A hybrid system for design space estimation of the continuous tableting process

Dr. Mohammad Shahab

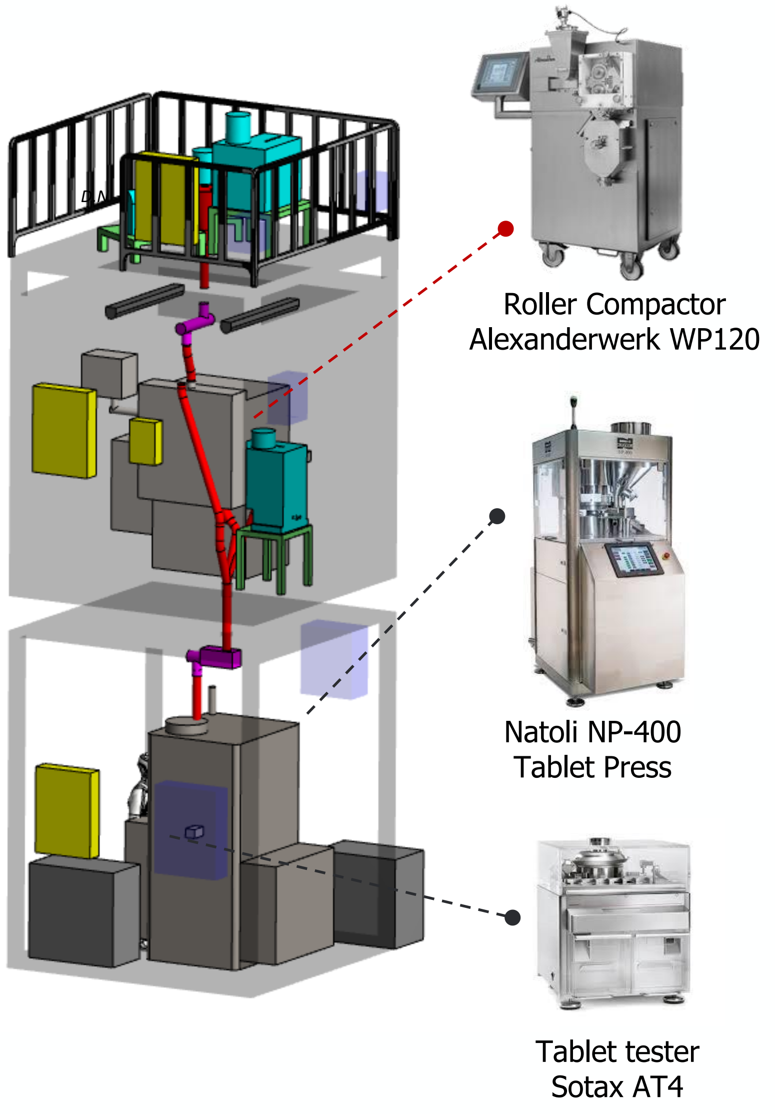
Postdoc Research Associate

Advisors: Profs. Gintaras Reklaitis & Zoltan Nagy

PURDUE UNIVERSITY | Davidson School of Chemical Engineering

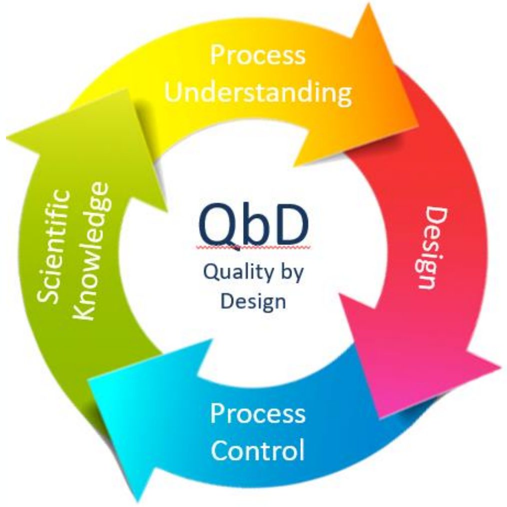
Outline

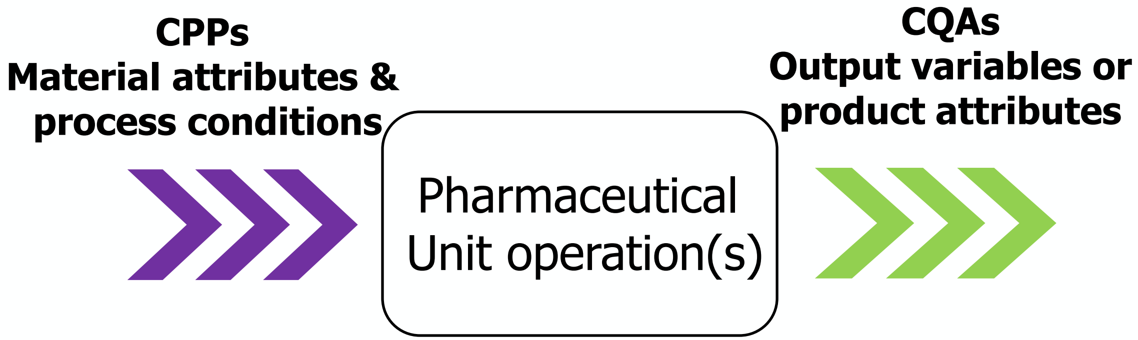
- Introduction/Motivation
- Continuous tableting & Experimental Procedure
- Tablet Press Semi Mechanistic Model
  - Effects of Glidants and Lubricants
- Design Space Estimation
- Multi-Response Surface Model
  - Effects of Lubricants
- A hybrid system
- Summary



Introduction

- Quality by Design (QbD) is a strategic approach employed in various industries, including pharmaceuticals, manufacturing, and product development, to ensure the consistent delivery of high-quality products.
- QbD leads to a systematic understanding of intricate relationships between CPPs and CQAs for the manufacture of drugs within a broad operating regime called the design space.
- The regulatory approval provides boundaries within which CPPs can change without further approval.





