



NCS

Non-Clinical
Statistics
Conference

Padua, IT / 23 – 25 September 2026



Efficient experimental designs for pharmaceutical production

Kalliopi Mylona

Department of Mathematics, King's College London

01

**Design of
Experiments**

02

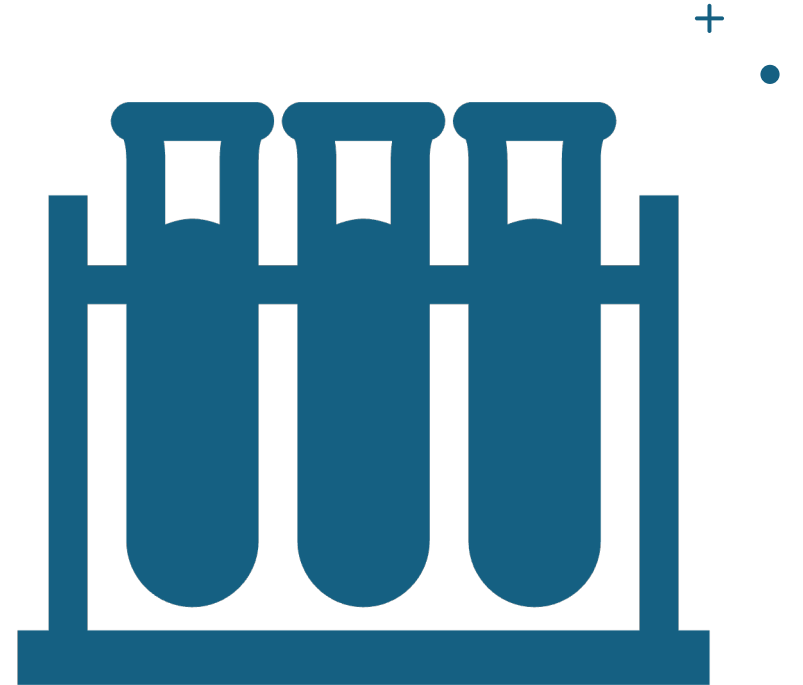
Case Studies

03

Discussion

Design of Experiments (DoE)

- **A structured plan** for conducting a series of experiments on a system, process, or product.
- **A systematic alternative** to trial-and-error testing.
- **A framework for data collection** to model, understand, and optimise performance.



Split-plot experiments

hard-to-change factors

- Levels of hard-to-change factors are changed infrequently.
- This creates groups of observations collected under the same hard-to-change factor settings.
- These groups are called *whole plots*.

two-stage experiments

- Many industrial experiments span several stages of production.
- Certain factors can only be applied at an early stage.
- Additional factors are applied in subsequent stages.

Response Surface Model

$$Y = X\beta + Z\gamma + \varepsilon$$



**whole-plot
effect**

i.i.d. $N(0, \sigma_\gamma^2)$



random error

i.i.d. $N(0, \sigma_\varepsilon^2)$

Variance-Covariance matrix

$$Y = X\beta + Z\gamma + \varepsilon$$

$$V = \text{var}(Y)$$

$$= \sigma_{\varepsilon}^2 \mathbf{I}_n + \sigma_{\gamma}^2 \mathbf{Z}\mathbf{Z}'$$

$$= \sigma_{\varepsilon}^2 (\mathbf{I}_n + \eta \mathbf{Z}\mathbf{Z}') \text{ where } \eta = \frac{\sigma_{\gamma}^2}{\sigma_{\varepsilon}^2}$$

Estimation

- generalized least squares regression

$$\hat{\beta} = \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X} \right)^{-1} \mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{Y}$$

- covariance matrix

$$\text{var}(\hat{\beta}) = \left(\mathbf{X}' \hat{\mathbf{V}}^{-1} \mathbf{X} \right)^{-1}$$

- $\hat{\mathbf{V}}$ is the estimated covariance structure of the observations
- $\hat{\mathbf{V}}$ can be estimated using restricted maximum likelihood estimation (REML)

Case Study I

The presented work has been conducted as a part of the
"Multi-Objective Optimal Design of Experiments" EPSRC grant
(EP/T021624/1)

Olszewska, A.; Egorova, O.; Gaggia, G.; Mylona, K.; Pitchford, S.; Rickard, J.; Forbes, B. A Design of Experiments Approach to Identify Critical Processing Parameters for Manufacture of an Autologous Platelet Gel for Diabetic Foot Ulcer. *Pharmaceutics* **2025**, *17*, 1482

- **Target:** Determine which processing parameters have the greatest impact on the effectiveness of a burn wound gel.
- **Measured outcomes:** Total Exudate Release (Total ER) and Gel Time (GT).

Table 1. Process parameters as experimental factors

Factor	Process parameter	Range of values	Levels	Scaled levels
x_{MT}	Temperature, °C	[15, 37]	{15, 24, 37}	{−1, −0.18, +1}
x_T	Mixing Time, rotations	[0, 30]	{0, 15, 30}	{−1, 0, +1}
x_W	WBC Content, %	[2, 8]	{2, 5, 8}	{−1, 0, +1}
x_t	Time to Thrombin use, min	[5, 20]	{5, 12.5, 20}	{−1, 0, +1}
x_{FT}	Filtered Thrombin	{No, Yes}	{0, 1}	{−1, +1}
x_M	Mixing	{No, Yes}	{0, 1}	{−1, +1}

Case Study I

- Human whole blood was sourced from a certain company and delivered overnight on ice.
- The experiment was completed over 8 days, with one donor blood bag used per day.
- Each day, four gel-manufacturing runs were performed under different process settings.
- Because manufacturing temperature was difficult to change between runs, it was treated as a whole-plot factor.

