in consultation with the researchers. The latter can be investigated as part of minimizing the objective function $\Psi(\cdot)$ (see Sections 4 and 5 for examples).

For the basis functions, $c_1(t), \ldots, c_J(t)$, the proposal is to use B-splines. These offer two distinct advantages. Firstly, in combination with B-spline or monomial basis functions for the functional parameter basis expansions, the integrals in the model matrix are available in closed form. Secondly, as shown below, it is straightforward to implement constraints on the profile variables $x_i(t)$, for $i = 1, \ldots, n$, via simple constraints on the elements of design Γ.

Often, the profile variable $x_{ij}(t)$, for $i = 1, \ldots, n$ and $j = 1, \ldots, J$, will be constrained in the experiment such that $x_{ij} : \mathcal{T} \to [a_j, b_j]$. For example, if $x_{ij}(t)$ represented the temperature at time $t \in \mathcal{T}$ for the ith run, then a_j and b_j are the minimum and maximum possible temperatures, respectively.

Suppose, for $j = 1, \ldots, J$, the basis functions $c_{j1}(t), \ldots, c_{jm_{x,j}}(t)$ are B-spline functions. Using the property of B-spline functions that $\sum_{l=1}^{m_{x,j}} c_{jl}(t) = 1$ for $t \in \mathcal{T}$, if the elements of γ_{ij} are such that $\gamma_{ijl} \in [a_j, b_j]$, for all $l = 1, \ldots, m_{x,j}$, it follows that $x_{ij}(t) = \sum_{l=1}^{m_{x,j}} \gamma_{ijl} c_{jl}(t) \in [a_j, b_j]$ for $t \in \mathcal{T}$.

To implement B-spline basis functions for $x_{ij}(t)$, the number, k_j, and location of the internal knots need to be specified. It is common to choose uniformly spaced internal knots and we use this approach in this paper. However, the location of the knots could be incorporated into the design problem. For maximally smooth basis functions, $m_{x,j} = k_j + d_j + 1$, where d_j is the degree. Thus the number of internal knots is specified by the number of basis functions, $m_{x,j}$, and the degree. The following lemma can aid in choosing the number of basis functions. It gives a lower bound on the complexity of the profile variables for the chosen complexity of the scalar-on-function linear model. The proof is provided in Section SM3 of the Supplementary Material.

Lemma 1: *A necessary, but not sufficient condition, for the least squares estimates of the functional parameter coefficients, θ, to exist, is $m_{\beta,q} \leq \prod_{a \in \mathcal{F}_q} m_{x,a}$ for all $q = 1, \ldots, Q$.*

Lemma 1 states that the degree of complexity of the basis functions for the functional parameters demands a minimum degree of complexity in the basis functions for the associated profile variables. However, this result says nothing about the maximum degree of complexity for the profile variables. From empirical evidence, for example from the illustrative example in Section 4, it appears that continuing to increase the complexity of the profile variables whilst keeping the complexity of the functional parameters, and amount of experimental resources (n), fixed leads to a, perhaps anticipated, plateau in the expected gain in information from the experiment.

3.2 Objective functions via Bayesian decision-theory

We formulate objective functions to find designs by using a Bayesian decision-theoretic approach. Bayesian decision-theoretic approaches for optimal experimental design are well established in the broader literature (e.g., Chaloner and Verdinelli, 1995; Rainforth et al., 2024; Ryan et al., 2016). Our contribution is to develop this framework for scalar-on-function linear models, which to our knowledge has not previously been addressed. The Bayesian approach has the advantages that the experimental goal can be incorporated via the choice of loss function and all assumed sources of uncertainty are accounted. The post-experiment analysis does not necessarily need to be undertaken under a Bayesian approach.

Bayesian decision-theoretic design (or Bayesian design) starts with specification of a loss function, denoted $\ell(\theta, y, \Gamma)$, giving the loss in estimating parameters θ, using responses y obtained from an experiment with design Γ. The objective function, $\Psi(\Gamma)$, follows from the expected loss $L(\Gamma) = \mathrm{E}_{y,\theta|\Gamma}[\ell(\theta, y, \Gamma)]$, where expectation is with respect to the joint probability distribution of the responses y and parameters θ. This joint probability distribution of y and θ is given by the distribution of the responses y given θ, and the prior distribution of θ. Note that the expected loss can always be written as $L(\Gamma) = g[\Psi(\Gamma)]$, where $g(\cdot)$ is a monotonically increasing function. Therefore, it is sufficient to minimize the objective function $\Psi(\Gamma)$ in place of $L(\Gamma)$.

The Bayesian design is given by minimizing the objective function $\Psi(\Gamma)$ over the space of all designs. Recall from Section 3.1, that if the profile variables are constrained such that $x_{ij} : \mathcal{T} \to [a_j, b_j]$, for $i = 1, \ldots, n$ and $j = 1, \ldots, J$, then it is sufficient for $\gamma_{ijl} \in [a_j, b_j]$, for $l = 1, \ldots, m_{x,j}$, $i = 1, \ldots, n$ and $j = 1, \ldots, J$. This defines the space of all designs.

