Application of beta regression response surface model for multidimensional design space construction
20/Oct/2022

Jessica Riley, *Shells*
Artwork from Reflections Art in Health

Chellafe Ensoy-Musoro

Manufacturing and Applied Statistical Sciences (MAS)

janssen | PHARMACEUTICAL COMPANIES OF Johnson&Johnson

Slide Overview

- Background
- Simulated Data
- Response Surface Model
- Design Space
  - Multivariate clasification tree
  - Parallel coordinates plot

Quality by Design

Design space construction:

- typically follows the **development of a response surface model** that relates different process parameters with various product quality attributes or process performance objectives (PPO) and
- **finding the set of process conditions** where acceptance criteria of the objectives are met with required level of assurance.

Quality by Design

- Normal response surface model
- Visualization of model results using contour plots and various 2D/3D plots.

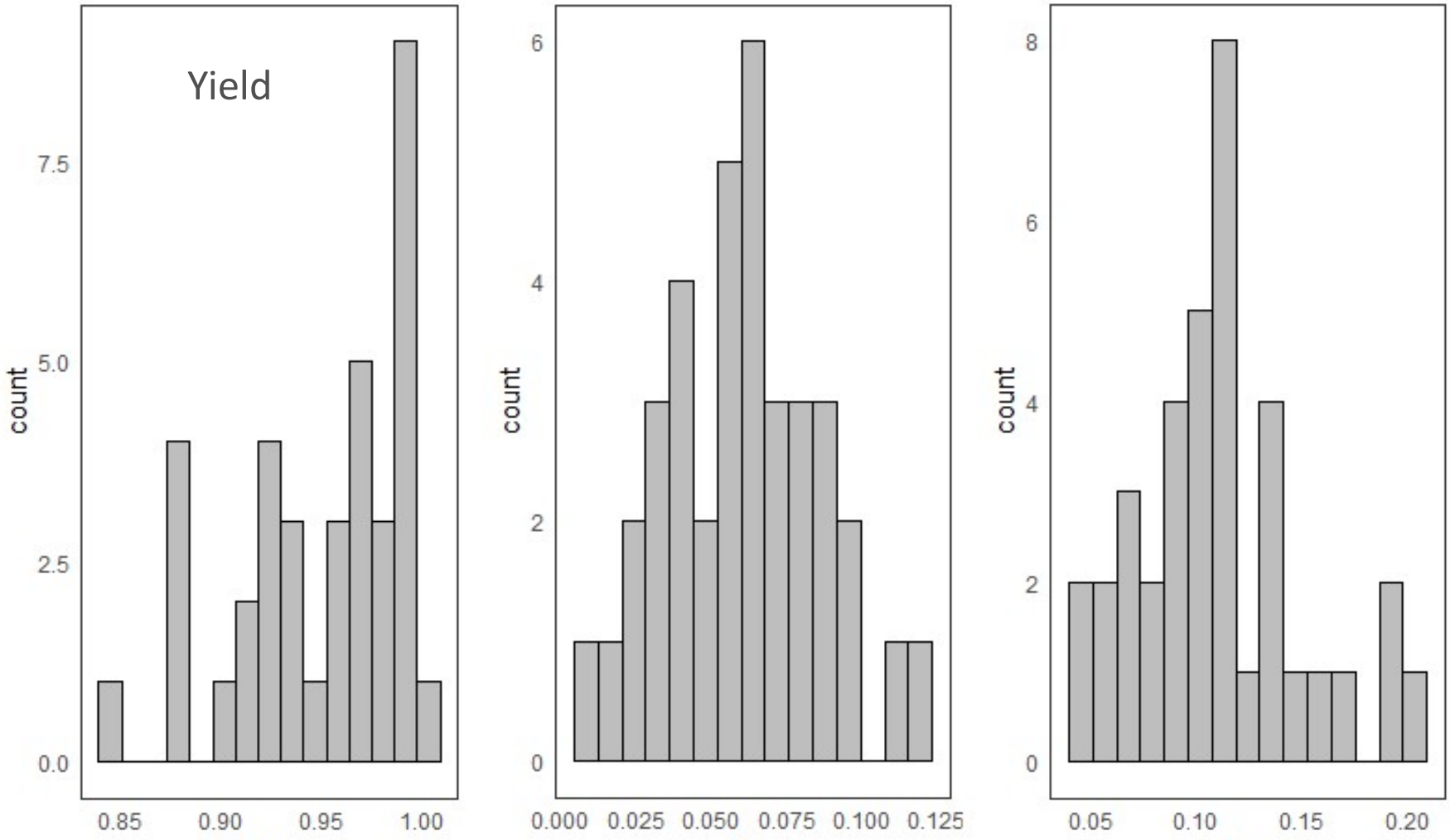
- **Challenge:**
  - process involves many important process parameters
  - simultaneous evaluation of several (normal and non-norma) outcomes