Approach to design space identification

1. Identifying the knowledge space
2. DoE measuring the relations between CPPs and CQAs within the knowledge space
3. Computational methods to determine the boundaries of the design space
4. Validation experiments to confirm the design space.

Design Space

Design space (DS) - "the multidimensional combination and interaction of input variables (e.g., materials attributes) and process parameters that have been demonstrated to assure quality." – ICH Q8 guideline, 2009

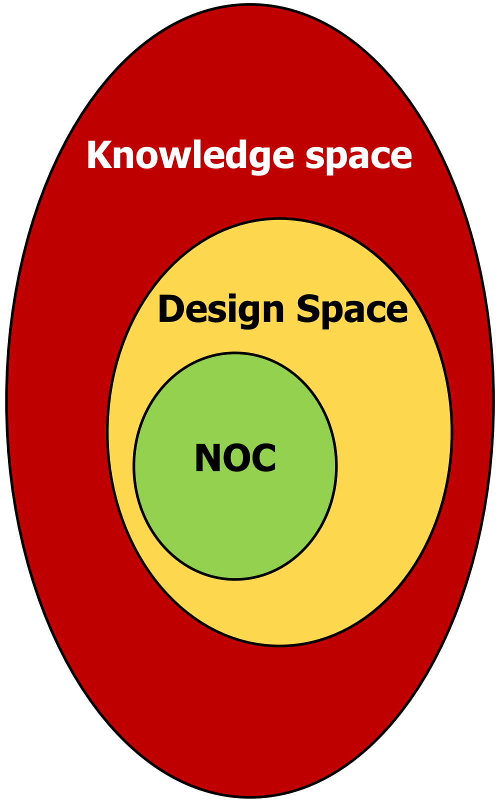
- The design space defines the range of CPPs that ensure the product's CQAs are within acceptable limits.
- It provides flexibility for process optimization while maintaining the required quality attributes.
- Within this design space, manufacturers can establish appropriate process controls, monitoring techniques, and quality assurance systems to ensure consistency and predictability in product performance.

Broad methodic approaches to design space:

1. **Optimization algorithms:** Identify operational parameters, which ensures that the process acquiesces to a constraint set by performing the flexibility analysis.
2. **Sampling algorithms:** Limited process data and variability and can incorporate prior knowledge about the process to define a design space that includes uncertainty.

- Widely available literature uses an empirical approach for DS estimation.
- Empirical approach is highly favorable when **model is complicated** and DS analysis is computationally challenging.

-Multiple unit operations
-Integrated flowsheet, etc.



Third paradigm

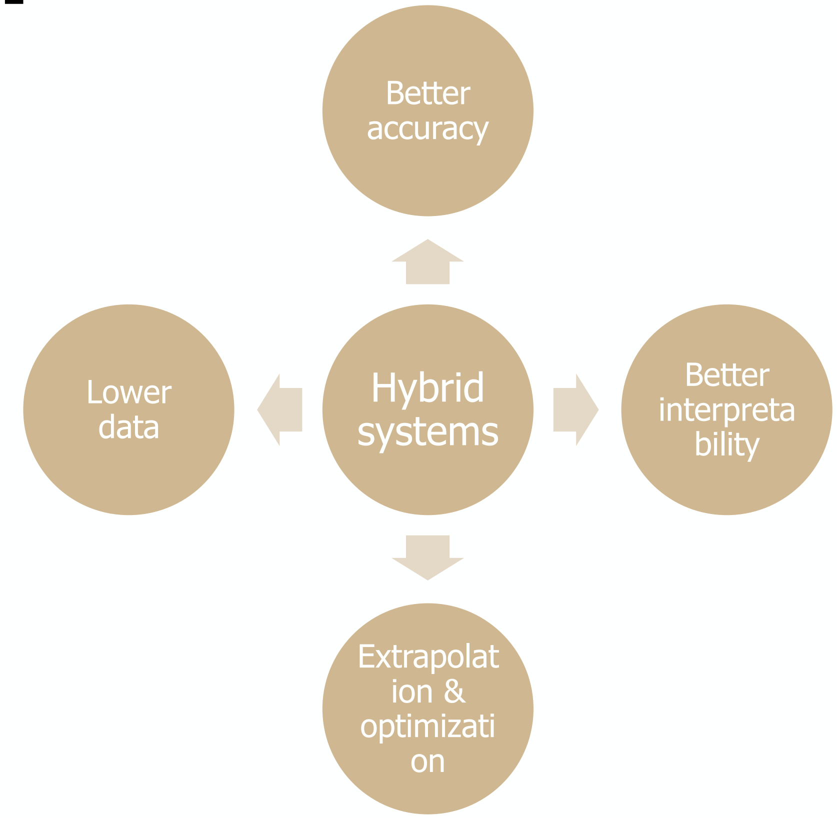
- In the presence of inexpensive comp. model, the original mechanistic model can be directly used to compute design space.
- Equations in these mechanistic models are constructed from a series of presumptions about the physical systems and conservation principles of physics and chemistry.
- A third paradigm exists when a trade-off is found between **the flexibility to discover possibly unknown interactions** between the process conditions and **restricting their associations**.
- Hybrid systems are quite helpful for systems where several contributing components and intricate relations are only partially understood.
- This paradigm looks at how the mechanistic and data-driven approaches may work together, utilizing information obtained from experimental data as well as process expertise already in place.