For the scalar-on-function linear model, the distribution of y given θ follows from specification of a distribution for the errors $\epsilon_1, \ldots, \epsilon_n$. For analytical tractability, we assume that the errors are normally distributed. For the prior distribution of θ, we assume a multivariate normal distribution with mean μ and variance $\sigma^2 \Sigma$, where μ and Σ are a specified $p \times 1$ vector and $p \times p$ matrix, respectively. The matrix Σ is termed the prior scale matrix. We also assume an inverse-gamma prior distribution for the nuisance parameter of the error variance σ^2. The shape and rate parameters are $a/2$ and $b/2$, respectively, where a and b are specified scalar constants.

Let $\hat{\mu} = \hat{\Sigma}(Z^{\mathrm{T}} y + \Sigma^{-1} \mu)$, $\hat{\Sigma} = (Z^{\mathrm{T}} Z + \Sigma^{-1})^{-1}$, $\hat{a} = a + n$ and $\hat{b} = b + \mu^{\mathrm{T}} \Sigma^{-1} \mu + y^{\mathrm{T}} y - \hat{\mu}^{\mathrm{T}} \hat{\Sigma}^{-1} \hat{\mu}$. Under the above prior specification, the marginal posterior distribution of θ is a multivariate t-distribution with mean $\hat{\mu}$, variance $\hat{b} \hat{\Sigma} / (\hat{a} - 2)$ and $\hat{a}$ degrees of freedom.

There are several prior hyperparameters, μ, Σ, a and b that require specification. However, for the loss functions considered in this paper, the expected loss only depends on the prior scale matrix Σ. There are various approaches to specifying Σ. It can be specified using prior information: either subjectively via prior elicitation (e.g., O'Hagan and Forster, 2004, Chapter 6) or using the results of a previous experiment, that is use $\hat{\Sigma} := \Sigma$. The prior scale matrix can also be specified to represent prior ignorance. Specifying Σ^{-1} to be the $p \times p$ matrix of zeros implements a uniform prior distribution for θ. Under such a prior distribution, the posterior distribution still exists (i.e., is proper), as long as $p = \sum_{q=1}^{Q} m_{\beta,q} \leq n$. Also, as discussed below, under such a prior distribution, and choice of certain loss function, the expected loss objective function is equivalent to commonly-

used objective functions from the classical design literature (e.g., Chaloner and Verdinelli, 1995). Lastly, Σ can be used to implement a roughness penalty on the functional parameter basis functions (see Section 6 and Section SM7 of the Supplementary Material).

In Sections 4 and 5, we consider the squared error (SE) loss function given by the sum of squared differences of the elements of $\boldsymbol{\theta}$ and their marginal posterior means, that is

$$\ell_{SE}(\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{\Gamma}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\mu}})^{\mathrm{T}} (\boldsymbol{\theta} - \hat{\boldsymbol{\mu}}).$$

It follows (see Section SM4 in the Supplementary Material for derivation) that the SE objective function given by the expected SE loss is $\Psi_{SE}(\boldsymbol{\Gamma}) = \mathrm{tr}(\hat{\Sigma})$. Note that finding the Bayesian SE design does not require the prior mean, $\boldsymbol{\mu}$, nor the prior shape, a, or rate, b, parameters to be specified.

If an improper uniform prior distribution is assumed for $\boldsymbol{\theta}$, whereby the elements of Σ^{-1} are zero, then the SE objective function reduces to $\Psi_{SE}(\boldsymbol{\Gamma}) = \mathrm{tr}\left[(Z^{\mathrm{T}}Z)^{-1}\right]$, and the corresponding Bayesian design is equivalent to the classical A-optimal design.

As presented above, the loss function is in terms of the functional parameter coefficients $\boldsymbol{\theta}$. However, these may not be of direct interest. Rather, interest may lie in implied functions $\boldsymbol{\beta}(t) = \left[\beta_1(t), \ldots, \beta_Q(t)\right]^{\mathrm{T}}$. Therefore, we propose an alternative loss, in terms of $\boldsymbol{\theta}$, that is given by integrating a loss, in terms of $\boldsymbol{\beta}(t)$, over $\mathcal{T}$.

