From rigid plans to adaptive learning: modernizing Design Of Experiments in chemical process optimization

Authors

Gréta TÁLOSI[1,2], Alex KUMMER[1], Zoltán ELLER[2], Bálint Levente TARCSAY[1]

Affiliations

[1]University of Pannonia, Department of Process Engineering, H-8200 Veszprém, Egyetem u. 10., www.uni-pannon.hu
[2]MOL Hungarian Oil ans Gas Plc., H-1117 Budapest, Dombóvári út 28., www.mol.hu

Research Lab

Abstract

Industrial process development is increasingly constrained by cost and time pressures, while demanding accurate and robust optimization. Traditional Design of Experiments (DOE) offers structured, interpretable results but is limited in highly nonlinear or complex systems. This study benchmarks classical DOE against model-based strategies using Bayesian Optimization in a simulated first-order exothermic batch reaction. Reactor temperature, wall temperature, and initial concentration defined the design space, and Gaussian Process surrogate models enabled consistent evaluation across all designs. Performance was assessed via predictive accuracy, global uncertainty, optimization reliability and experimental cost. Classical DOE methods showed low uncertainty only in the vicinity of the sampled points, while failing to reliably identify the global optimum even with more experimental runs. In contrast, model-based DOE outperformed classical designs, yielding 27% lower global predictive uncertainty and achieving a 50% reduction in the distance between the predicted and true optimum, demonstrating superior predictive accuracy and optimization reliability. These findings underscore the potential of adaptive, model-guided DOE to accelerate process development under stringent industrial constraints.

Keywords: Design of Experiments · Bayesian Optimization · Gaussian Process · surrogate modelling · chemical process optimization

1 Introduction

DOE has long provided systematic frameworks to quantify factor effects and interactions, yet many industrial settings still rely on one-factor-at-a-time testing, limiting efficiency and insight. Advances in computational power and statistical learning now enable model-driven approaches that adaptively guide experimentation, offering greater data efficiency and reliability. This study benchmarks classical DOE against model-based strategies in a simulated reaction system to highlight their relative strengths and limitations.

2 Objectives

- Compare classical vs. model-based DOE methods
- Benchmark performance: predictive accuracy, uncertainty, optimization reliability, experimental cost
- Demonstrate how adaptive learning reduces number of experiments while improving results

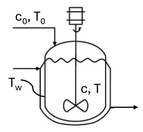
3 Methodology

- **Case study:** First-order exothermic batch reaction in a wall-cooled reactor (simulated)
- **Inputs:** c_0, T_0, T_w
- **Responses:** conversion, max temperature (soft constraint at 340 K)
- **Approaches tested:**
  - 7 classical designs (full/fractional factorial, CCD, BBD, Taguchi L9, L18) – all fitted with GP surrogate models
  - Model-based DOE: Bayesian Optimization with multiple acquisition functions and GP surrogates
- **Evaluation criteria:**
  - Number of runs (efficiency)
  - R^2 on validation grid (accuracy)
  - Mean GP uncertainty $\bar{\sigma}$ (confidence)
  - Euclidean distance to true optimum (optimization quality)

Reactor model:

$$\frac{dc}{d\tau} = -\exp\left(\gamma - \frac{\delta}{T}\right)\cdot c$$

$$\frac{dT}{d\tau} = \beta\cdot\exp\left(\gamma - \frac{\delta}{T}\right)\cdot c - \alpha\cdot(T - T_w), \qquad \tau \in [0,1]$$



γ, δ: dimensionless kinetic parameters [-]

$\beta = \frac{-\Delta H_r}{\rho c_p}$: parameter [m³K/mol] combining heat of reaction ($-\Delta H_r$ [kJ/kmol]), density (ρ [kg/m³]) and specific heat (c_p [kJ/m³K])

$\alpha = \frac{UA}{V\rho c_p}$: parameter characterizing heat transfer efficiency [1/h] via heat transfer coefficient (U [W/m²K]), heat transfer surface (A [m²]), reactor volume (V [m³]), and previously described density and specific heat factors

Objective function:

$$obj(c_0, T_0, T_w) = \begin{cases} -conv(c_0, T_0, T_w) & if\ T_{max} \le T_{\max\ limit} \\ -conv(c_0, T_0, T_w) + \omega_p (T_{max} - T_{\max\ limit})^2 & if\ T_{max} > T_{\max\ limit} \end{cases}$$

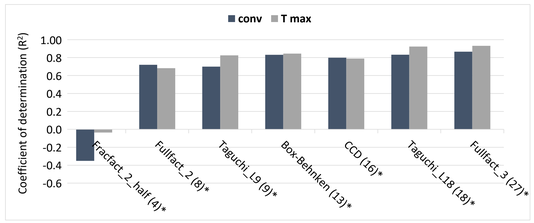
4 Results

Classical DOE methods demonstrated good predictive performance when sufficient design points were available, but improvements diminished beyond 16-run Central Composite Designs (**Fig. 1**).

Gaussian Process maps highlighted that predictive uncertainty was lowest near sampled points and highest in unexplored regions, as illustrated for a full factorial design (**Fig. 2**).

In contrast, Bayesian Optimization adaptively reduced uncertainty around promising regions within only 10 iterations, leading to efficient identification of the optimum (**Fig. 3**).