3rd paradigm → Pharmaceutical processes

Resource-intensive

Fundamentally unresolved

- **Objective to develop a hybrid system to estimate DS for continuous tablet press (TP) unit operation where the effects of process conditions, glidants, and lubricants are investigated to achieve the desired tablet CQAs.**



Continuous tableting line

Direct compaction line at Purdue's Pharmaceutical Continuous Manufacturing Pilot Plant (FLEX lab)

Active ingredient:
Acetaminophen (APAP) 10%
Excipient:
Microcrystalline Cellulose (MCC)

K-Tron and Schenck Feeders

Control Variables
- **Flow rate**
- **Mixing time**

Control Variables
- **Turret speed**
- **Feeder speed**
- **Dosing position**
- **Pre-compression thickness**
- **Main compression thickness**

Process Variables
- **Main compression force**
- **Pre-compression force**
- **Ejection force**

Product data
- **Weight**
- **Thickness**
- **Diameter**
- **Breaking force**

K-Tron MT12
Micro-Screw Feeder

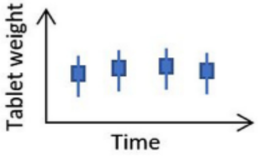
Glidant:
Colloidal Silica (0-0.2%)
Lubricant:
Magnesium Stearate (MgSt) 0-2%

Natoli NP-400
Tablet Press

Tablet

At-line (5 min)

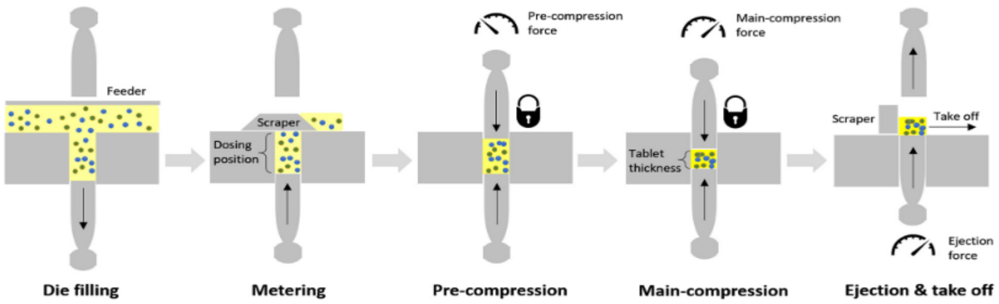
Sotax AT4



Product data: Weight, thickness, diameter, breaking force

- **30 steady state** experiments for Silica and **20 steady state** experiments for MgSt were performed with Latin Hypercube experimental design.
- Mechanistic reduced order models to integrate lubricant and glidant feeding with tableting were developed
- These models can be used to create DOE to design a continuous operation and implement end-to-end control

Semi-mechanistic tablet press model



Integration of Glidants and Lubricants:

Glidant concentration (0-0.2%) - c_L , Shear strain - γ
Lubricant concentration (0-2%) - c_L , Shear strain - γ

Tablet weight and production rate

$$\frac{W}{\rho_b V^{\text{fill}}} = -\xi_1 \frac{n_{\text{F}}}{n_{\text{T}}} + \xi_2 \frac{H^{\text{fill}}}{D} + \xi_3 \left(\frac{H^{\text{fill}}}{D}\right)^2 \longrightarrow \dot{m} = N_{\text{d}} n_{\text{T}} W$$

At higher H/D the filling efficacy diminishes. As the n_F/n_T ratio increases, the filling efficacy decreases

In-die relative density and compaction force (modified Kawakita model)

$$\rho^{\text{in-die}} = \frac{W}{\rho_t V^{\text{in-die}}} \longrightarrow \sigma_{\text{punch}} = \frac{4F_{\text{punch}}}{\pi D^2} = \frac{\rho^{\text{in-die}} - \rho_c}{[\rho^{\text{in-die}}(a-1) + \rho_c]\, b}$$

Compaction force decreases with increase in concentration

$$a = \frac{a_0 - a_\infty}{1 + C_{\text{c}}} + a_{0,\infty} \longrightarrow$$

Does not apply for the glidant case

Elastic recovery and tablet density (from Bommireddy-Gonzalez)

$$\epsilon_\rho = \epsilon_0 \frac{\rho^{\text{in-die}} - \rho_{\text{c},\epsilon}}{1 - \rho_{\text{c},\epsilon}} \longrightarrow \rho^{\text{tablet}} = \rho^{\text{in-die}}(1 - \epsilon_\rho)$$

Elastic recovery increases with increase in concentration

Tensile strength (modified Leunberger's model)

$$\sigma_{\text{t}} = \sigma_0 \left[1 - \left(\frac{1 - \rho^{\text{tablet}}}{1 - \rho_{\text{c},\sigma_{\text{t}}}}\right) e^{(\rho^{\text{tablet}} - \rho_{\text{c},\sigma_{\text{t}}})}\right]$$

$$C_c = \frac{c_L^{p_1}(\gamma + \gamma_0)^{p_2}}{p_3}$$

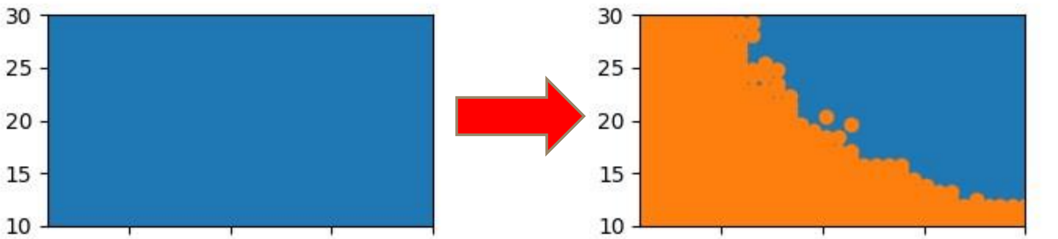
Tensile strength decreases with an increase in concentration or mixing time

S. Bachawala, M. Gonzalez, "Development of mechanistic reduced order models (ROMs) for glidant and lubricant effects in continuous manufacturing of pharmaceutical solid-dosage forms", ESCAPE32, June 12-15, 2022, Toulouse, France.