Case Study I

32 runs

8 whole plots of size 4

1 whole-plot factor

5 subplot factors

full quadratic model

Accommodate for the possibility of degree 3 polynomial terms

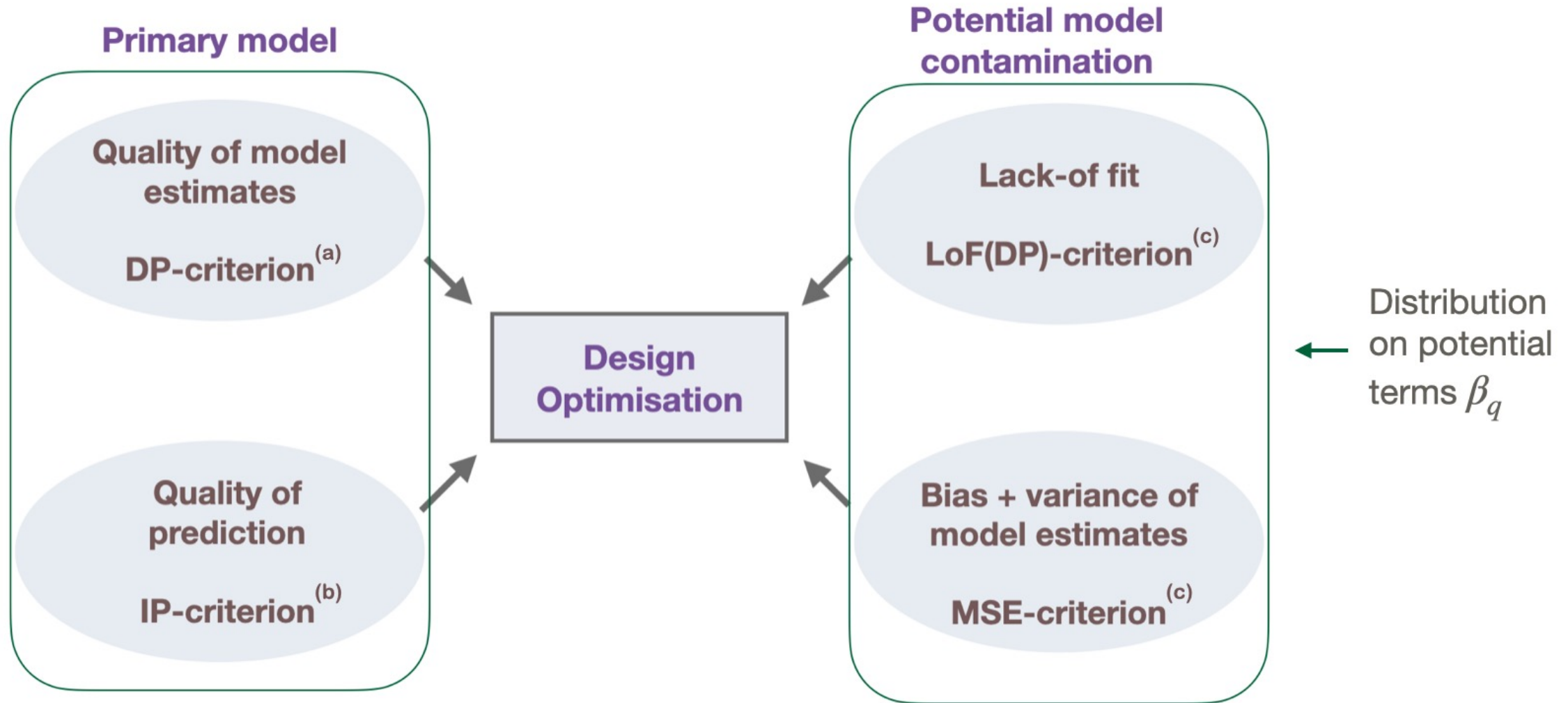
Statistical model

Response Surface Methodology. Restricted randomisation.

- Fitted polynomial model: $Y = \sum_{i=1}^s Z_i \epsilon_i + X_p \beta_p$, p primary terms.
- Potentially better model: $Y = \sum_{i=1}^s Z_i \epsilon_i + X_p \beta_p + X_q \beta_q$, with "extra" q potential terms: model contamination.

Inference based on the primary model with mitigating the potential effects of the contamination.

Optimality Criteria



- (a) Gilmour & Trinca, 2012
- (b) de Oliveira et al., 2022
- (c) Egorova & Gilmour, 2022

Design search

$$\text{Eff}_{DP}^{\kappa_{DP}} \times \text{Eff}_{IP}^{\kappa_{IP}} \times \text{Eff}_{LoF}^{\kappa_{LoF}} \times \text{Eff}_{MSE}^{\kappa_{mse}} \longrightarrow \max$$

Constructing a single compound criterion with weights

As an alternative to Pareto search; optimality criterion, various sets of weights

Algorithm: stratum-by-stratum design construction (Trinca & Gilmour, 2001)

Some optimal design performancies



Weights				Efficiencies, %			
DP	IP	LoF	MSE	DP	IP	LoF	MSE
1	0	0	0	100.00	96.86	100.00	97.63
0	1	0	0	86.99	100.00	100.00	85.33

