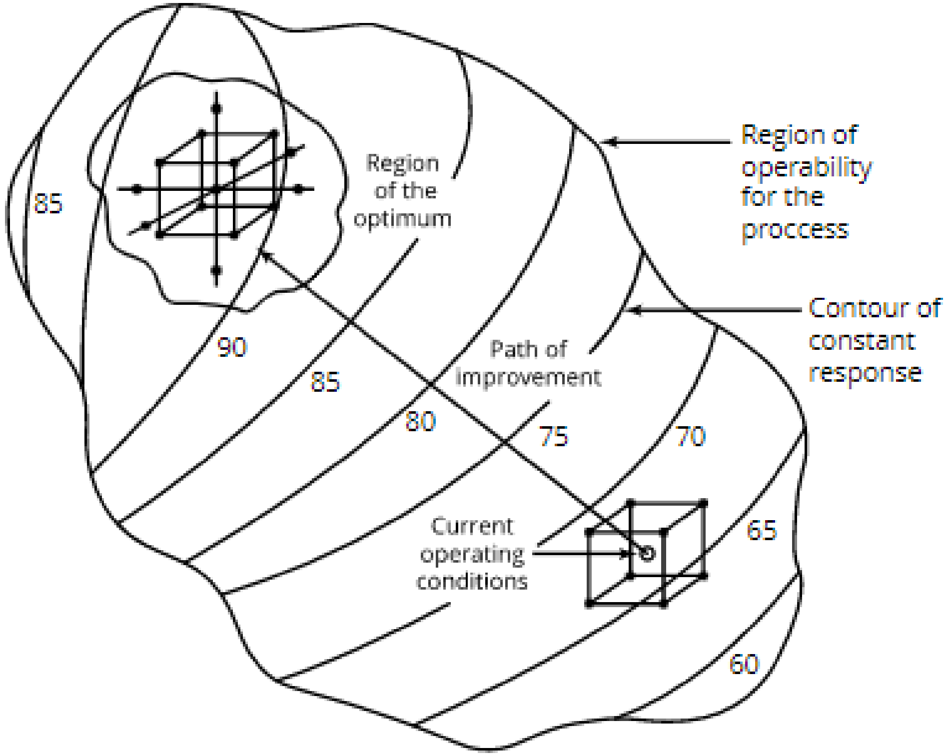
Optimization using DoE

Response Surface Methods (RSM)

- We are now shifting from screening designs to trying to optimize a process
- This is done by looking at the factor level combinations that give us the maximum yield at a minimum of cost.
- The objective of Response Surface Methods (RSM) is optimization, finding the best set of factor levels to achieve some goal

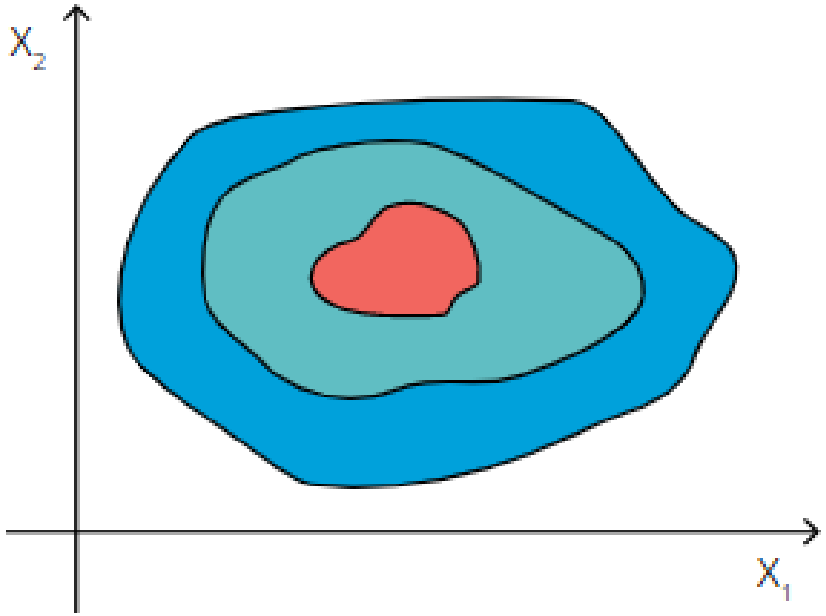
RSM as a sequential process

The figure depicts a response surface method in three dimensions, though it is actually represented in four dimensional space since the three factors are in 3-dimensional space the response is the 4th dimension.