DS estimation - semi mechanistic model

- Deterministic DS:
  - Create a fine mesh of sample points in CPP space
  - Simulate each of these points to determine if the predicted quality of the product violates any CQA constraint

- Probabilistic DS:
  - At each CPP, run Monte Carlo simulations by sampling model parameter values within their uncertainty bounds.



Deterministic   Deterministic + Probabilistic

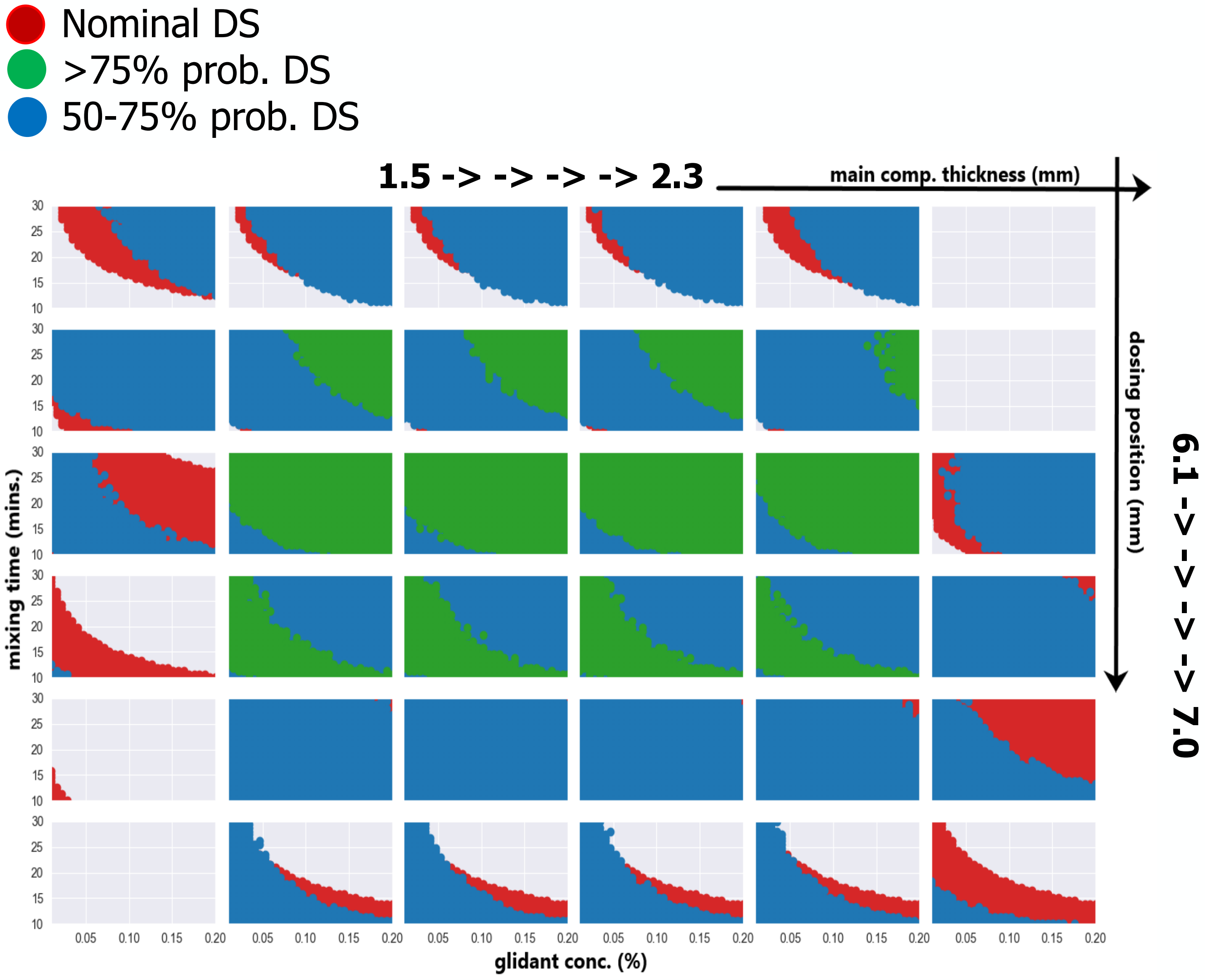
- Specify the realizable bounds of the CPPs

CPP	Low limit	High limit
Dosing position (h_{fill}, mm)	6	11
Comp. thickness (h_{in-die}, mm)	1.5	3
Silica conc. (c_l, %)	0.1	0.2
MgSt conc. (c_l, %)	0.1	2
Mixing time (γ, mins.)	10	30