**DP-optimal
design used
to run the
experiment**

Bag	Run	Temperature	Mixing time	WBC content	Time	Filtered Thrombin	Mixing
1	1	15	15	8	20	No	No
1	2	15	30	8	5	Yes	Yes
1	3	15	0	2	5	No	No
1	4	15	0	5	20	Yes	Yes
2	5	37	30	2	20	Yes	No
2	6	37	0	8	5	Yes	Yes
2	7	37	30	2	5	No	No
2	8	37	0	2	20	No	Yes
3	9	37	30	8	20	No	Yes
3	10	37	0	8	20	Yes	No
3	11	37	30	8	5	Yes	No
3	12	37	30	2	5	Yes	Yes
4	13	15	30	2	20	No	Yes
4	14	15	30	2	5	Yes	No
4	15	15	15	8	12.5	Yes	Yes
4	16	15	30	8	5	No	No
5	17	15	30	5	12.5	No	No
5	18	15	0	2	5	Yes	Yes
5	19	15	0	8	5	Yes	No
5	20	15	0	2	20	Yes	No
6	21	24	30	8	20	Yes	No
6	22	24	15	2	20	Yes	Yes
6	23	24	0	5	20	No	No
6	24	24	0	8	20	No	Yes
7	25	24	0	2	12.5	Yes	No
7	26	24	30	8	20	Yes	Yes
7	27	24	30	2	20	No	No
7	28	24	15	5	5	No	Yes
8	29	37	30	8	5	Yes	No
8	30	37	0	8	5	No	No
8	31	37	0	2	5	Yes	No
8	32	37	0	8	20	Yes	No

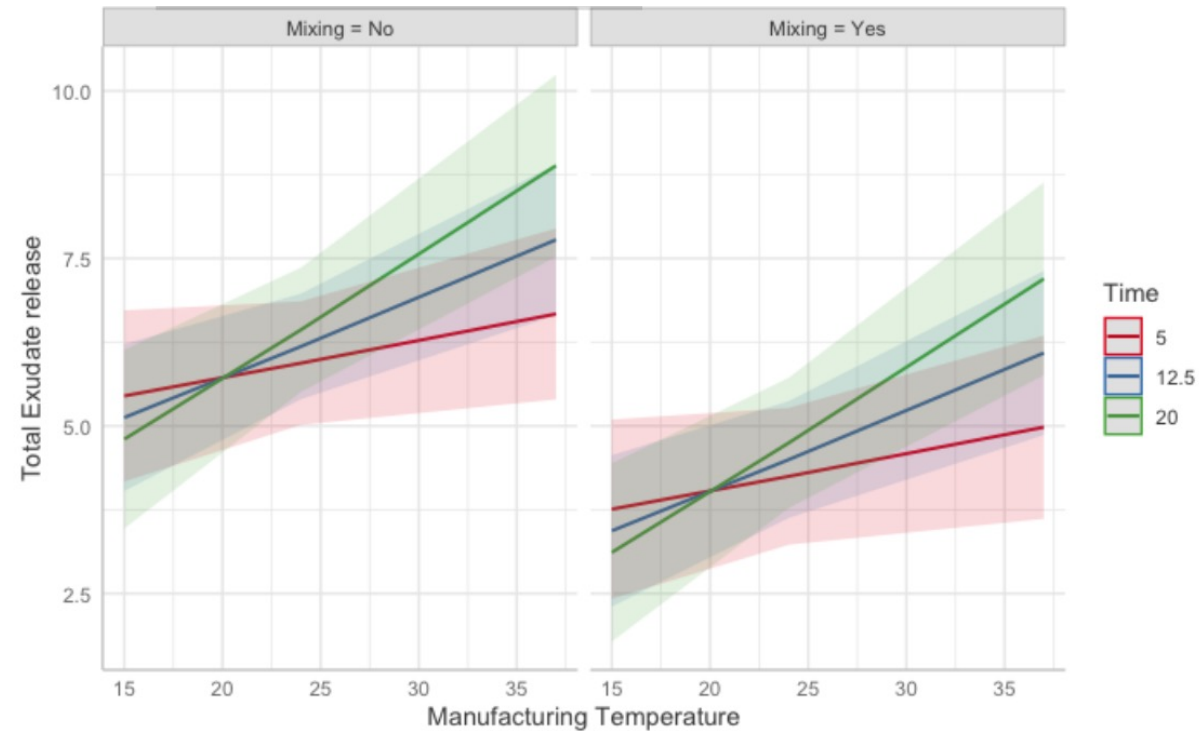
Statistical inference

- A linear mixed-effects model with a full quadratic response surface was fitted.
- REML was used to estimate variance components, and GLS was used to estimate fixed effects.
- All analyses were conducted in R.

Linear mixed model for Total Exudate Release

Parameter	Value	Standard Error	Degrees of freedom	t-value	p-value
Intercept, β_0	6.4499	0.3957	8.8723	16.236	0.0000
Temperature, β_{MT}	1.3260	0.4057	5.8767	3.268	0.0176
Time to Thrombin use, β_t	0.3917	0.2470	24.2309	1.586	0.1257
Mixing, β_M	-1.6888	0.4611	22.9054	-3.662	0.0013
Temperature $\times$ Time to Thrombin use, β_{MT*t}	0.7144	0.2650	21.8555	2.696	0.0133

Predicted Total Exudate Release

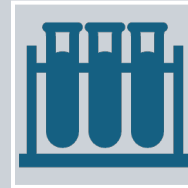


Case Study II

Otava, M. & Mylona, K. (2025). Split-plot experiments with replicated runs in pharmaceutical synthesis, *Quality and Reliability Engineering International*, 41, 641-651.