Data

Simulated Data

- Synthesis of an API
- **Input parameters** (e.g. Temperature, Water level)
  - X1 – X6
  - Coded as -1 to 1
- **Outcomes**: Y1 – Y3
  - Yield: Y1
  - By-products/ residue: Y2 – Y3
- Total number of experimental runs: 37

Responses

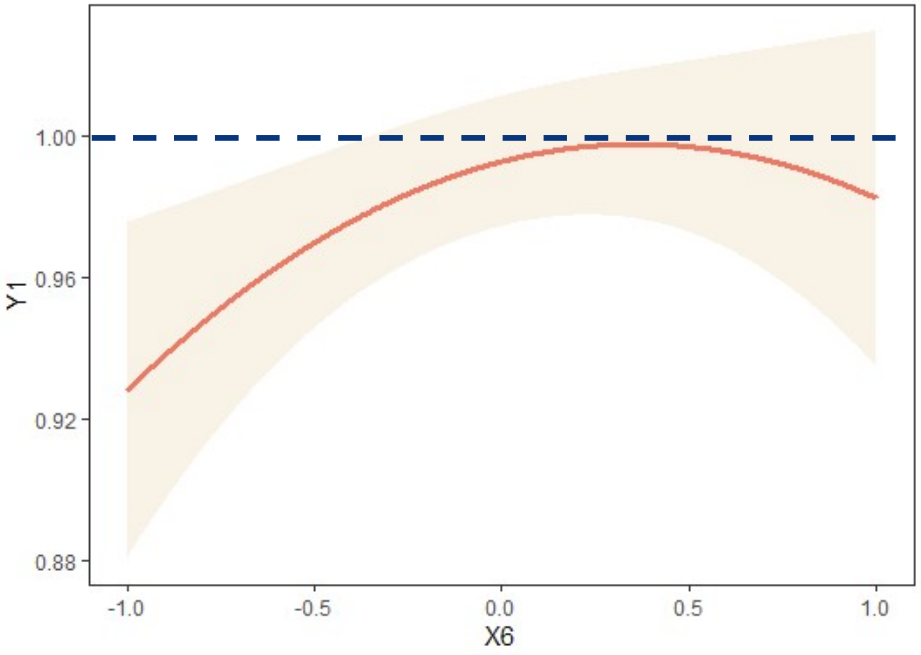

Yield
count
7.5
5.0
2.5
0.0
0.85
0.90
0.95
1.00
count
6
4
2
0
0.000
0.025
0.050
0.075
0.100
0.125
count
8
6
4
2
0
0.05
0.10
0.15
0.20