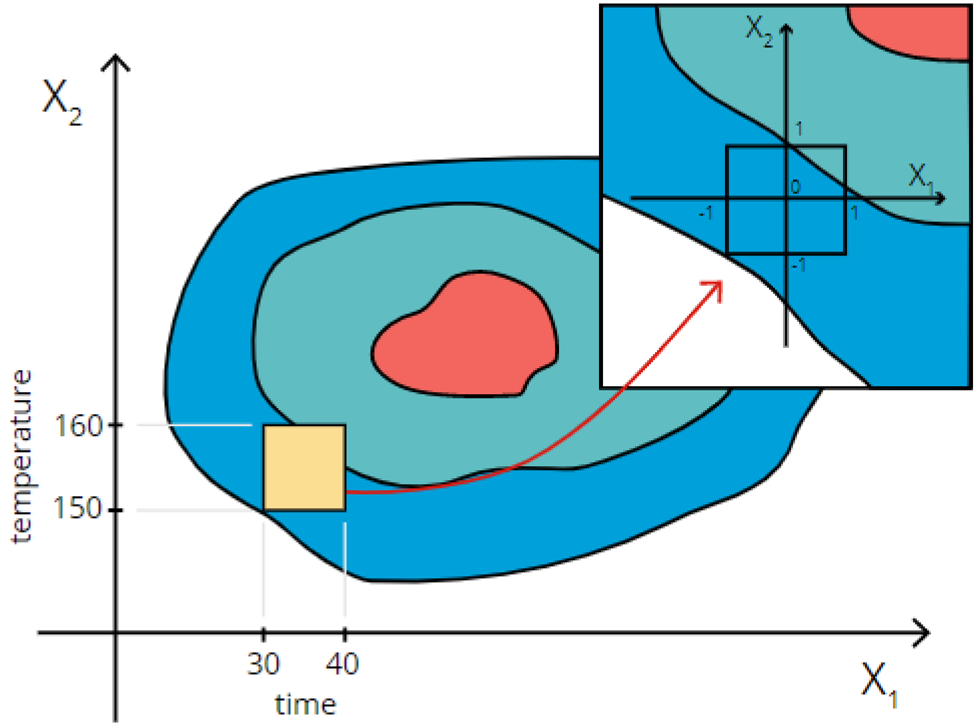
RSM as a sequential process

- Let us instead start in 2D
- The response surface here is an ideal case where there is a 'hill' which has a nice centered peak.
- Our task is to find factor level values where the response is at its peak



RSM as a sequential process

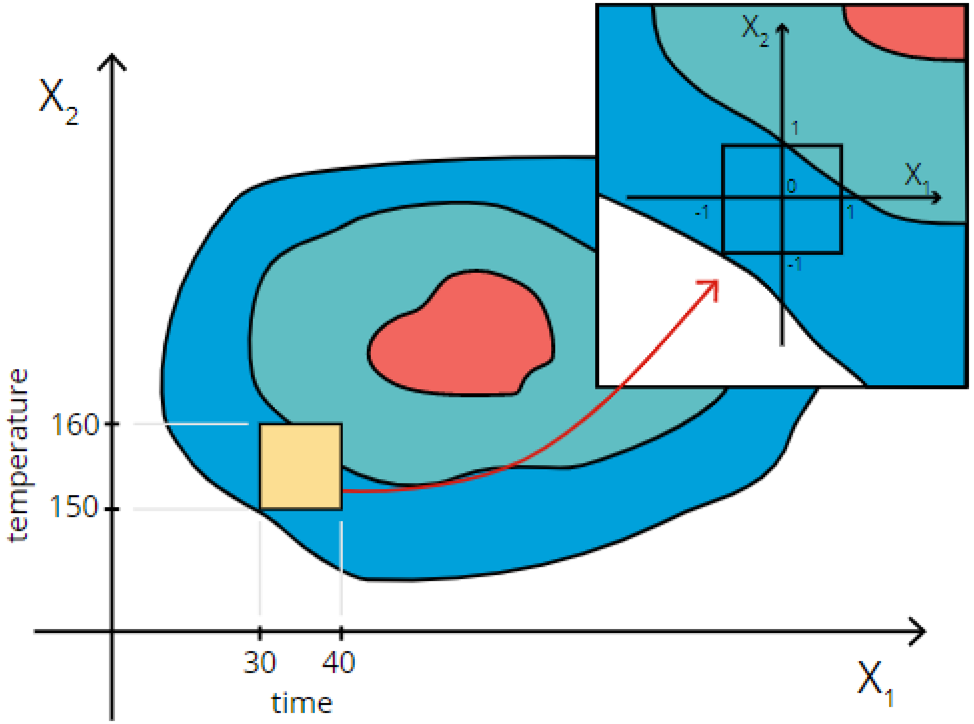
- Let us instead start in 2D
- The response surface here is an ideal case where there is a 'hill' which has a nice centered peak.
- Our task is to find factor level values where the response is at its peak
- Our goal is to start somewhere using our best prior or current knowledge and search for the optimum spot where the response is either maximized or minimized.