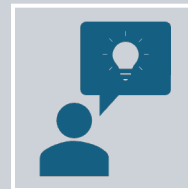
The objective was to optimise the synthesis of an active pharmaceutical ingredient



A prior screening study on a similar material had already been completed.



Therefore, the experiment was designed for process optimisation rather than factor screening.



Factor levels were expressed on a coded scale from -1 to $+1$.

Case Study II

The experiment was highly time-intensive, requiring a full day for each run, with a maximum of three runs completed per week.

Weekly preparation and operational constraints made it impractical to change the level of Factor A within the same week.

Seven full weeks were available, allowing a total of 21 experimental runs.

An additional week included a public holiday, reducing capacity to two runs.

The process was modelled using a Response Surface Model (RSM).

Bayesian composite optimality criterion

Mylona, K., Gilmour, S. G., & Goos, P. (2019). Optimal blocked and split-plot designs ensuring precise pure-error estimation of the variance components. *Technometrics*, 61(4), 481–492.

- Assume prior distribution for η
- Find design that maximises

$$D_B = \int_0^{\infty} \log (D(\eta)) \pi(\eta) d\eta$$

with $\pi(\eta)$ the prior distribution

Composite optimality criterion

- Find design that maximises

$$D = |\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|^{w/p} \times \left| \begin{array}{cc} \text{tr}((\mathbf{P}_t\mathbf{Z}\mathbf{Z}')^2) & \text{tr}(\mathbf{P}_t^2\mathbf{Z}\mathbf{Z}') \\ \text{tr}(\mathbf{P}_t^2\mathbf{Z}\mathbf{Z}') & \text{tr}(\mathbf{P}_t^2) \end{array} \right|^{(1-w)/2}$$

with

$$\mathbf{P}_t = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}_t(\mathbf{X}_t'\mathbf{V}^{-1}\mathbf{X}_t)^{-1}\mathbf{X}_t'\mathbf{V}^{-1}$$

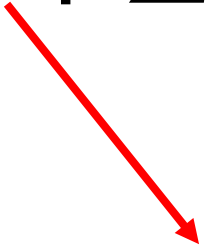
- depends on η^* too, so D should read $D(\eta^*)$ where $\eta^* = \frac{\sigma_{pure-\gamma}^2}{\sigma_{pure-\varepsilon}^2}$

Treatment Model

$$Y = X_t \tau + Z \gamma + \varepsilon$$



full treatment
design matrix



vector of
treatment means

$$X\beta = X_t \tau$$



Assumption

RSM=TM

The equal-interest criterion

$$D = \left| \mathbf{X}' \mathbf{V}^{-1} \mathbf{X} \right|^{1/p+2} \times \left| \begin{array}{cc} \text{tr} \left((\mathbf{P}_t \mathbf{Z} \mathbf{Z}')^2 \right) & \text{tr} \left(\mathbf{P}_t^2 \mathbf{Z} \mathbf{Z}' \right) \\ \text{tr} \left(\mathbf{P}_t^2 \mathbf{Z} \mathbf{Z}' \right) & \text{tr} \left(\mathbf{P}_t^2 \right) \end{array} \right|^{1/p+2}$$

- The special case of the previous one in the event w is set to $p/(p + 2)$.
- For the example involving a 4-factor response surface model, $p = 15$, in which case the alternative composite criterion is equivalent to the previous one when $w = 15/17 \approx 0.88$.
- We found that the $w = 0.5$ often lead to designs involving the maximum number of replicates, $n-p$, and the minimum number of distinct treatments, p .

Lognormal prior for η

- A log-normal prior can be conveniently specified using its geometric mean (u) and geometric standard deviation (v). For example, 68.3% and 99.7% of the probability mass lie within $[u/v, uv]$ and $[u/v^3, uv^3]$, , respectively.
- Three alternative priors were selected in consultation with the experimenters.
- All three priors led to the same optimal design.
- Gauss-Hermite quadrature to numerically evaluate the Bayesian criterion.

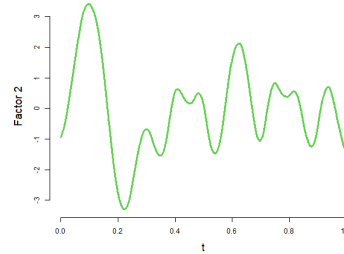
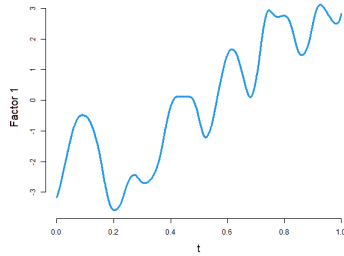
Design Stage

- Our criterion was used to construct the optimal design for seven weeks, with three experimental runs available per week.
- The week, in which only two runs were possible, was reserved for fine-tuning
- The design was generated using the coordinate-exchange optimisation algorithm employed in the original paper.