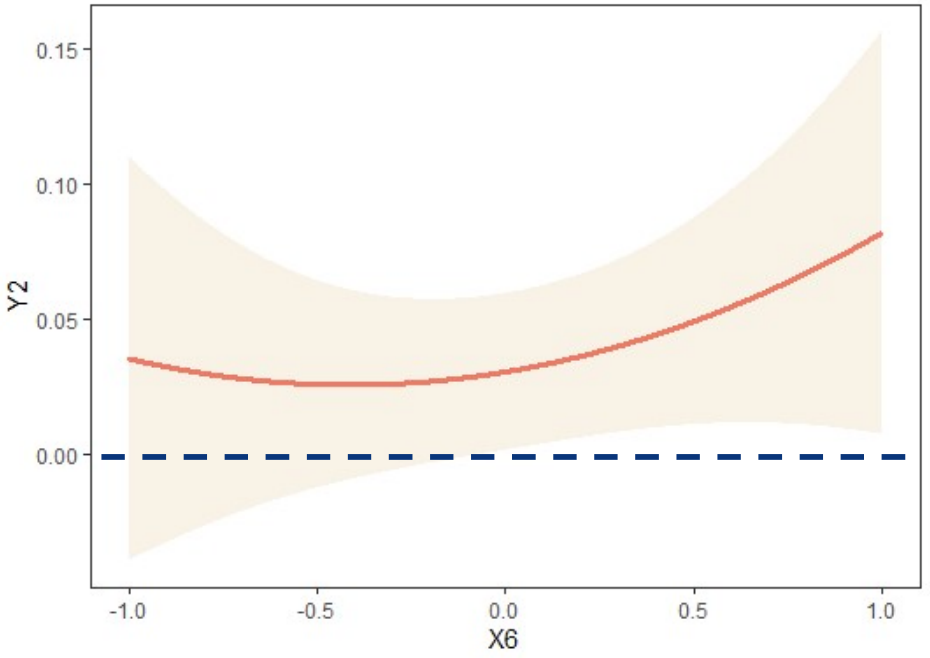
Response Surface Model

Response Surface Model

Second-order polynomial normal regression model

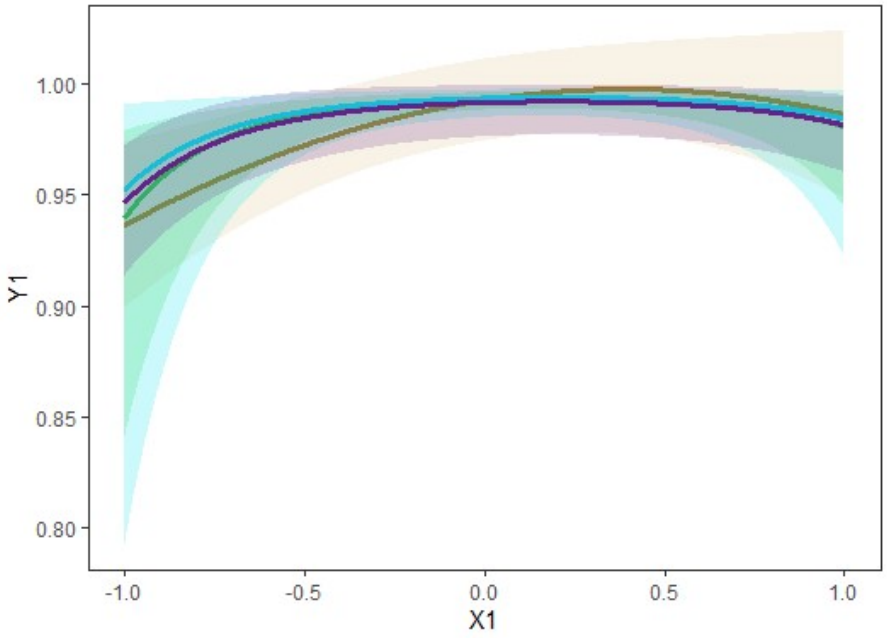
- Large 95% confidence prediction intervals due to relatively large residual error of the model
- Prediction intervals goes out of range of what was rationally possible (e.g. yield prediction interval exceeds 100%, prediction intervals for other by-products goes below 0)

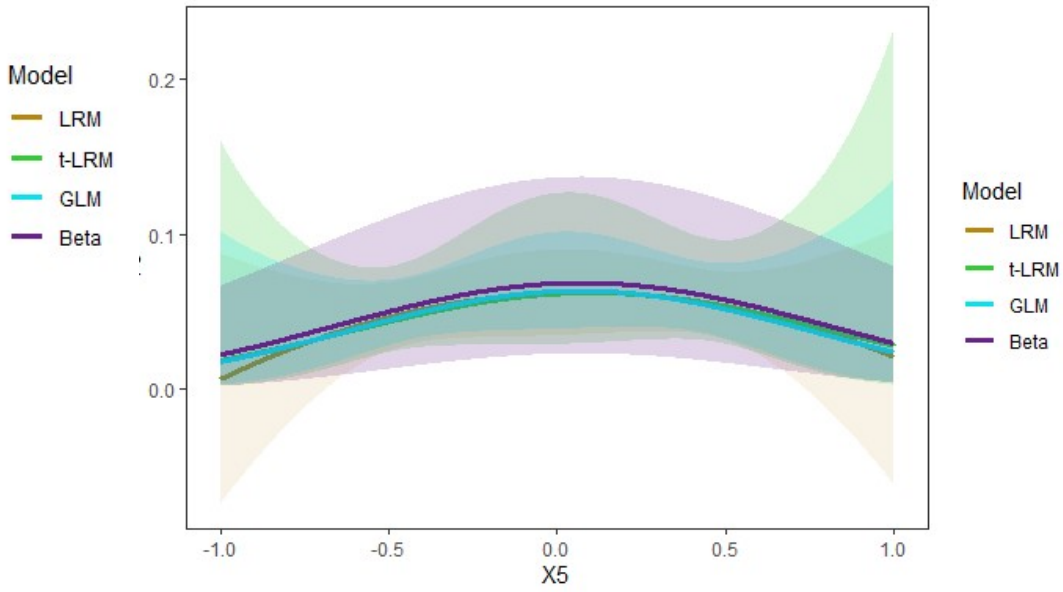




Alternative models

- Alternative models investigated:
  - Logit-transformed linear regression model
  - GLM fractional response model
  - Beta regression model