As an example, consider the SE loss in terms of $\boldsymbol{\beta}(t)$ given by

$$\ell_{SE}(\boldsymbol{\beta}(t), \boldsymbol{y}, \boldsymbol{\Gamma}) = \left[\boldsymbol{\beta}(t) - \mathrm{E}_{\boldsymbol{\beta}(t)|\boldsymbol{y}, \boldsymbol{\Gamma}}(\boldsymbol{\beta}(t))\right]^{\mathrm{T}}$$
$$\left[\boldsymbol{\beta}(t) - \mathrm{E}_{\boldsymbol{\beta}(t)|\boldsymbol{y}, \boldsymbol{\Gamma}}(\boldsymbol{\beta}(t))\right],$$

where $\mathrm{E}_{\boldsymbol{\beta}(t)|\boldsymbol{y}, \boldsymbol{\Gamma}}\left[\boldsymbol{\beta}(t)\right]$ is the marginal posterior mean of $\boldsymbol{\beta}(t)$. Using properties of the multivariate t-distribution, the posterior distribution of $\boldsymbol{\beta}(t)$, for $t \in \mathcal{T}$, is multivariate t-distribution with posterior mean $\mathrm{E}_{\boldsymbol{\beta}(t)|\boldsymbol{y}, \boldsymbol{\Gamma}}\left[\boldsymbol{\beta}(t)\right] = B(t)^{\mathrm{T}}\hat{\boldsymbol{\mu}}$. The implied loss in terms of $\boldsymbol{\theta}$ is now

$$\ell_{WSE}(\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{\Gamma}) = \int_{\mathcal{T}} \ell_{SE}(\boldsymbol{\beta}(t), \boldsymbol{y}, \boldsymbol{\Gamma}) \, \mathrm{d}t$$
$$= \int_{\mathcal{T}} \left[B(t)^{\mathrm{T}}\boldsymbol{\theta} - B(t)^{\mathrm{T}}\hat{\boldsymbol{\mu}}\right]^{\mathrm{T}}$$
$$\left[B(t)^{\mathrm{T}}\boldsymbol{\theta} - B(t)^{\mathrm{T}}\hat{\boldsymbol{\mu}}\right] \, \mathrm{d}t$$
$$= (\boldsymbol{\theta} - \hat{\boldsymbol{\mu}})^{\mathrm{T}} B_I (\boldsymbol{\theta} - \hat{\boldsymbol{\mu}}),$$

where $B_I = \int_{\mathcal{T}} B(t)B(t)^{\mathrm{T}} \, \mathrm{d}t$ is a $p \times p$ matrix. Thus the implied loss in terms of $\boldsymbol{\theta}$ is a weighted squared error (WSE) loss. The corresponding objective function is $\Psi_{WSE}(\boldsymbol{\Gamma}) = \mathrm{tr}\left[B_I \hat{\Sigma}\right]$ (see Section SM4 in the Supplementary Material for derivation). Under an improper uniform prior distribution for $\boldsymbol{\theta}$, the Bayesian WSE design is equivalent to a classical L-optimal design (e.g., Atkinson et al., 2007, Section 10.5).

4 ILLUSTRATIVE EXPERIMENT WITH A SINGLE PROFILE VARIABLE

In this section, we consider an illustrative example where Bayesian designs are found for an experiment with $J = 1$ pro-

file variable. The time interval is $\mathcal{T} = [0, 1]$, for example, $T = 1$. The regression functional is $\boldsymbol{f}\left[x_{i1}(t)\right] = [1, x_{i1}(t)]$, so that $Q = 2$ and the model is $y_i = \beta_1 + \int_0^1 \beta_2(t)x_{i1}(t) \, \mathrm{d}t + \epsilon_i$, for $i = 1, \ldots, n$.

For the basis expansion of $x_{i1}(t)$, we consider degree-zero B-spline basis functions with uniformly spaced internal knots. If there are $m_{x,1}$ basis functions, then this amounts to each $x_{i1}(t)$ being a step function on $\mathcal{T} = [0, 1]$, where the discontinuities correspond to the internal knots $\lambda_{1,1}, \ldots, \lambda_{1,k_1}$, where $k_1 = m_{x,1} - 1$. Explicitly, for $l = 1, \ldots, m_{x,1}$,

$$c_{1l}(t) = \begin{cases} 1 \text{ if } t \in [\lambda_{1,l-1}, \lambda_{1,l}); \\ 0 \text{ if otherwise;} \end{cases}$$

where $\lambda_{1,0} = 1$ and $\lambda_{1,m_{x,1}} = 1$.

Step functions are an important and useful tool in practice. For instance, Rameez et al. (2014) considered an Ambr bioreactor application with a temperature step shift from the high to the low level. We will vary the number of basis functions, $m_{x,1}$, which is equivalent to varying the complexity of the profile variable, to investigate the effect on the performance of the design through the expected loss.

For the basis expansion of the functional parameters, β_1 is a constant so $m_{\beta,1} = 1$ and $b_1(t) = 1$. For $\beta_2(t)$, we consider two cases with monomial basis functions. In Case I, $m_{\beta,2} = 2$, $\boldsymbol{b}_2(t) = (1, t)^{\mathrm{T}}$, giving $p = 3$. In Case II, $m_{\beta,2} = 3$, $\boldsymbol{b}_2(t) = (1, t, t^2)^{\mathrm{T}}$, giving $p = 4$.