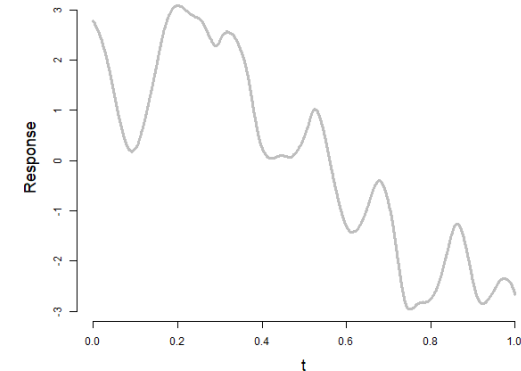
• Final Design

Run	Week	Factor A	Factor B	Factor C	Factor D	Comment
1	1	1	-0.05	-1	1	
2	1	1	-1	1	-1	
3	1	1	1	-0.05	-0.05	
4	2	-0.37	-0.05	-0.05	-0.05	target
5	2	-0.37	-0.05	-0.05	-0.05	target
6	3	1	1	-1	-1	
7	3	1	-1	-0.05	1	
8	3	1	-0.05	1	-0.05	
9	4	-0.37	-1	-1	-1	
10	4	-0.37	-1	1	1	
11	4	-0.37	1	-1	1	
12	5	-1	1	-1	-1	
13	5	-1	-1	1	-1	
14	5	-1	1	1	1	
15	6	1	-1	-1	-0.05	
16	6	1	1	1	1	
17	6	1	-0.05	-0.05	-1	
18	7	-1	1	1	1	replicated
19	7	-1	-1	1	-1	replicated
20	7	-1	-1	-1	1	
21	8	-0.37	-1	1	1	replicated
22	8	-0.37	1	1	-1	
23	8	-0.37	-1	-1	-1	replicated

Dynamic factors



Statistical
model



Continuous variables

Categorical variables

May, C., Ladas, T., Pigoli, D. and Mylona, K. Optimal design of experiments for functional linear models with dynamic factors. *TEST* (2026), 35, 469–497.

Case study III (motivated by a pharmaceutical manufacturing process described in Van Snick et al. (2019))

- Material is fed into a continuous process via a gravimetric feeder, while outlet concentration is measured in-line.
- Each experimental run can be viewed as a dynamic input-output system.
- The input is the feeder dosing profile, typically a stepwise function over time, as the feeder may be turned on or off, or operated at a few different levels in successive time intervals.
- The output is the outlet concentration profile, which responds with a delay and is smoother due to transport and mixing effects.

Estimation of the parameters in a fully function linear model: *matrix form*

$$\mathbf{y}(t) = \int_a^b \boldsymbol{\mathcal{X}}(s) \boldsymbol{\beta}(s, t) ds + \boldsymbol{\varepsilon}(t)$$

where

$$\mathbf{y}(t) = (y_1(t), \dots, y_N(t))^T$$

$$\boldsymbol{\varepsilon}(t) = (\varepsilon_1(t), \dots, \varepsilon_N(t))^T$$

$$\boldsymbol{\beta}(s, t) = (\mu(t), \beta_1(s, t), \dots, \beta_p(s, t))^T$$

$$\boldsymbol{\mathcal{X}}(s) = (\mathbf{1}_N, X(s))$$

$$X(s) = \begin{pmatrix} x_{11}(s) & \cdots & x_{1p}(s) \\ \vdots & \ddots & \vdots \\ x_{N1}(s) & \cdots & x_{Np}(s) \end{pmatrix}$$

$$\mathbf{1}_N = (1, \dots, 1)^T$$

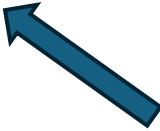
Estimation of the parameters in a fully function linear model: *basis expansion*

Let us consider some sets of basis functions:

$$\{\theta_l(t), l \geq 1\} , \{\eta_k^1(s), k \geq 1\}, ..., \{\eta_k^p(s), k \geq 1\}$$

and let
$$\mu(t) = \sum_l b_l^0 \theta_l(t) = \mathbf{b}^{0T} \boldsymbol{\theta}(t)$$

and
$$\beta_i(s, t) = \sum_{k,l} b_{k,l}^i \eta_k^i(s) \theta_l(t) = \boldsymbol{\eta}^i(s)^T \mathbf{B}^i \boldsymbol{\theta}(t)$$



Separable representation

Estimation of the parameters in a fully function linear model: *basis expansion*

$$\mathbf{y}(t) = \int_a^b \boldsymbol{\mathcal{X}}(s) \boldsymbol{\beta}(s, t) ds + \boldsymbol{\varepsilon}(t)$$



$$\mathbf{H}(s)^T = \text{diag}(1, \boldsymbol{\eta}^1(s)^T, \dots, \boldsymbol{\eta}^p(s)^T)$$

$$\mathbf{y}(t) = \int \boldsymbol{\mathcal{X}}(s) \mathbf{H}(s)^T \mathbf{B} \boldsymbol{\theta}(t) ds + \boldsymbol{\varepsilon}(t)$$

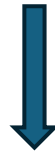


$$\mathbf{Z} = \int \boldsymbol{\mathcal{X}}(s) \mathbf{H}(s)^T ds$$

$$\mathbf{y}(t) = \mathbf{Z} \mathbf{B} \boldsymbol{\theta}(t) + \boldsymbol{\varepsilon}(t)$$

Estimation of the parameters in a fully function linear model: *reduced rank estimator*

$$\hat{\mathbf{B}} = \arg \min_{\mathbf{B}} \int \|\mathbf{y}(t) - \mathbf{Z} \mathbf{B} \boldsymbol{\theta}(t)\|^2 dt$$