Beta Regression RSM

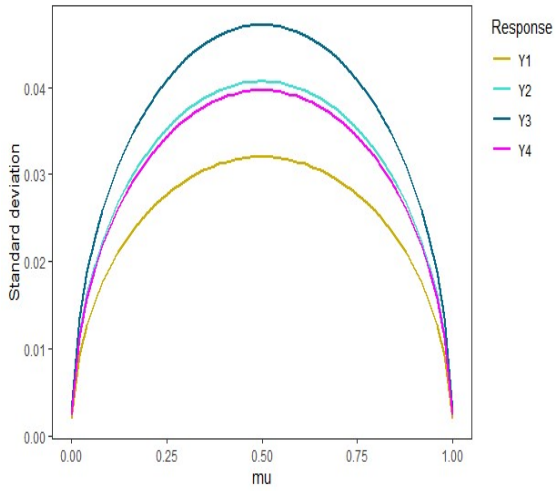
For each response, a beta-distributed second-order polynomial model was considered where,

$$y_r \sim Beta(\mu_r, \phi)$$

$$\text{logit}(\mu_r) = \beta_0 + \sum_{i=1}^{6} \beta_i x_i + \sum_{i=1}^{6} \beta_{ii} x_i^2 + \sum_{i<j} \beta_{ij} x_i x_j$$