- The CQAs and their desired specifications

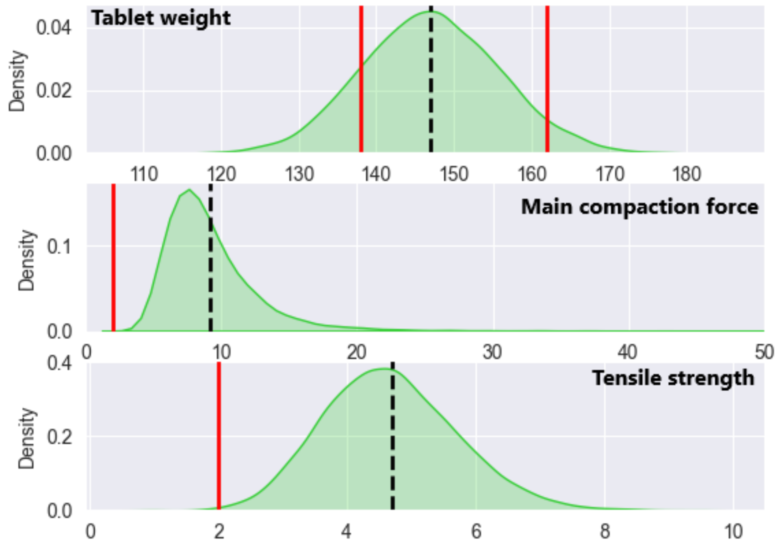
CQA	Low	High
Tablet weight (W, mg)	138	162
Compression Force (F_{main}, kN)	2	50
Tensile strength (σ_t, MPa)	2	15

Design space for glidant effects

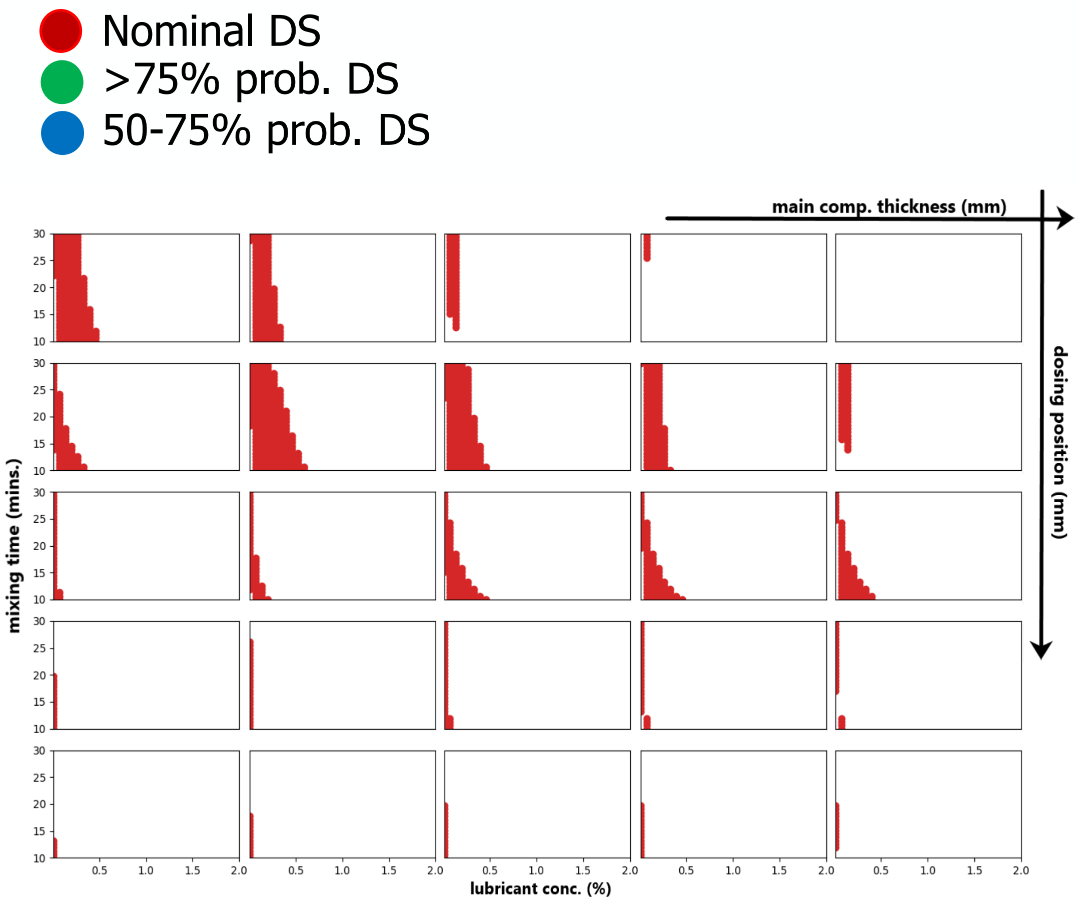


CPP	Feasible region
Dosing position (h_{fill}, mm)	**6.2-6.7**
Comp. thickness (h_{in-die}, mm)	**1.6-2.1**
Silica conc. (c_L, %)	**0.01-0.2**
Mixing time (γ, mins.)	**10-30**

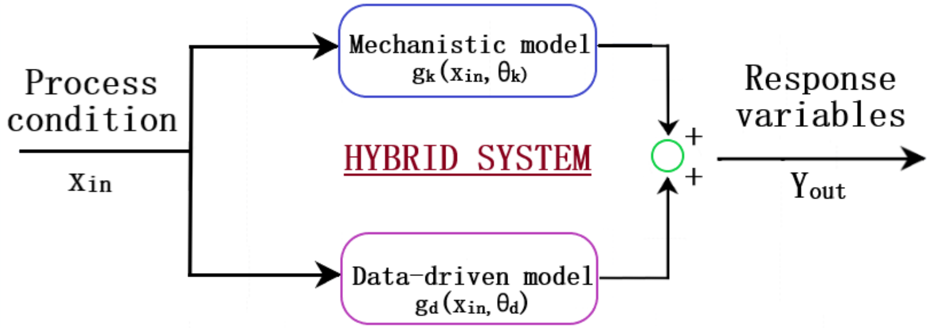
Optimal CPP (h_{fill}, h_{in-die}, c_L, γ) – (6.28, 1.82, 0.194, 28)



Design space for lubricant effects



- Maximum DS contours are minimal with nominal values
- When model uncertainty is included, the mechanistic TP model fails to generate any significant DS.
- Higher model parameters; high non-linearity; high model uncertainty; lower experimental samples.
- For some powder blends, more experimental samples might be needed to produce quality products.
- Limited resources and complex machinery is a challenge.
- In such cases, a data-driven alternative is the best choice for model development and validation.