Proposition (Ramsay and Silverman, 2005): If $\mathbf{Z}$ is full rank and $\{\boldsymbol{\theta}_l(t)\}$ is an orthonormal basis, then

$$\hat{\mathbf{B}} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \int \mathbf{y}(t) \boldsymbol{\theta}(t)^T dt$$

Proposition (May et al., 2026)

If Z is full rank and $\{\theta_l(t)\}$ is an orthonormal basis, then

$$\text{Var}(\text{vec } \hat{B}) = \Sigma \otimes (Z^T Z)^{-1}$$

Vectorisation of the
matrix
of the coefficients'
estimators, obtained by
stacking the columns

Matrix of the covariances of the projections
of the error on the basis functions, i.e.
 $\Sigma_{lq} = \text{cov}(\int \varepsilon(t)\theta_l(t)dt, \int \varepsilon(t)\theta_q(t)dt)$

Proposition (May et al., 2026)

If Z is full rank and $\{\theta_l(t)\}$ is an orthonormal basis, then

[A-optimality] The A-optimal design $X_A^*(s)$ satisfies

$$X_A^*(s) = \arg \min_{X(s)} \text{tr}(Z^T Z)^{-1}$$

[D-optimality] The D-optimal design $X_A^*(s)$ satisfies

$$X_D^*(s) = \arg \min_{X(s)} \frac{1}{\det Z^T Z}$$

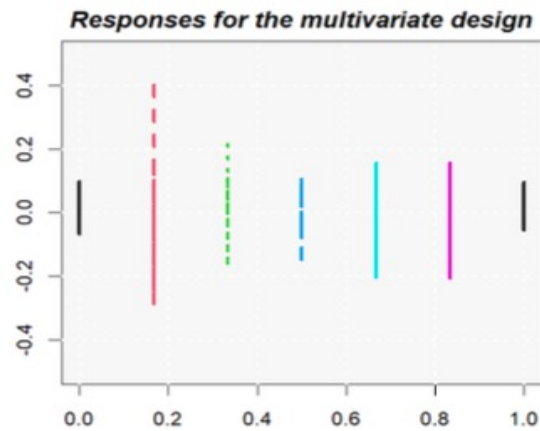
Simulations and designs under comparison.

- A bivariate coefficient surface was specified to generate simulated response data.
- The true coefficient is represented on a B-spline basis with 4×4 cubic splines and random coefficients.
- The error is a Gaussian process with radial-basis-function kernel and small variance, represented on a Fourier basis with 81 terms.

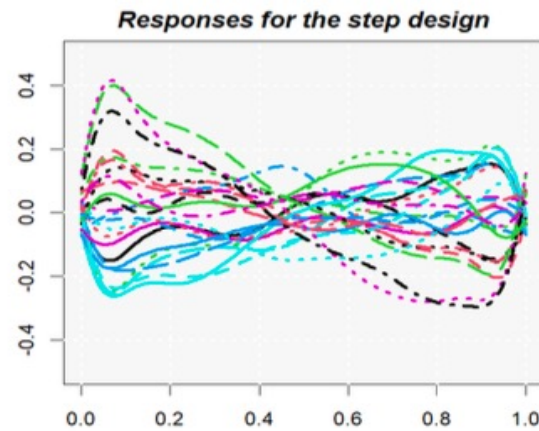
We compare three different designs, each one with $n = 24$ runs :

1. **Multivariate baseline** (naive approach)
2. **Step-step** (basis aligned with process operation)
3. **Spline-spline** (smooth dosing trajectories)