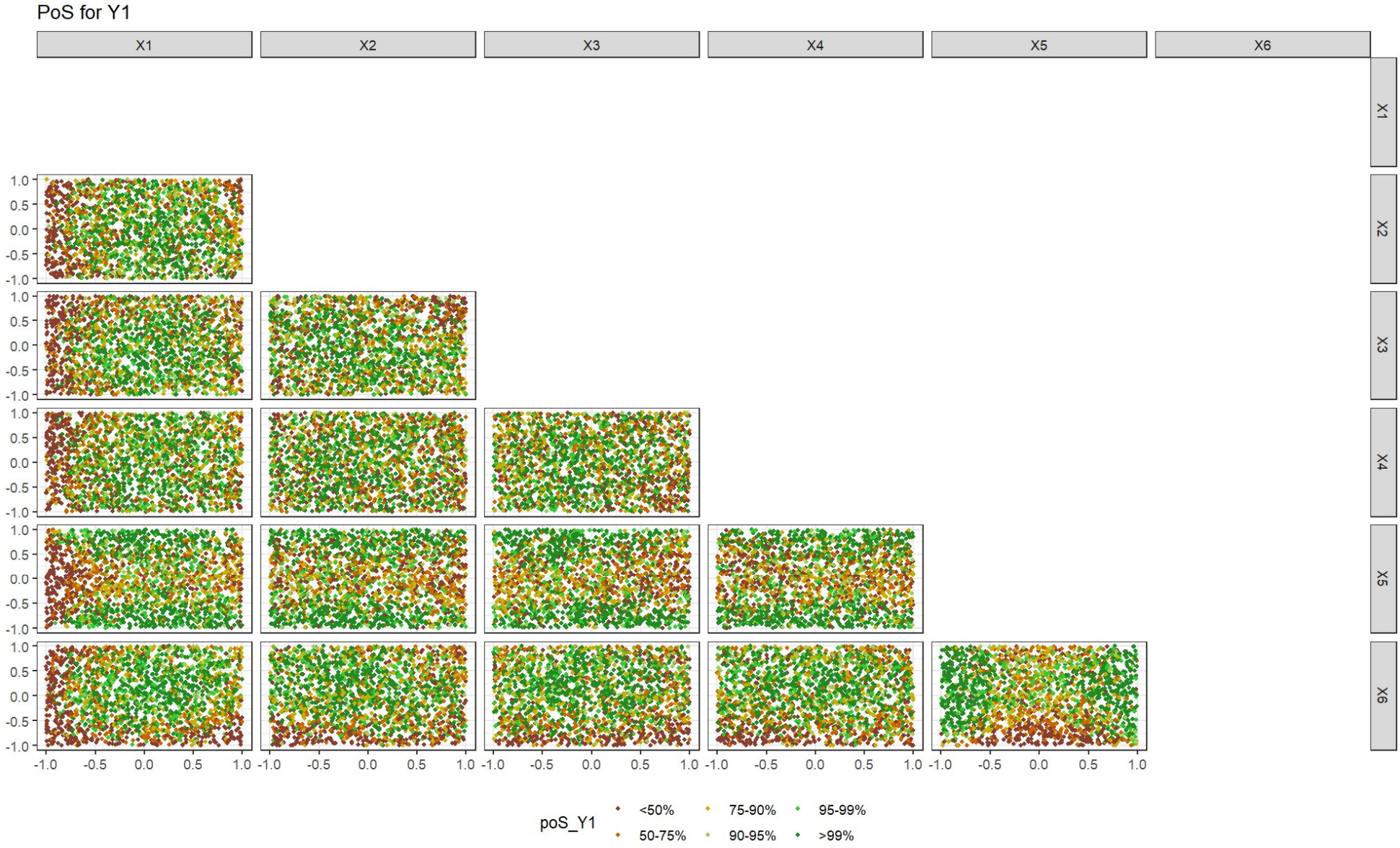
where

- y_k = observed response for experimental run r;
- β_0 = is the overall centre-point mean;
- β_i = represents the linear effect of the ith factor;
- β_{ii} = represents the quadratic effect of the ith factor;
- β_{ij} = represents the interaction effect of the ith and jth factors;
- ϕ = precision parameter, such that: $Var = \mu(1-\mu)/(1+\phi)$.