For both cases, we find Bayesian SE and WSE designs with number of basis functions for the profile variable $m_{x,1} = 2$ (not Case II), 3, 4, 8, 16, 100 and $n = 4, 12$. Note that $m_{x,1} = 2$ is forbidden for Case II since the least squares estimate of $\boldsymbol{\theta}$ does not exist for this value (see Lemma 1). Section SM5 of the Supplementary Material gives expressions for the two matrices, R_1 and R_2, which feature in the model matrix Z, and the weighting matrix, B_I, for Bayesian WSE designs.

The design, $\boldsymbol{\Gamma}$, is a vector with $n\sum_{j=1}^{J} m_{x,j} = nm_{x,1}$ elements. The profile variable $x_{i1}(t)$ is assumed to be constrained to $[-1, 1]$ meaning the elements of $\boldsymbol{\Gamma}$ are also constrained to $[-1, 1]$ (see Section 3.1). In each scenario, the coordinate exchange algorithm (Meyer and Nachtsheim, 1995) is used to minimize the objective function from 1000 random starts, as suggested by Goos and Jones (2011).

The methodology is implemented in the **fdesigns** R package (Michaelides et al., 2025). The Supplementary Material includes a vignette that demonstrates how to find the designs discussed in this section and in Section 5.

Table 1 shows the values of the SE and WSE objective function for the respective optimal designs for each value of $m_{x,1}$ and n. Clearly as n increases, the expected loss decreases. Typically, as the number of basis functions, $m_{x,1}$, increases, the expected loss decreases, since the finite-dimensional representation of the profile variable becomes a better approximation to the optimal infinite-dimensional profile variable. The rate of this decrease slows as $m_{x,1}$ increases allowing a compromise to be made between profile variable complexity and gain in information. This sensitivity to $m_{x,1}$ also depends on the choice of functional parameter basis. In Case I, even small values of $m_{x,1}$ already achieve near-minimal expected loss, so relatively simple profile-variable

TABLE 1 SE and WSE objective function values for optimal designs for $m_{x,1} = 2, 3, 4, 8, 16, 100$ and $n = 4, 12$. The profile variables for the scenario highlighted in bold are shown in Figure 1.

| | SE | | | | WSE | | | |
| | Case I | | Case II | | Case I | | Case II | |
$m_{x,1}$	$n = 4$	$n = 12$	$n = 4$	$n = 12$	$n = 4$	$n = 12$	$n = 4$	$n = 12$
2	8.750	2.583	–	–	1.417	0.472	–	–
3	8.828	2.778	386.408	126.409	1.581	0.499	3.363	1.120
4	8.750	2.570	246.869	67.735	1.417	0.472	3.243	1.022
8	**8.493**	2.539	218.479	65.217	1.417	0.472	3.147	1.021
16	8.427	2.520	208.843	63.610	1.417	0.472	3.099	1.016
100	8.404	2.512	206.884	63.028	1.417	0.472	3.094	1.010

bases provide designs that are close to optimal. In Case II, the designs are more sensitive. Increasing $m_{x,1}$ leads to more noticeable reductions in expected loss, and somewhat larger bases are required. Nevertheless, once a moderate number of basis functions is reached, further increases yield little additional gain. These results reinforce the discussion in Section 3.1, highlighting the importance of aligning the complexity of the profile variable basis with that of the functional-parameter basis to avoid unnecessary increases in dimensionality and computational burden. The exception to expected loss decreasing with $m_{x,1}$ is for $m_{x,1} = 2$ for Case I. The uniformly-spaced internal knots for $m_{x,1} = 2, 3, 4$ are $\{\frac{1}{2}\}$, $\{\frac{1}{3}, \frac{2}{3}\}$ and $\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$, respectively. A change at the knot at $t = \frac{1}{2}$ is important for small $m_{x,1}$ and its absence for $m_{x,1} = 3$, leads to higher expected loss than for $m_{x,1} = 2$.

Case II naturally yields larger expected-loss values than Case I (see Table 1), because the number of parameters increases with the quadratic term in the functional parameter basis. For this reason, absolute magnitudes of the objective functions are not directly comparable across cases. Instead, the focus should be on the structural features of the resulting profile variables. In Case I, most functions have a single step change, with some allowing two step changes. In contrast, in Case II, the designs have two step changes in most functions, with some profiles having three step changes. This additional flexibility reflects the need to capture the curvature of the quadratic parameter functions. These results indicate that the robustness of the design depends on the modeling choice, that is Case I designs are simpler and more stable, whereas Case II designs require more complex step functions which are still stable once sufficient flexibility is introduced.

As an example, Figure 1 shows the profile variables $x_{11}(t), \ldots, x_{41}(t)$ for the $n = 4$ SE design for Case I with $m_{x,1} = 8$ basis functions. This scenario is highlighted in bold in Table 1. Three profile variables have exactly one change and the remaining has exactly two changes.