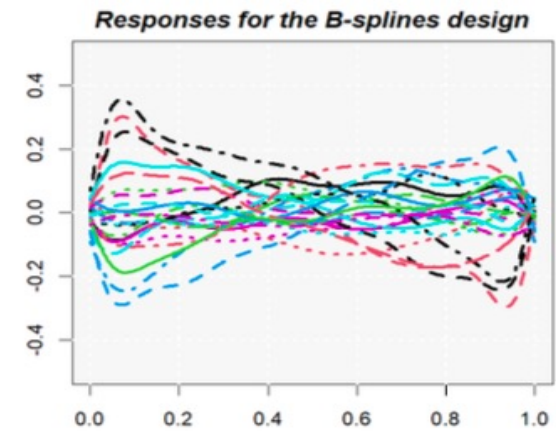
Simulated response curves



(a) Multivariate baseline

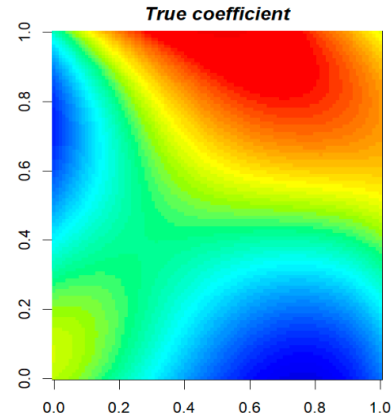


(b) Step-step design

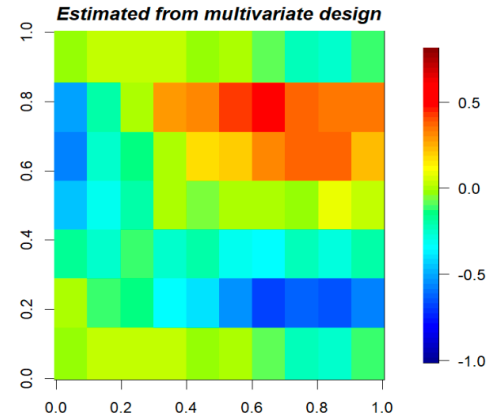


(c) Spline-spline design

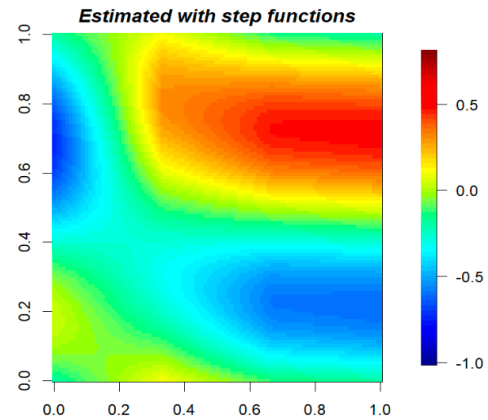
True coefficient and estimated coefficient surfaces under the three designs.