RSM as a sequential process

The models we will use are:

- Screening response model
- Steepest Ascent model
- Optimization model



Screening Response model

- For a first order model, it involves linear effects and a single cross product factor.
- Obtaind from a 2^2-factorial design

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

Steepest Ascent model

- If we ignore the cross product which gives an indication of the curvature of the response surface and just look at the first order model, this is called the steepest ascent model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Optimization model

- When we are somewhere near the top of the hill, we fit a second-order model. This includes the two second-order quadratic terms

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Example… Jupyter notebook

Multiple Responses

- Compromises need to be made
- Desirability function

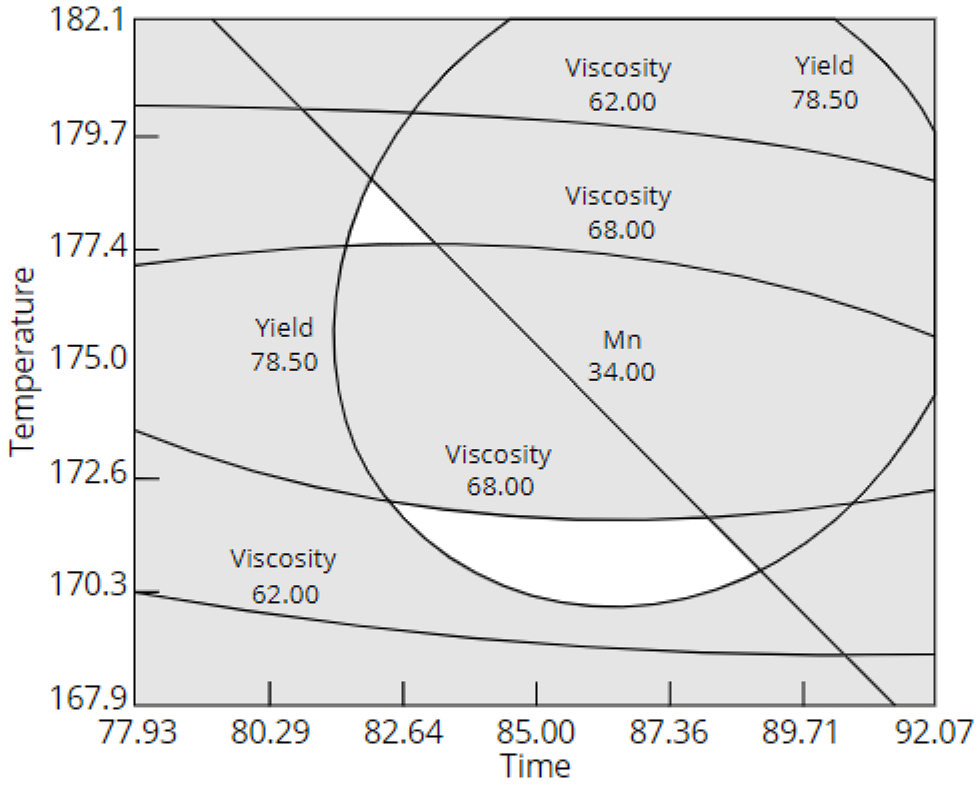
individual

$$d = \begin{cases} 0 & y < L \\ \left(\frac{y-L}{T-L}\right)^r & L \le y \le T \\ 1 & y > T \end{cases}$$

$$D = (d_1 d_2 \dots d_m)^{1/m}$$

Region of optimum found by overlaying yield, viscosity and molecular weight response surfaces

overlaid contour plots



The unshaded area is where *yield* > 78.5, 62 < *viscosity* < 68, and *molecular weight* < 3400

Figure 11.16 (Design and Analysis of Experiments, Douglas C. Montgomery, 7th Edition)