5 APPLICATION TO A DYNAMIC EXPERIMENT IN THE AMBR250 BIOREACTOR

We now apply the methodology to a real Ambr250 bioreactor experiment. GSK are investigating the relationship between titer content of a product (constrained to $(0,1)$) and $J = 4$ variables: feed volume (denoted $x_{i1}(t)$); IVCC (x_{i2}); pH (x_{i3}); and temperature (x_{i4}). Feed volume is a profile variable while the remaining variables are static, that is $m_{x,2} = m_{x,3} = m_{x,4} = 1$. The response, y_i, is the logit of titer content. GSK are able to run a dynamic experiment on the (scaled) time interval $\mathcal{T} = [0, 1]$ with $n = 12$ runs.

In consultation with GSK, the following regression functional is proposed

$$\boldsymbol{f}[\boldsymbol{x}_i(t)] = \left[1, x_{i1}(t), x_{i2}, x_{i3}, x_{i4}, x_{i2}^2, x_{i3}^2, x_{i4}^2\right]^{\mathrm{T}},$$

giving $Q = 8$. For all $q = 1, \ldots, Q$, apart from $q = 2$, $m_{\beta,q} = 1$ giving constant parameters. For the functional parameter $\beta_2(t)$, we consider the monomial basis function $\boldsymbol{b}_2(t) = (1, t, t^2)^{\mathrm{T}}$, that is $m_{\beta,2} = 3$. This leaves a model with $p = \sum_{q=1}^{Q} m_{\beta,q} = 10$.

For the profile variable $x_{i1}(t)$, for $i = 1, \ldots, n$, GSK are able to implement step functions with at most three changes during $\mathcal{T}$. Thus we consider zero-degree B-splines with $m_{x,1} = 4$ (allowing up to three changes). For simplicity, we use uniformly-spaced internal knots.

For the matrices $R_1, \ldots, R_Q$ that feature in the model matrix, following the same derivation as in Section SM5 of the Supplementary Material, $R_q = 1$, for $q \neq 2$, are all scalars, and R_2 is a $m_{\beta,2} \times m_{x,1} = 3 \times 4$ matrix given by

$$R_2 = \frac{1}{192} \begin{pmatrix} 48 & 48 & 48 & 48 \\ 6 & 18 & 30 & 42 \\ 1 & 7 & 19 & 37 \end{pmatrix}.$$

After discussion with collaborators at GSK, it was decided that focus would be on estimating the $\boldsymbol{\theta}$ parameters rather than $\boldsymbol{\beta}(t)$, motivating the goal of finding an Bayesian SE design. Also after discussion, it was decided to use a prior distribution for $\boldsymbol{\theta}$ that represented prior ignorance, in particular an improper uniform prior distribution. This means the Bayesian SE design is equivalent to a classical A-optimal design.

The design $\boldsymbol{\Gamma}$ has a total of $n \sum_{j=1}^{J} m_{x,j} = 84$ elements. All controllable variables are scaled to $[-1, 1]$ meaning elements of $\boldsymbol{\Gamma}$ are constrained to the same interval. The Bayesian SE design is found using coordinate exchange, as in Section 4. Table 2 shows the optimal design for the $n = 12$ runs. Five unique functions for feed volume, $x_{i1}(t)$, were found, labeled (a)–(e) in Table 2 and shown in Figure 2. Function (a) is repeated four times, function (b) repeated five times, and functions (c), (d), and (e) once. The optimal design for the scalar variables includes boundary points
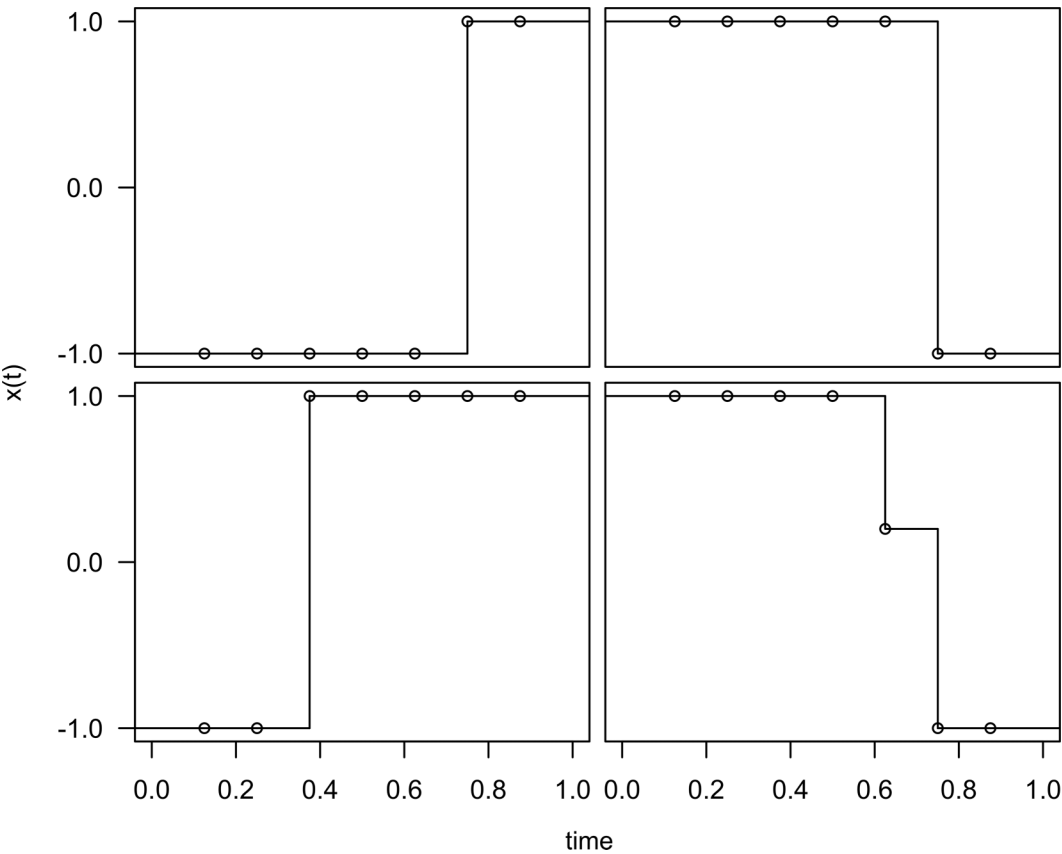
FIGURE 1 Profile variables for the $n = 4$ SE design for Case I with $m_{x,1} = 8$ basis functions.

