DATAHOW

CASE STUDY

Understanding and Predicting Product Quality (CQAs) in Process Development:

AI (hybrid) Models vs. Industry State-of-the-Art

Bristol Myers Squibb

Process Development | **mABS** | **Product Quality**

Controlling Product Quality

CQAs (Critical Quality Attributes) are the key physical, chemical, and biological product characteristics that must be within an appropriate limit, range, or distribution to ensure product safety, efficacy, consistency, and overall quality.

In development, process scientists are challenged to understand the relationship and impact of multiple process parameters on CQAs through the analysis of experimental data using advanced process data analytics.

The Challenge

48 (5 liter scale) experiments were designed and conducted by BMS to evaluate the impact of 12 process parameters on 18 different product CQAs.

DataHow was challenged to evaluate the ability of its AI enabled, hybrid process models to accurately predict CQAs compared to industry state-of the-art "black box" models.

Approach & Technology Deployed

BMS designed and executed 48 experiments using a classical fractional factorial design methodology.

With this data, a multiple linear regression model (MLR) was applied to understand the interrelationships between the process parameters and CQAs. These models are broadly considered state-of-the-art by the industry and are commonly referred to as "black-box" models.

Using the same data, DataHow deployed its AI-powered hybrid models.

Objectives

1. **Evaluate ability of hybrid models to predict CQAs**
   How do Hybrid models compare to industry black box models in understanding and predicting the relationship between process parameters and CQAs.

2. **Assess how much experimental data is needed to accurately predict CQA's for each approach**
   Determine how many experiments are needed for each modeling approach to be able to reliably predict CQAs.

3. **Assess other benefits of Hybrid Models**
   Determine whether Hybrd models offer any additional benefits over "black-box" models.

What is a Hybrid Model?

AI traditionally requires a high volume of data to be reliable and representative - a limitation for process development where data is scarce and expensive to create.

DataHow's hybrid models solve this by fusing a hard-coded "mechanistic" backbone, which code known process dynamics, with machine learning models to learn what is unknown.
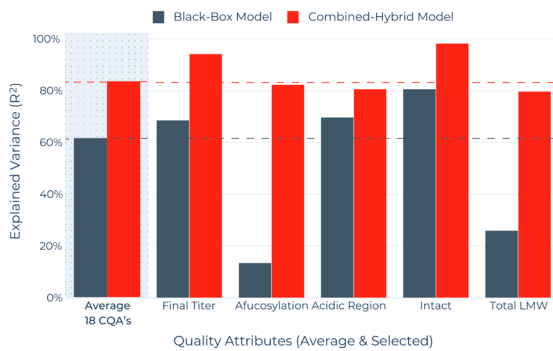
The result is AI applied to the unique requirements and data environment of bioprocessing.

Results

## Understanding & Predicting CQA's: Black Box vs. Hybrid

KEY INSIGHT 1:

Hybrid Models predict CQA more accurately than the industry state of the art "black-box" models

- On average, hybrid models improved CQA prediction by 22% (Explained variability of 83% vs 61% for black-box models)
- Even after 34 experiments, black-box models were unable to reliably predict 5 of the 18 CQAs (highlighted: Afucosylation / Total LMW)
- For some CQAs, the predictive ability of hybrid models was approaching 100% (highlighted: Titer / Intact)

of Experiments required to predict CQA's: Black Box vs. Hybrid
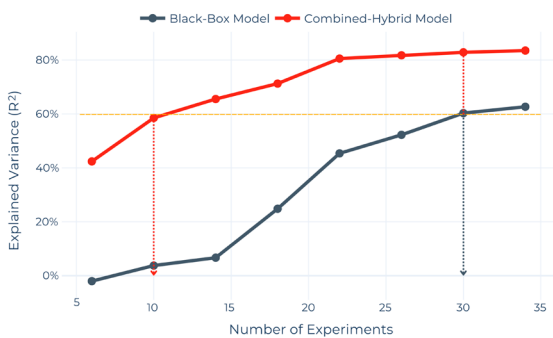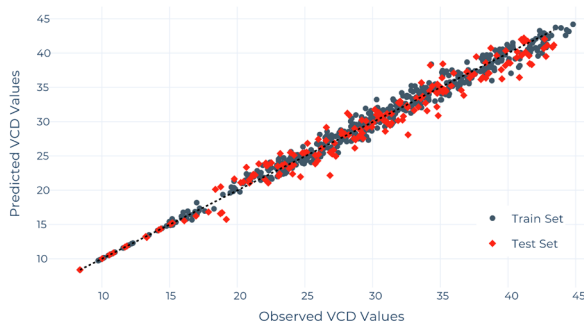

KEY INSIGHT 2:

Hybrid models can reliably predict CQAs with 3x fewer experiments

- Black box models needed 30 experiments before they could understand the CQA / process parameter interrelationships and reliably predict CQA values
- Hybrid models only required 10 experiments to reach the same level of predictive accuracy
- After 30 experiments, hybrid models offered 34% greater understanding, and thereby control, of CQAs.

Hybrid Models: Predicting Process Parameters (VCD)

KEY INSIGHT 3:

Hybrid Models return more process knowledge and insight to support greater process control and reduce risk

- Black-box models are able to predict CQAs but can not support in predicting process parameters. Additional analysis is required, with no industry standard approach.
- DataHow's hybrid models can accurately predict both CQAs and process parameters, offering greater process insight and control
- Highlighted example (VCD): Hybrid models were able to predict 70% of VCD variability

Impact of hybrid (AI) models
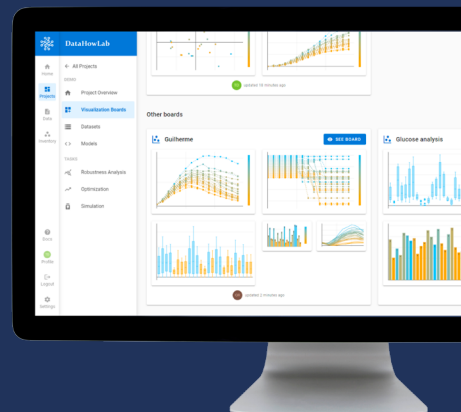
+22%
Improved CQA Prediction Accuracy

3x
Fewer Experiments needed to Predict CQA

More Process Knowledge, Insight & Control

Learn more about DataHow

DATAHOW