Multi-Response Surface modeling

- RSM is a statistical technique that relates controllable variables to response(s) through an empirical model **that is not available or is very complex**.
  - Generates knowledge in the experimental domain of interest
  - Suggests sequential strategies for conducting experiments with different alternatives based on the results (Model-based DoE)
  - Allows decision-making under uncertain conditions.

$$\boldsymbol{y} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_1} + \boldsymbol{\beta_2 x_2} + \boldsymbol{\beta_3 x_1^2} + \boldsymbol{\beta_4 x_1 x_2} + \ldots$$

- MRS techniques
  - Dimension reduction – the absence of correlation among multiple responses
  - Loss function – does not consider process parameter fluctuations, model parameter uncertainty
  - Posterior predictive function – does not consider model parameter uncertainty
  - OLS, WLS for each response.

- Such models fail to perform when the responses are correlated and precise estimates are needed.

Seemingly Unrelated Regression (SUR) Models

- Seemingly unrelated regression (SUR) is very useful when response variables in a multi-response RSM are correlated.
- SUR yields more precise estimates when response correlation exists.
- Multi response experiments:

$$y_i = \boldsymbol{X}_i \boldsymbol{\beta}_i + \varepsilon_i\, , i = 1, 2, \dots, r \qquad \textbf{(1)}$$
$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_1 x_2 + \dots$$

*y_i is an n×1 vector of observations, $\boldsymbol{X}_i$ is an n×p_i matrix of known functions, $\boldsymbol{\beta}_i$ is a p_i×1 vector of unknown parameters, and ε_i is a random error vector **on ith response**.*

- In SUR, it is assumed that errors for any individual model have const. variance but the errors in different models are correlated. (E(ε_i) = 0; Var(ε_i) = $\sigma_{ii} I_n$; Cov(ε_i, ε_j) = $\sigma_{ij} I_n$)
- The *r* polynomial regression models can be converted to a matrix form $\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ **(2)**

Y = [y_1' : y_2' :…: y_r']′ ; β = [β_1' : β_2' :…: β_r']′ ; ε = [ε_1' : ε_2' :…: ε_r']′ ; X is the block diagonal matrix $\begin{bmatrix} X_1 & 0 & 0 \\ 0 & X_2 & 0 \\ 0 & 0 & X_r \end{bmatrix}$

FGLS: $\beta_{est} = [X'(\Sigma_{est}^{-1} \otimes I_n)X]^{-1}\, X'(\Sigma_{est}^{-1} \otimes I_n)Y$ *and*

$\Sigma_{est} = \sigma_{ij} = \frac{(Y_i - X_i\beta_i)^T (Y_j - X_j\beta_j)}{n}$*, i, j = 1,2, … r* **(3)**

Shah, H.K. et al., 2004. Response surface modeling and optimization in multiresponse experiments using seemingly unrelated regressions. Quality Engineering, 16(3), pp.387-397.

RSM methodology

1. The RSM is fitted for each response variable to determine the primary form of the response model

   i. The primary form (polynomial complexity) is obtained by splitting the data into train and validation sets and the complexity (degree) which yields minimum RMSE across both data sets is the required form. (RMSE vs. complexity plot)

   ii. The multivariate multiple linear regression model is obtained once the complexity is defined. (A degree of 2 is considered optimal)

2. OLS is used to compute initial parameter estimates and residuals, and the quality of fit is evaluated.
3. The statistical significance analysis of parameter estimates is done based on a 5% significance level.
4. The estimation of β and Σ for each response is obtained by fitting the models through SUR.

- Specify the realizable bounds of the CPPs

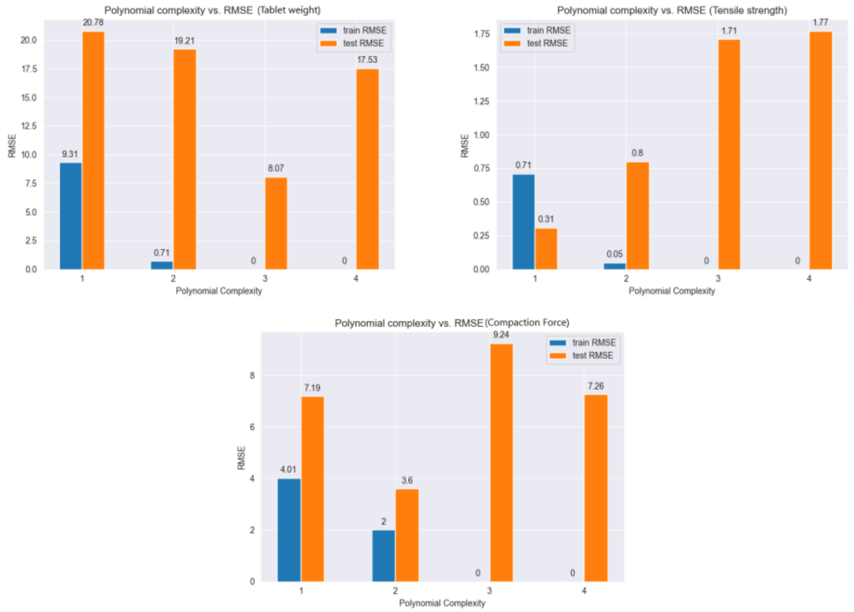
CPP	Low limit	High limit
Dosing position (h_{fill}, mm)	6	11
Turret speed (n_T, rpm)	10	20
MgSt conc. (c_L, %)	0.1	2
Mixing time (γ, mins.)	10	30