TABLE 2 SE design for the Ambr250 bioreactor experiment. The feed volume profile variables labeled (a)-(e) are shown in Figure 2.

Run i	Feed volume $x_{i1}(t)$	IVCC x_{i2}	pH x_{i3}	Temperature x_{i4}
1	(a)	0	−1	1
2	(a)	1	1	1
3	(b)	0	1	−1
4	(c)	0	0	0
5	(d)	0	0	0
6	(a)	−1	1	0
7	(e)	0	0	0
8	(b)	1	−1	0
9	(b)	−1	0	1
10	(a)	1	0	−1
11	(b)	1	1	0
12	(b)	0	0	1

and center points in order to be able to estimate the parameters associated with the quadratic terms. This is similar to the behavior of the quadratic function for the profile variable, where the functions of the profile variable changes at most twice.

As a sensitivity analysis, we find the SE optimal design for $m_{x,1} = 3, 4, 8, 16, 100$. The value of the objective function is shown in Table SM1 in the Supplementary Material. There is limited gain in information in increasing $m_{x,1}$ beyond $m_{x,1} = 4$. However, if $m_{x,1} = 3$, then there is a large increase in the expected loss. From Figure 2, the optimal profile variables make at most two changes over $\mathcal{T}$. Therefore, the increase in expected

loss for $m_{x,1} = 3$ is due to the position of the internal knots rather than their number.

On running the experiment, there were two complications. Firstly, run $i = 7$ failed to return a titer value, so there are only responses from the remaining eleven runs. Secondly, the Ambr250 experimental apparatus used by GSK provides titer content to three decimal places. For run $i = 1$, a titer content of one was returned. This means the true observed titer content was in the interval $[0.9995, 1)$, meaning the true observed response y_1 (after logit transform) was in the interval $[7.600, \infty)$ (to three decimal places). To address this complication, we assumed that the true
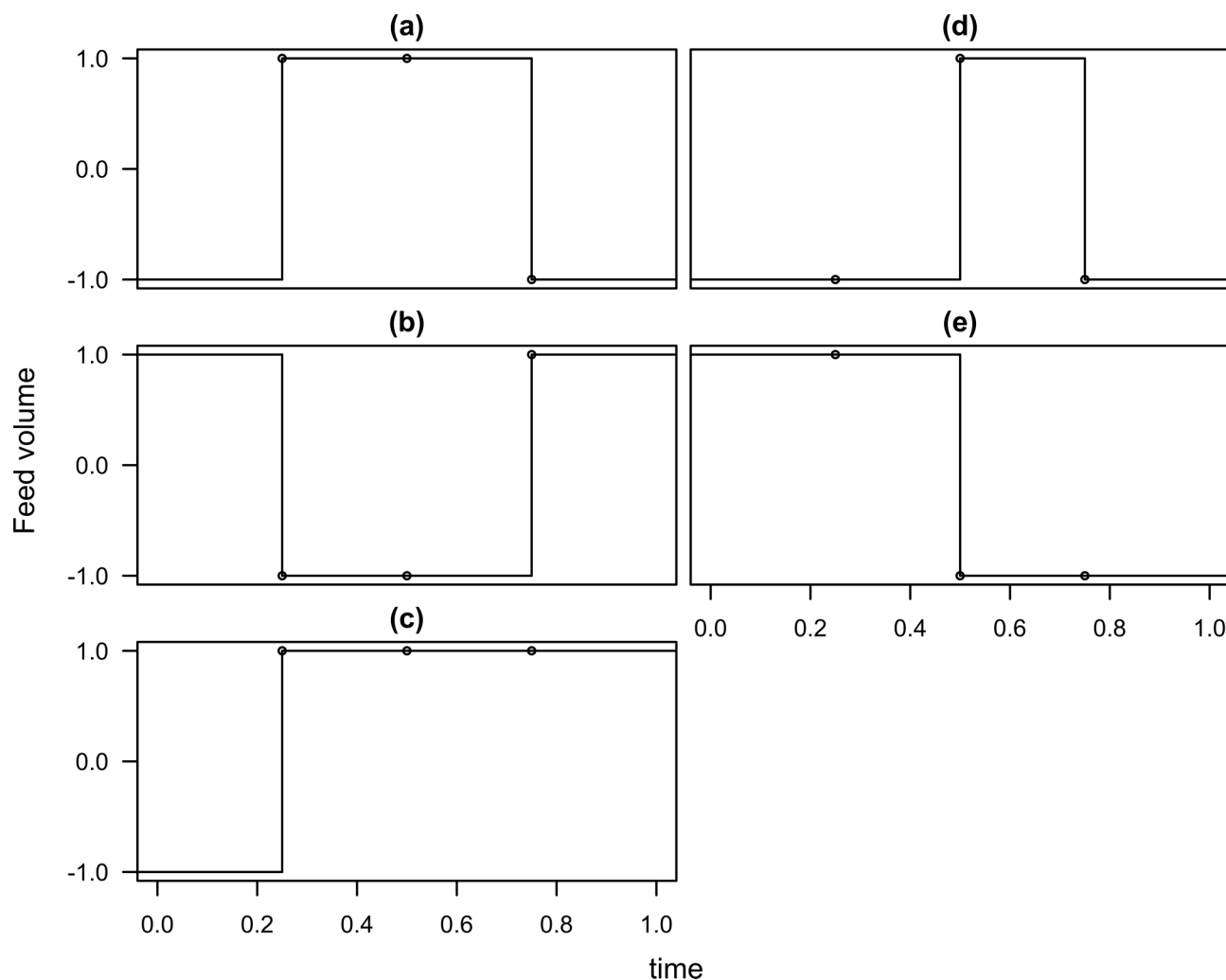
FIGURE 2 Five unique Feed volume profile variables of the Bayesian SE design for the Ambr250 bioreactor experiment.