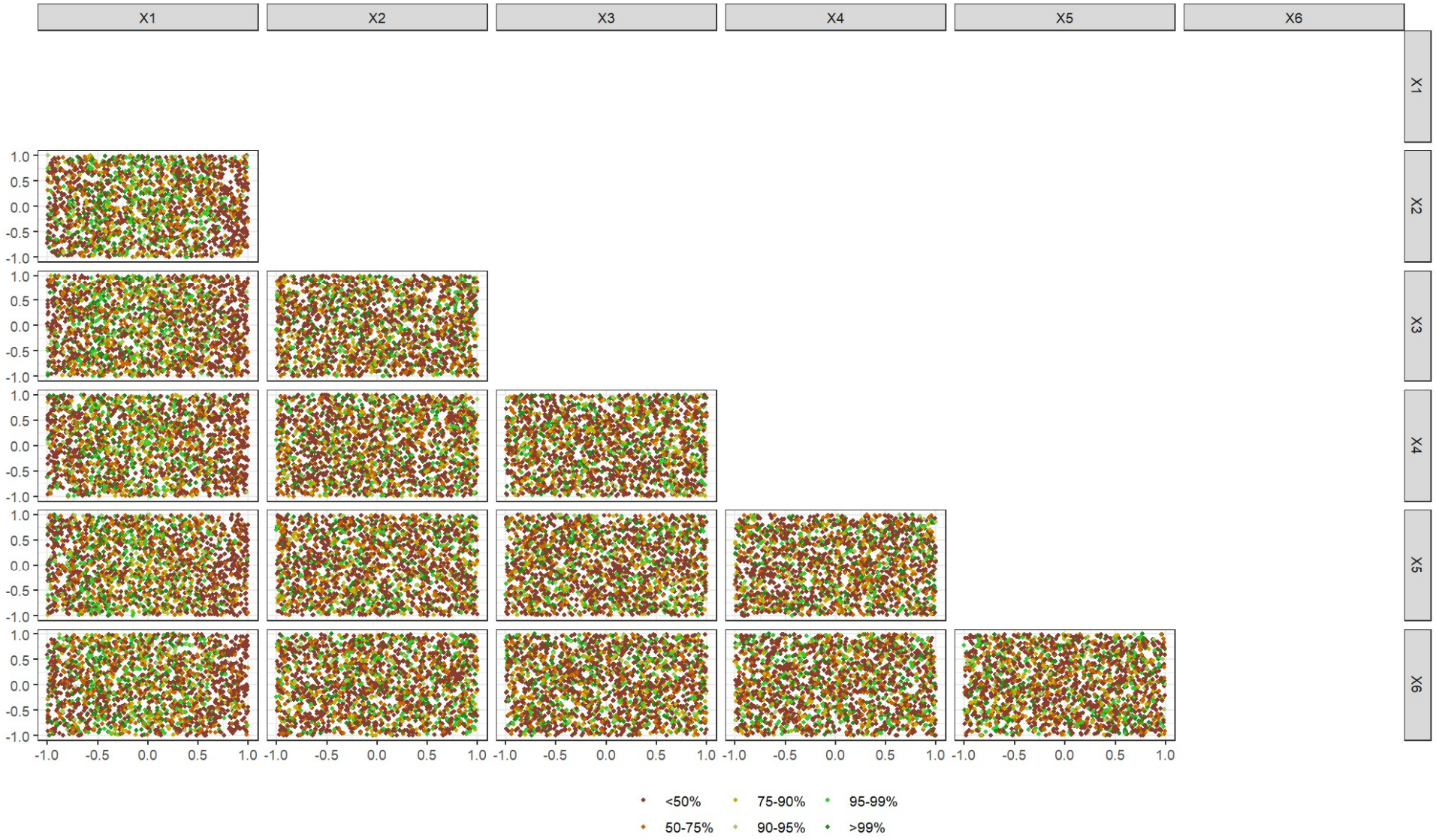


Design Space

P(Yield>0.96):

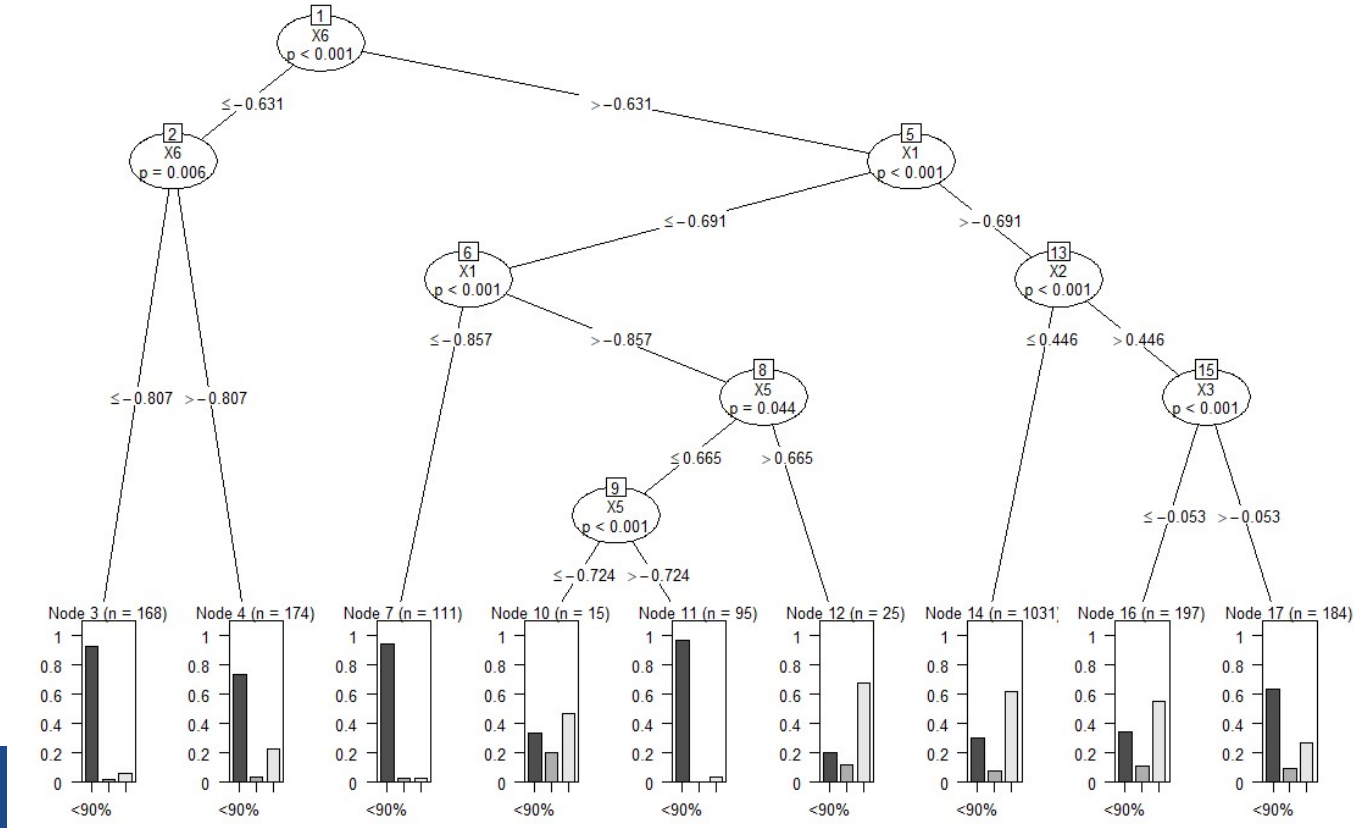

PoS for Y1
X1
X2
X3
X4
X5
X6
1.0
0.5
0.0
-0.5
-1.0
poS_Y1
<50%
50-75%
75-90%
90-95%
95-99%
>99%