- The CQAs and their desired specifications

CQA	Low	High
Tablet weight (W, mg)	138	162
Compression Force (F_{main}, kN)	2	50
Tensile strength (σ_t, MPa)	2	15

SUR model development

Polynomial complexity



PC = 2

Final SUR model

n_T (x_1), dos. pos. (x_2), c_L (x_3), shear time (x_4)

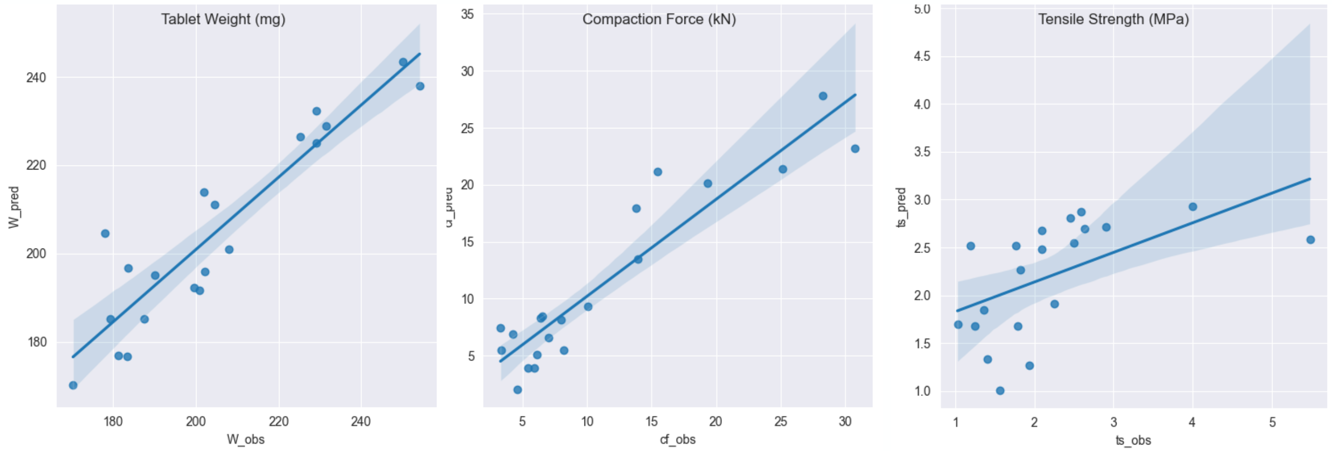
Lubricants SUR model: 13 (27*) params

$\mathbf{Y_{weigh}} = \boldsymbol{f}(x_1, x_1^2, x_1x_3, x_2^2, x_3^2, x_3x_4)$

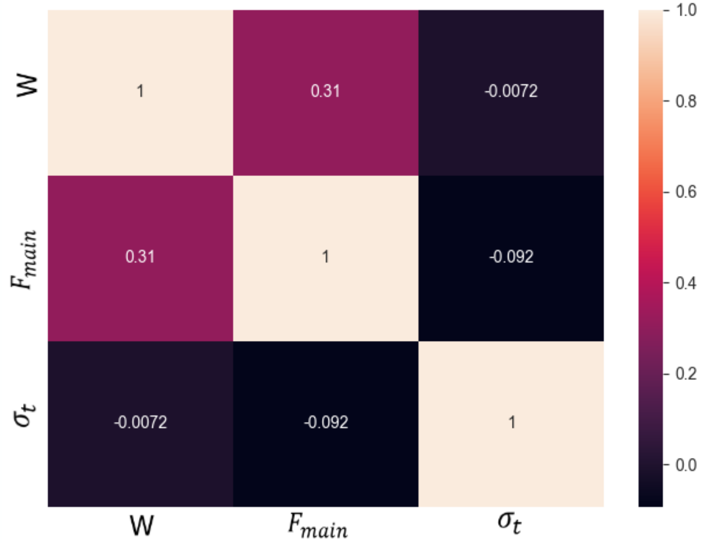
$\mathbf{Y_{cf}} = \boldsymbol{f}(\mathrm{C}, x_1, x_2, x_1x_2)$