response was censored and estimated $\boldsymbol{\theta}$ (and σ^2) via maximum likelihood. This means the likelihood contribution for runs $i > 1$ is $\phi \left\{ \left(y_i - \int_{\mathcal{T}} \boldsymbol{f}[\boldsymbol{x}_i(t)] \right)^{\mathrm{T}} \boldsymbol{\beta}(t) \mathrm{d}t \right) / \sigma \right\}$, and for run $i = 1$ is $1 - \Phi \left\{ \left(7.600 - \int_{\mathcal{T}} \boldsymbol{f}[\boldsymbol{x}_i(t)] \right)^{\mathrm{T}} \boldsymbol{\beta}(t) \mathrm{d}t \right) / \sigma \right\}$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density and cumulative distribution functions, respectively, of the standard normal distribution.

To investigate the relationship between the variables and titer content, we used model selection. Specifically, the model that minimized the Akaike information criterion (AIC), found using backward selection, retained main effects for IVCC, pH and temperature and a quadratic effect for temperature, with regression functional $\boldsymbol{f}[\boldsymbol{x}_i(t)] = (1, x_{i2}, x_{i3}, x_{i4}, x_{i4}^2)^{\mathrm{T}}$. The profile variable of feed volume was found not to significantly affect titer content. The final estimated fitted model is

$$\widehat{\mathrm{logit}(\mathrm{titer})} = 1.199 + 1.324 \, \mathrm{pH} + 1.684 \, \mathrm{IVCC}$$
$$+ \, 0.506 \, \mathrm{Temperature} - 1.851 \, \mathrm{Temperature}^2,$$

from which, in scaled units, optimal titer content is estimated to occur when pH is $x_2 = 1$, IVCC is $x_3 = 1$ and temperature is $x_4 = 0.137$ (to three decimal places).

6 DISCUSSION

In this paper, new methodology for finding optimal designs has been proposed for dynamic experiments where a scalar response depends on profile variables. The methodology uses basis function expansions of the profile variables and the functional parameters in a scalar-on-function linear model. It is flexible and can be applied assuming a variety of different expansions for both profile variables and parameter functions.

As discussed in Section 3.1, the complexity of the model needs to be specified before the experiment. Part of that specification involves the number of basis functions for the expansion of the functional parameters. An approach is to set the number of basis functions as large as possible and to penalize via a roughness penalty (Ramsay and Silverman, 2005, Chapter 5) implemented through the prior variance matrix, Σ. Details of this are given in Section SM7 of the Supplementary Material.

Ongoing work is to extend the current methodology to non-normally distributed responses and finding designs for scalar-on-function generalized linear models. In this case, a challenge will be the objective function given by the expected loss will not be available in closed form. However, it is antici-

pated that computational advancements in Bayesian optimal design over the last decade (e.g., Rainforth et al., 2024) can be employed.