P(Y2+Y3<0.21):



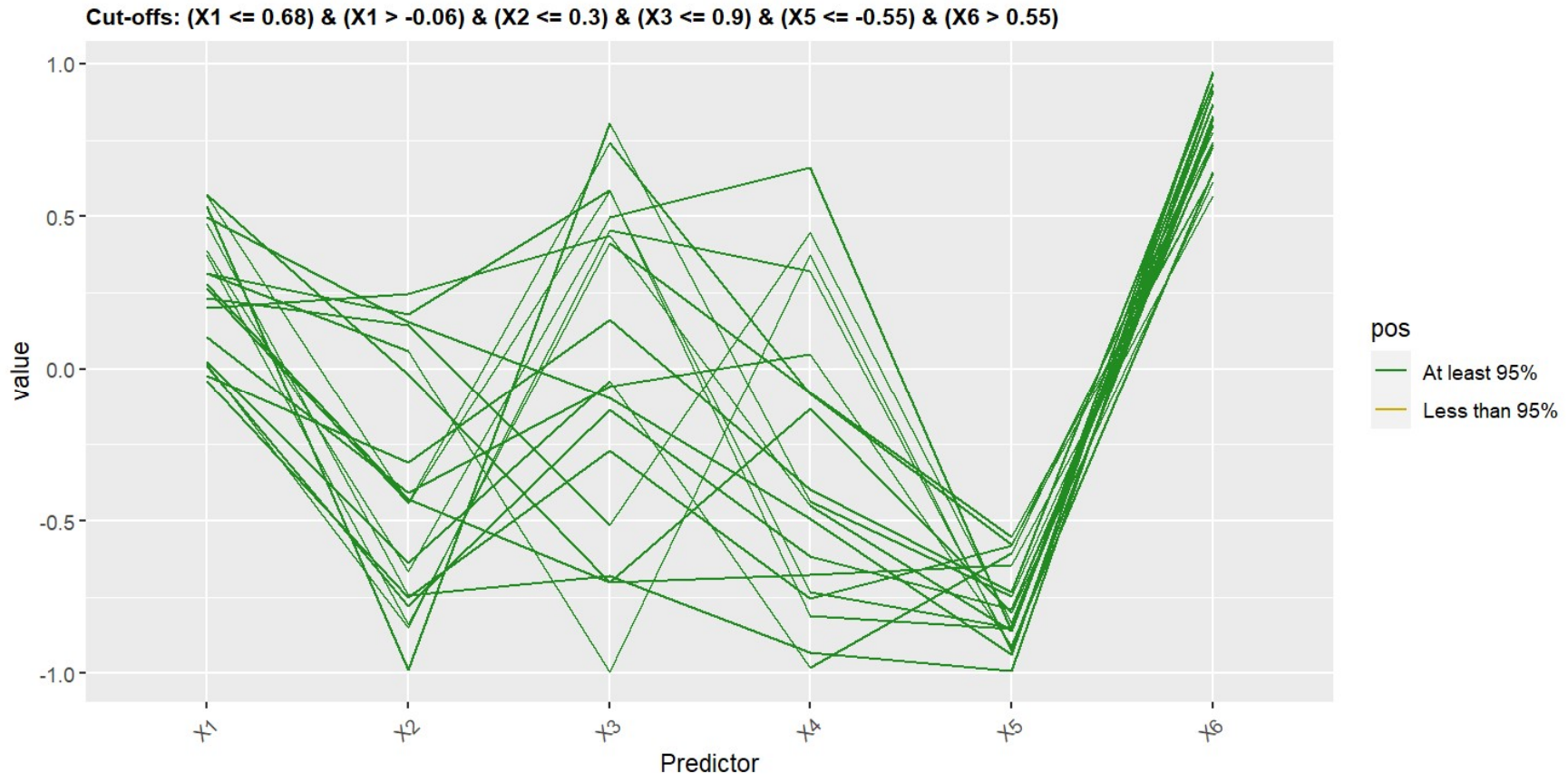
Multivariate classification tree

- To determine initial cut-offs for the design space, multivariate classification tree was fitted, taking the (categorized) PoS as response.
  - Library “ctree” was used
- Some good-performing nodes with the lowest classification error was selected (the error here represents the percentage of observations that are not classified within the >95% group)

