

FORMULATION ADVISOR · AN INTRODUCTION

DECODEXIS

Optimum formulations with fewer experiments

Bayesian optimisation applied to product and process development

1 Example use case

2 Working with BO

3 Use cases

4 Next steps

Example use case: A body lotion with a new emulsifier

Ten historical trials on the original emulsifier, no data on the new grade, and a budget of five new bench trials. The task is to match the current product and reach specification within that budget.

OBJECTIVE

Reformulate around a bio-based emulsifier and match the current product.

TARGET SPECIFICATION

Viscosity	**4,000–6,000 cP**
Droplet size, d50	**below 10 µm**
Stability	**4 weeks at 40 °C**

CORE FORMULATION

Oil phase	8–14 wt%
Bio-based emulsifier, new grade	1.0–2.5 wt%
Thickener	0.3–0.8 wt%
Water, to balance 100 wt%	

PROCESS

Homogeniser speed	3,000–7,000 rpm
Mixing temperature	50–80 °C
Cooling rate	0.5–1.5 °C/min

DATA AVAILABLE

- Ten trials on the original emulsifier, with the same variables and outcomes
- No trials on the new grade
- The golden batch settings for the original formulation
- Specification data for both grades: HLB value, active content, viscosity

Example use case: Applying BO

Steps

01 Transfer the history

The ten trials on the original emulsifier are treated as a related task. The model learns the shape of the response surface from them and starts with no information specific to the new grade.

02 Establish the offset

Trial 1 is a direct substitution at the golden batch settings. Its measured result quantifies the difference between the two materials.

03 Refine under constraints

Each result updates the offset. The next trial balances exploration against exploitation, with viscosity and droplet size applied as hard constraints.

TRIAL	EMULSIFIER	THICKENER	HOMOGENISER	VISCOSITY	d50
1	1.6 wt%	0.50 wt%	4,500 rpm	2,400 cP	9.0 µm
2	1.6 wt%	0.68 wt%	5,600 rpm	3,600 cP	9.5 µm
3	1.7 wt%	0.74 wt%	6,200 rpm	**4,600 cP**	8.0 µm
4	1.65 wt%	0.70 wt%	6,000 rpm	4,900 cP	8.4 µm
5	1.65 wt%	0.70 wt%	6,000 rpm	4,700 cP	8.3 µm

Trial 3 reaches the specification; trials 4 and 5 refine and confirm it.

RESULT

Trial 1 held droplet size but fell below the viscosity target. The model attributed the gap to the emulsifier change, raised thickener and homogeniser speed for trial 2, and reached specification at trial 3. The campaign closed within the five-trial budget, with a documented reason for every setting and the ten historical trials retained as the basis of the model.

Working with BO: Black box and surrogate model

THE BLACK BOX

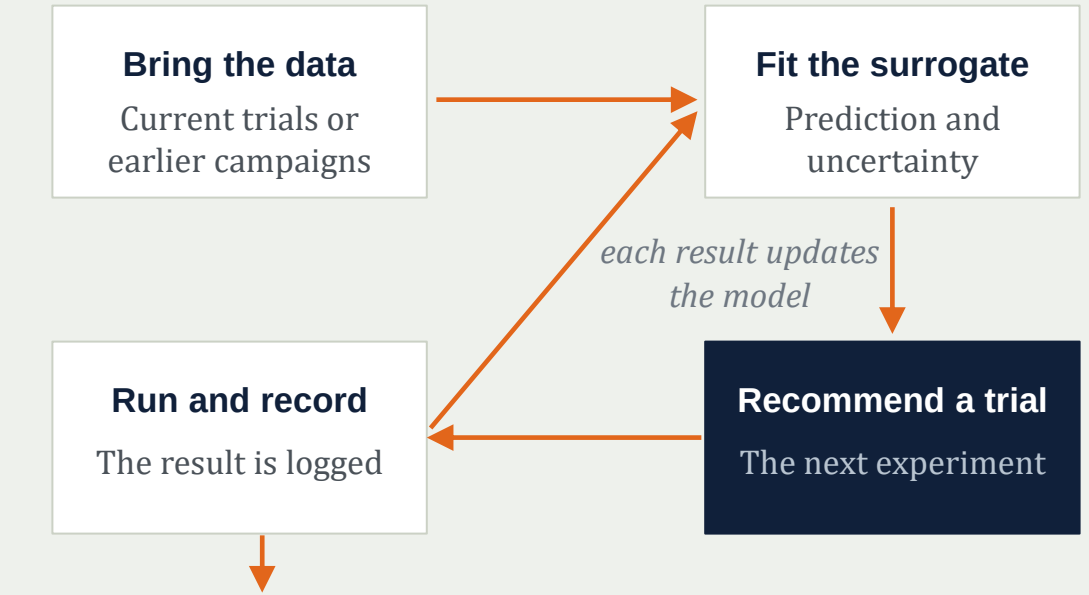
The formulation or process is treated as a black box: known inputs produce measured outcomes, and the relationship between them is unknown and expensive to evaluate.

THE SURROGATE

Bayesian optimisation fits a surrogate model — a fast statistical approximation of that black box — to the trials already completed. At any untested setting it returns a predicted outcome and an estimate of the uncertainty.

THE LOOP

A model is fitted to the experiments already available, and the next experiment is chosen from it. The loop repeats after every result.



Closed-loop optimisation. Each recorded result updates the model before the next recommendation.

1 Bring the data

Current trials, earlier campaigns, or both.