In this paper it is assumed that time $t \in [0, \mathcal{T}]$ is the continuous single input to profile variables. The methodology can be naturally extended to scenarios where profile variables have multiple inputs, such as in spatio-temporal studies.

A limitation of Bayesian optimal design is that it requires assumptions about the data-generating process to be made prior to observing responses. Gibbs optimal design (Overstall et al., 2025) is an alternative approach to experimental design which has been developed to be less sensitive to such assumptions. This is a further avenue of future research.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web Appendices and R code referenced in Sections 4 and 5 are available with this paper at the *Biometrics* website on Oxford Academic. The supplementary archive `SM_Code.zip` contains the scripts used to generate the optimal designs in Section 4 (illustrative examples) and Section 5 (Ambr250 application).

The Ambr250 data are proprietary to GlaxoSmithKline and cannot be shared, however, the R code illustrates the optimal design construction process. The proposed methodology is implemented in the R package **fdesigns** (version 1.2), available on CRAN at https://cran.r-project.org/package=fdesigns.

FUNDING

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY

The Ambr250 data that support the findings in this paper are proprietary to GlaxoSmithKline and cannot be shared.

REFERENCES

Atkinson, A., Donev, A. and Tobias, R. (2007). *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press, 2 edition.

Braniff, N. and Ingalls, B. (2018). New opportunities for optimal design of dynamic experiments in systems and synthetic biology. *Current Opinion in Systems Biology*, 9, 42–48.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 10, 273–304.

Espie, D. and Macchietto, S. (1989). The optimal design of dynamic experiments. *AIChE Journal*, 35, 223–229.

Georgakis, C. (2013). Design of dynamic experiments: A data-driven methodology for the optimization of time-varying processes. *Industrial and Engineering Chemistry Research*, 52, 12369–12382.

Georgakis, C. (2024). Dynamic data-driven models for complex pharmaceutical reactions - the dynamic response surface methodology. *Current Opinion in Chemical Engineering*, 45, 101045.

Goos, P. and Jones, B. (2011). *Optimal Design of Experiments: A Case Study Approach*. Hoboken, New Jersey: John Wiley and Sons.

Lu, F., Toh, P. C., Burnett, I., Li, F., Hudson, T., Amanullah, A. et al. (2013). Automated dynamic fed-batch process and media optimization for high productivity cell culture process development. *Biotechnology and Bioengineering*, 110, 191–205.

Meyer, R. K. and Nachtsheim, C. J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics*, 37, 60–69.

Michaelides, D., Overstall, A. and Woods, D. (2025). *fdesigns: Optimal Experimental Designs for Dynamic/Functional Models*. R package version 1.2.

O'Hagan, A. and Forster, J. (2004). *Kendall's Advanced Theory of Statistics, Volume 2B, Bayesian Inference*. London, UK: Arnold.

Overstall, A., Holloway-Brown, J. and McGree, J. (2025). Gibbs optimal design of experiments. *Statistical Science*, To appear.

Rainforth, T., Foster, A., Ivanova, D. and Smith, F. (2024). Modern Bayesian experimental design. *Statistical Science*, 39, 100–114.

Rameez, S., Mostafa, S. S., Miller, C. and Shukla, A. A. (2014). High-throughput miniaturized bioreactors for cell culture process development: Reproducibility, scalability, and control. *Biotechnology Progress*, 30, 718–727.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis*. New York, NY, USA: Springer, 2nd edition.

Reiss, P. T., Goldsmith, J., Shang, H. L. and Ogden, R. T. (2017). Methods for scalar-on-function regression. *International Statistical Review*, 85, 228–249.

Roche, A. (2018). Local optimization of black-box functions with high or infinite-dimensional inputs: application to nuclear safety. *Computational Statistics*, 33, 467–485.

Ryan, E. J., Drovandi, C., McGree, J. M. and Pettitt, A. N. (2016). A review of modern computational algorithms for Bayesian optimal design. *International Statistical Review*, 84, 128–154.

Tai, M., Ly, A., Leung, I. and Nayar, G. (2015). Efficient high-throughput biological process characterization: definitive screening design with the Ambr250 bioreactor system. *Biotechnology Progress*, 31, 1388–1395.

Titterington, D. (1980). Aspects of optimal design in dynamic systems. *Technometrics*, 22, 287–299.

Trummer, E., Fauland, K., Seidinger, S., Schriebl, K., Lattenmayer, C., Kunert, R. et al. (2006). Process parameter shifting: Part I. Effect of DOT, pH, and temperature on the performance of Epo-Fc expressing CHO cells cultivated in controlled batch bioreactors. *Biotechnology and Bioengineering*, 94, 1033–1044.

Yao, F. and Müller, H.-G. (2010). Functional quadratic regression. *Biometrika*, 97, 49–64.

Yoon, S. K., Kim, S. H. and Lee, G. M. (2003). Effect of low culture temperature on specific productivity and transcription level of anti-4–1BB antibody in recombinant Chinese hamster ovary cells. *Biotechnology Progress*, 19, 1383–1386.