Comparative analysis of all methods showed that model-based DOE achieved lower predictive uncertainty and smaller distances to the true optimum under the same budget, confirming its superior efficiency (**Fig. 4**).



*notation: method (number of teaching samples)

Figure 1. Predictive accuracy (R^2) of classical DOE designs as a function of number of runs. Increasing sample size improves model fit, but gains diminish beyond 16-run CCD-level design.

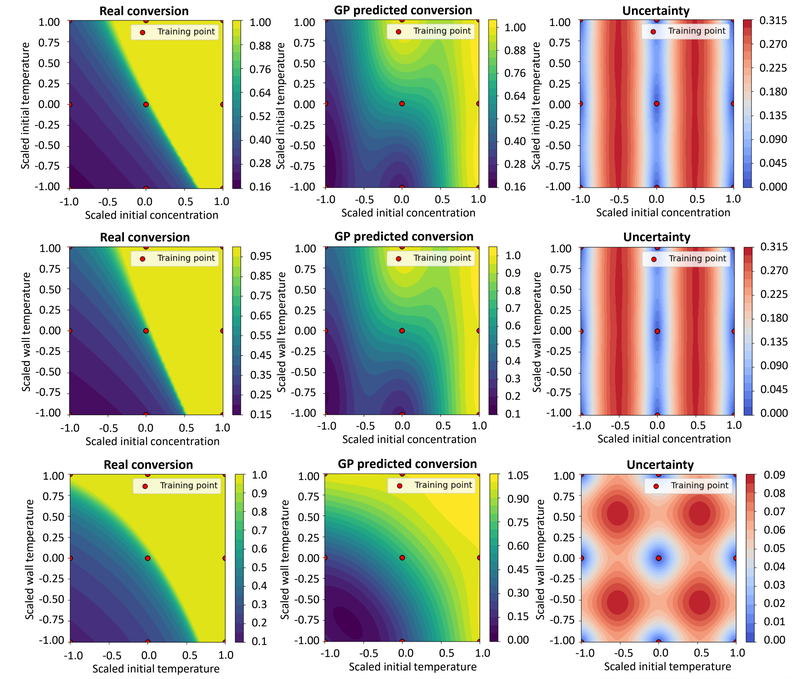
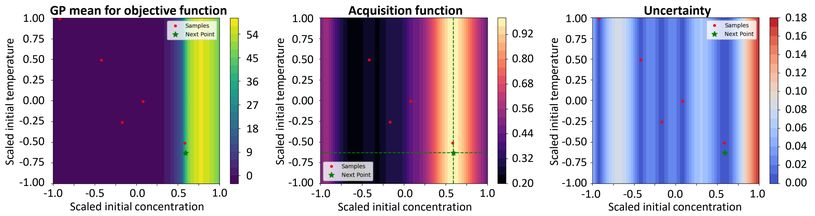


Figure 2. Gaussian Process surrogate uncertainty map (conversion, full factorial 3-level design). High confidence near sampled points, but elevated uncertainty in untested regions.

1. Iteration



10. Iteration

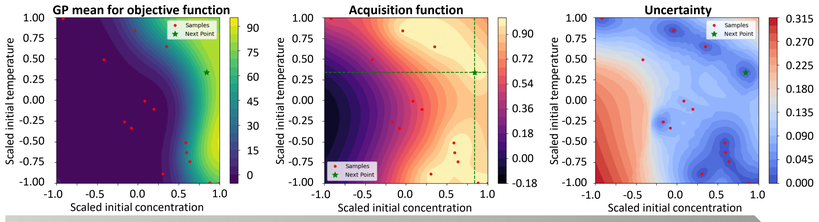


Figure 3. Evolution of uncertainty during Bayesian Optimization (LCB acquisition, 10 iterations). Adaptive sampling progressively reduces uncertainty in regions of interest.

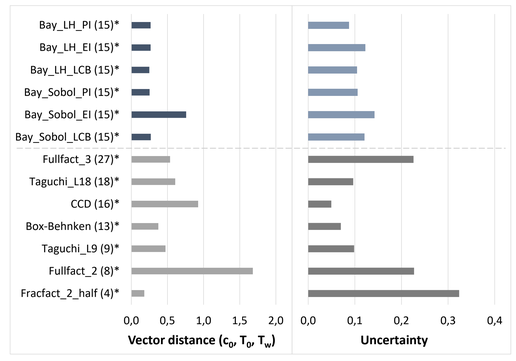


Figure 4. Comparative performance of DOE methods. Euclidean distance to true optimum and mean predictive uncertainty are shown, highlighting the superior efficiency of model-based DOE.

5 Conclusions

Classical DOE supports structured screening, but adaptive model-based strategies deliver superior efficiency and reliability. This benchmark highlights the promise of Bayesian Optimization for modern chemical process optimization.

Key takeaways:

- Classical DOE: simple but limited, misses the global optimum even with more runs
- Model-based DOE: accurate and efficient, 27% lower uncertainty, 50% closer to optimum
- Adaptive workflows: accelerate process development under tight industrial constraints

6 Future focus areas

- Experimental validation in reactor system under industrial conditions
- Hybrid mechanistic–data-driven modelling
- Development of scalable, transferable frameworks for industrial use

Together, these directions aim to accelerate the shift toward adaptive, data-driven experimentation in chemical process development.