2 Warm up if needed

With no data, the tool proposes roughly two trials per variable, plus one.

3 Run the suggestion

The model identifies the next trial and gives the reason for it.

4 Record and repeat

The result updates the model before the next suggestion.

Working with BO: Exploration and exploitation

Every recommendation does two jobs at once: it improves the predicted outcome, and it reduces uncertainty where nothing has yet been measured.

EXPLOITATION

Improve the predicted outcome

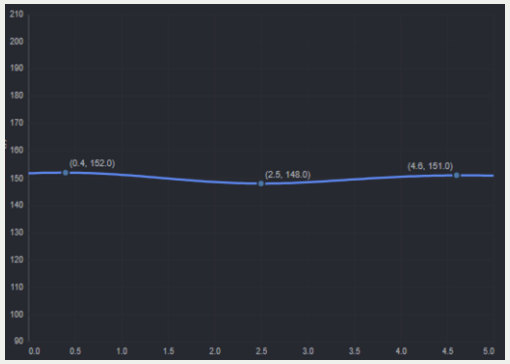
Run where the model already expects a good result.

EXPLORATION

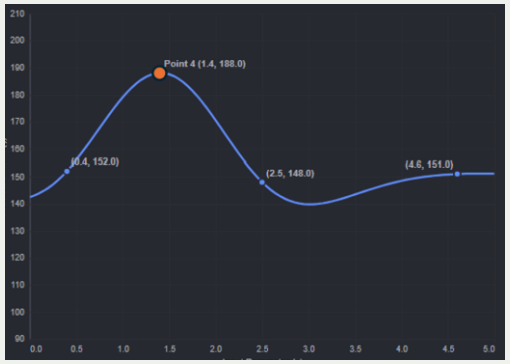
Reduce uncertainty

Run where nothing has been measured and the model is least certain.

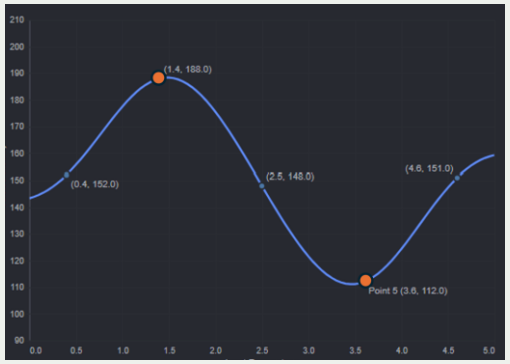
Excessive exploitation converges on a local optimum; excessive exploration amounts to unstructured search. The method balances the two, so each trial either improves the result or improves the model's knowledge of the system.



Three pins



Four pins



Five pins

The surrogate at three, four and five measured points. Each trial pins the sheet at the measured value; between pins, uncertainty remains, and the next recommendation is chosen to reduce it.

This is the elastic sheet: pinned where measured, free where not. For more elaborate explanation check article [BO for Product and Process Developers](#).

Working with BO: How it differs from other methods

Bayesian optimisation is built for an unknown, non-linear system with a limited experimentation budget.

	DESIGN OF EXPERIMENTS	LINEAR PROGRAMMING	PREDICTIVE ANALYTICS	BAYESIAN OPTIMISATION
What it requires	A designed matrix of trials	A known linear objective and known constraints	A large, standardised historical database	A small project table and a defined search space
Learning from results	After the full batch	None; coefficients are set in advance	After model training	After every trial
Uncertainty	Estimated from replication	Not represented	Model-dependent	Explicit at every untested setting
Output	Effects across the design space	An optimal allocation under known constraints	Predictions at scale	The next experiment to conduct

The distinction from design of experiments is examined in the article Too Many Experiments, Too Little Time.

Further use cases

Because the method learns from few observations, it suits any setting where experiments are expensive and time is short.

1

Screening candidate ingredients

Candidates are encoded by two or three specification descriptors, so the model can rank untested options after a few physical trials.

2

A formulation from scratch

Where no data exist, the first trials are spread across the search space; a viable formulation follows within about ten trials.

3

Process scale-up and transfer

Scale-invariant parameters such as tip speed replace equipment settings; two or three pilot trials quantify the offset.

4

Troubleshooting a failure

Failed batches are recorded as pass or fail, so the model maps the feasible boundary and finds the likely cause within three or four trials.

5

Reducing raw-material cost

Cost is applied as a hard constraint while quality remains the objective; varying the ceiling produces a cost-versus-performance frontier.

6

Improving robustness

Replicates with process drift, or a worst-case objective, steer the search to a broad operating plateau rather than a narrow peak.

Campaigns are not confined to the laboratory. Pilot runs, scale-up and production all supply usable data: routine records serve as history, and each deliberate change counts as the next observation.

Next steps

The workbench is free to use, and a first recommendation follows within minutes.

→ **Open the workbench**

Begin with a pre-configured example, or with a table from a current project. No installation and no payment.

→ **Arrange a walkthrough**

Twenty minutes on one of your own tables: what to vary, what to measure, what to hold as hard limits.

→ **Request the PDF guide**

Four foundation modules and seven use-case playbooks, with worked examples, available on request.

PROJECT DATA IS NOT STORED

The workbench keeps no product database and retains no project data after processing. The table is used to fit the model and to produce the next suggestion; any working copy, including a saved draft, remains in the browser on the scientist's own computer and can be deleted at any time.

Vidya Ranade
Decodexis Consulting Services · Rotterdam, The Netherlands