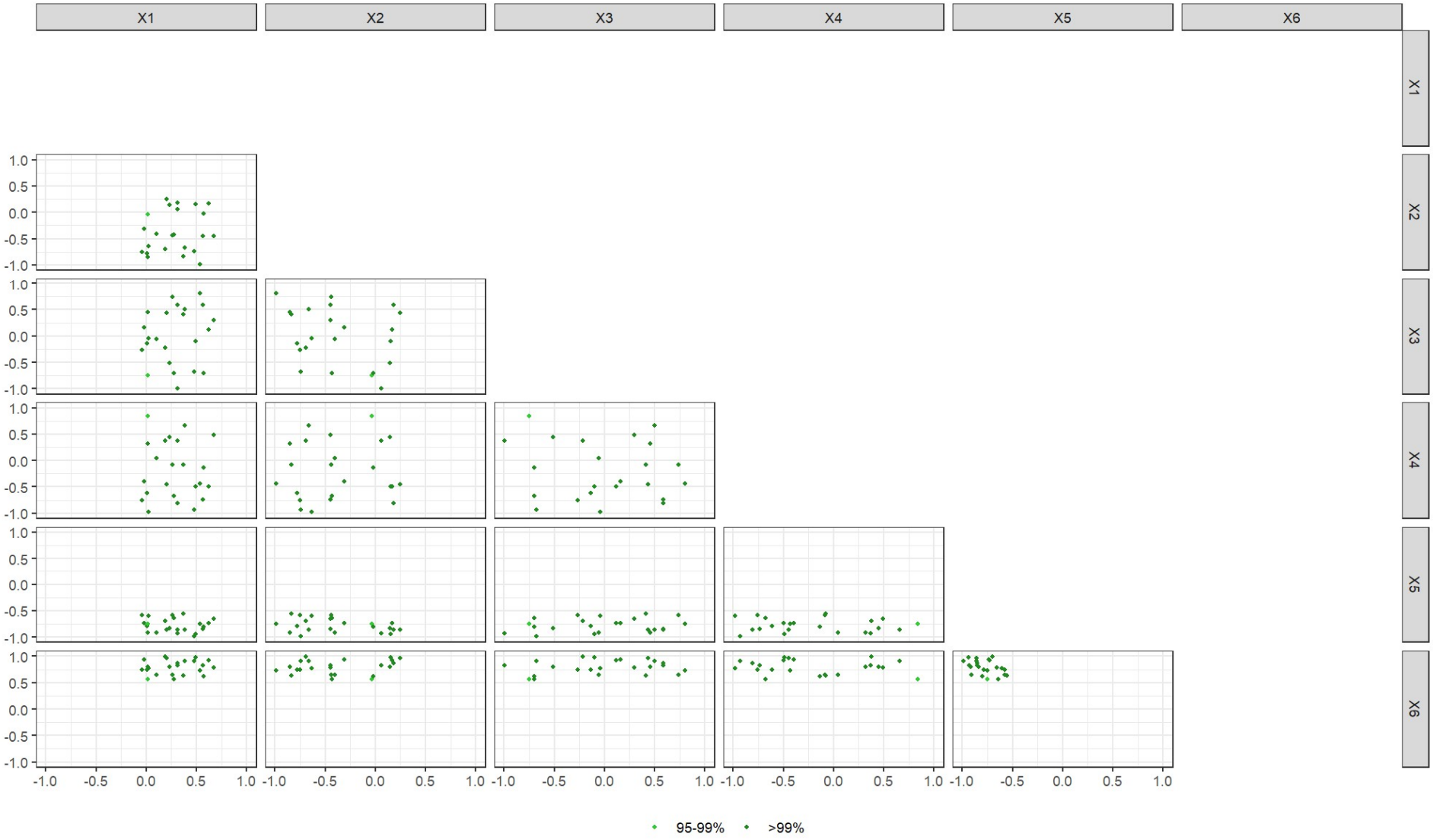


Multivariate classification tree + Parallel coordinates:

- Parallel coordinates plot with the cut-offs selected
- Each line represents a simulated (normalized) predictor combination
  - **Green color** indicates at least 95% probability of meeting the specifications for all criteria,
  - **Gold/orange** line indicates that at least one of the probability was below 95%.



Design Space:



Summary

Summary

- Not properly accounting for the natural data boundary can result to misleading results
- Beta-regression model can easily be applied for response surface modelling
- Multivariate classification tree and parallel coordinates plot are useful tools in multidimensional design space exploration, especially in situations where the design space is not immediately clear

Acknowledgement

- Martin Otava
- Matthew Mower

Jessica Riley, *Shells*
Artwork from Reflections Art in Health

janssen | PHARMACEUTICAL COMPANIES OF Johnson & Johnson