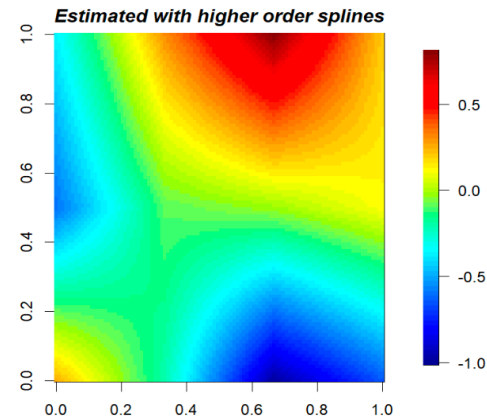
(a) True coefficient



(b) Estimated from multivariate design

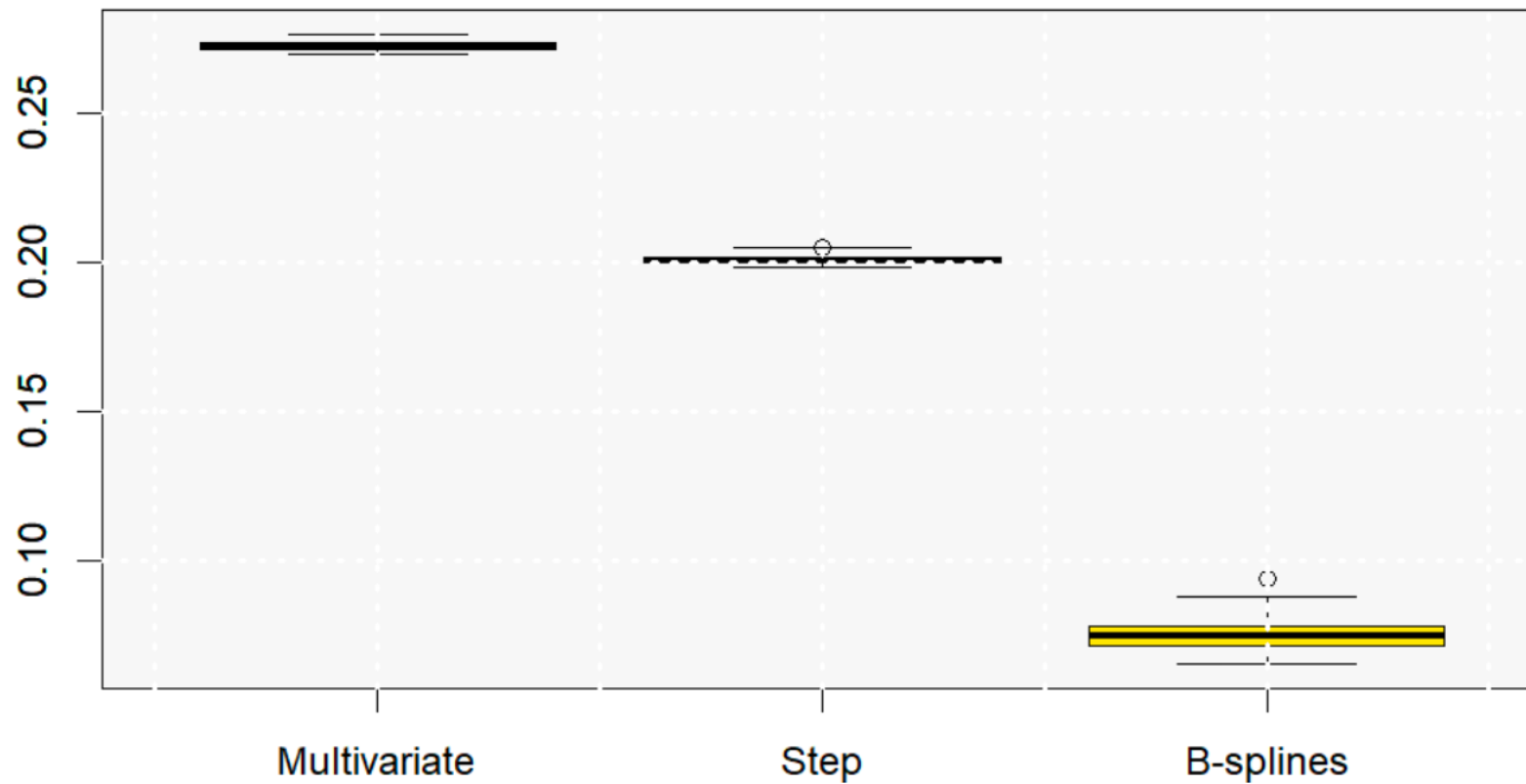


(c) Estimated from step-step design

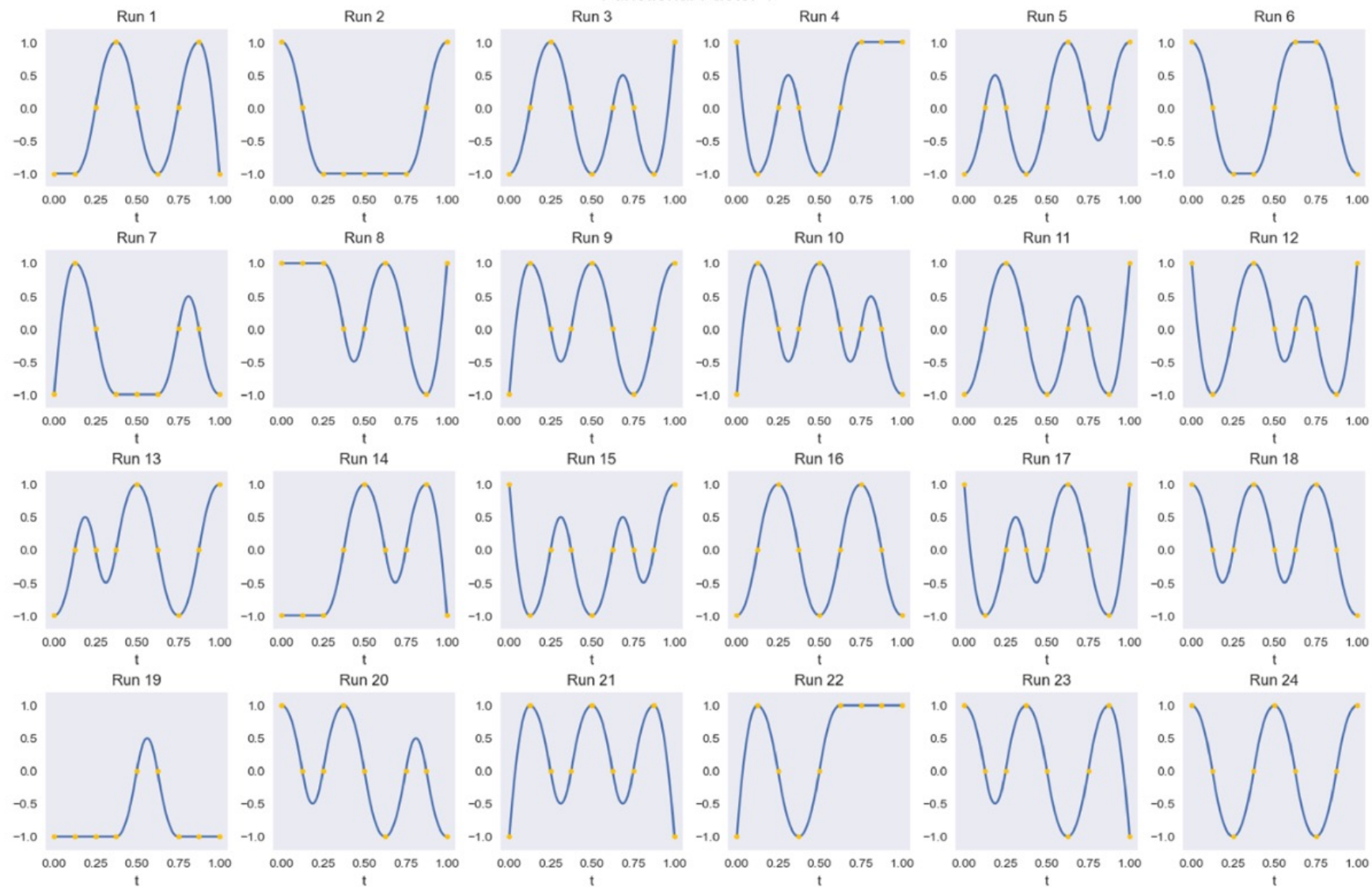


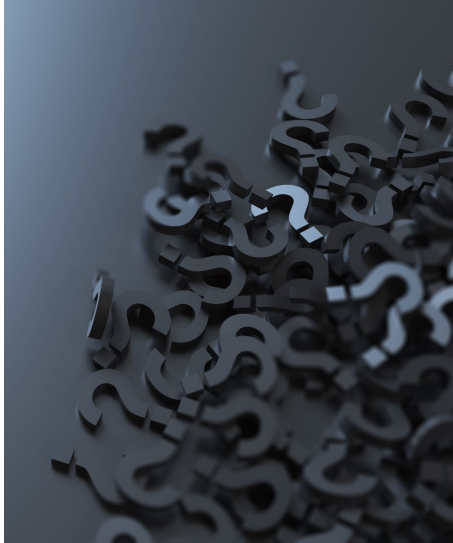
(d) Estimated from spline-spline design

RMSE over 100 simulated experiments. Left to right: multivariate baseline, step-step, spline-spline.



Functional Factor 1





Discussion

- R package: multiDoE: Multi-Criteria Design of Experiments for Optimal Design (<https://cran.r-project.org/web/packages/multiDoE/index.html>)
- Python: Neuro Bayes Design Optimiser

The presented work has been conducted as a part of the "Multi-Objective Optimal Design of Experiments" EPSRC grant (EP/T021624/1)