$\mathbf{Y_{tens}} = \boldsymbol{f}(x_1, x_1^2, x_3^2)$

Predicted vs. observed values

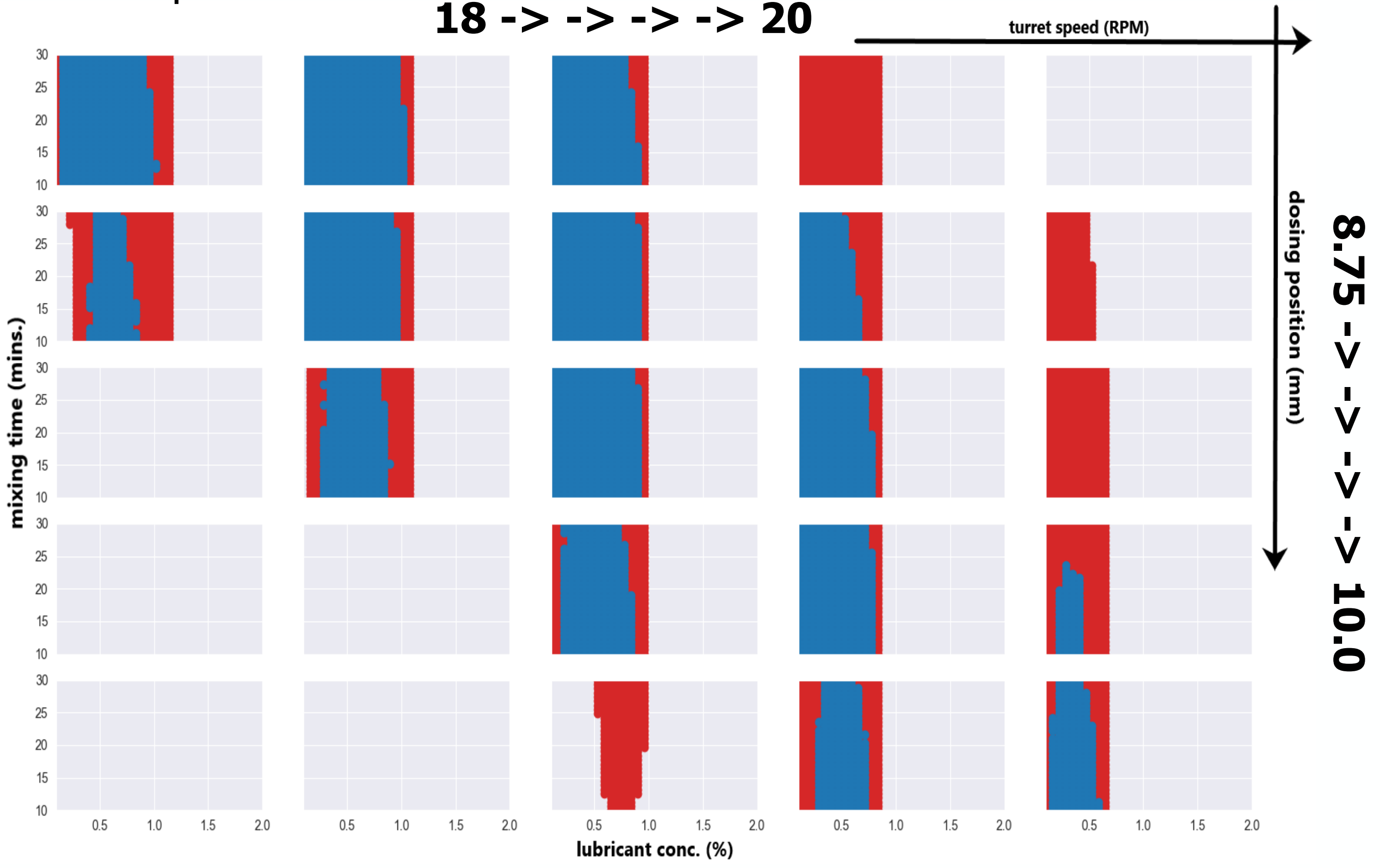


Responses correlation



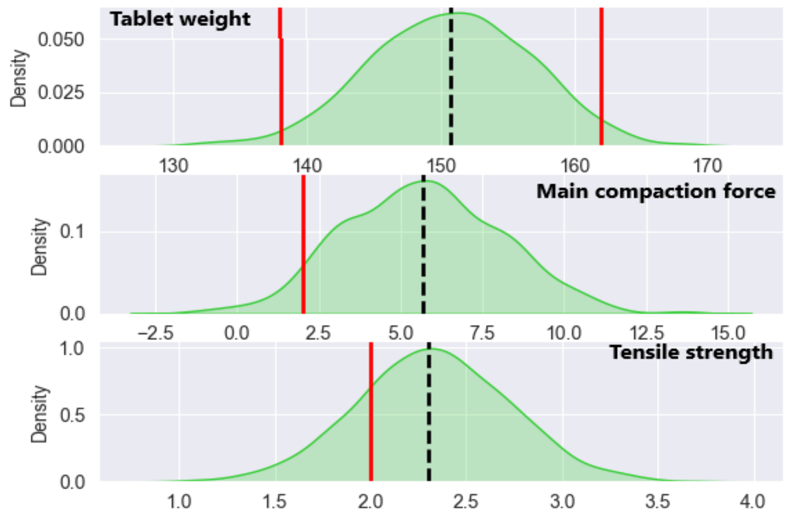
RSM - Design space for lubricant effects

- Nominal DS
- >75% prob. DS
- 50-75% prob. DS



CPP	Feasible region
Dosing position (h_{fill}, mm)	**8.5-9.7**
Turret speed (n_T, rpm)	**18-19.5**
MgSt conc. (c_L, %)	**0.3-0.8**
Mixing time (γ, mins.)	**10-30**

Optimal CPP (h_{fill}, n_T, c_L, γ) = (9.1, 19, 0.4, 11)



A hybrid system

- From the design space regions generated by the hybrid system, the two additives affect the blend properties in a significantly different manner.
- Experimental investigation requires different CPPs that would result in desirable CQAs

Glidants-based tablet DS

- 50 tablets were manufactured for two separate blends (c_L = 0.1 & 0.2%, γ = 20 mins)
- CPP were selected from the extracted DS

Run	CPPs h_{fill}, h_{in-die},	CQAs W, F_{main}, σ_t	Fail prob (%)
1	6.25, 1.7	143, 3.4, 2.2	1.73
2	6.45, 1.75	146, 3.5, 2.1	1.75

Tablet tester

Lubricant-based tablet DS

- 50 tablets were manufactured for two separate blends (c_L = 0.75 & 0.5%, γ = 15 mins)
- CPP were selected from the extracted DS

Dos. pos. (h_{fill})	Turret speed (n_T)	CQAs (W, F_{main}, σ_t)
9.4	18	215, 12.3, 3.3
8.5	18	199, 7.4, 2.2
8.0	18	188, 5.4, 1.67

- SUR model for lubricants can be improvised by building model-based DoE and/or change of TP tooling.
- Nonetheless, the SUR-based model results in meaningful DS with significant overlap with experimental tablet properties.
- This is a significant improvement from the semi-mechanistic TP lubricant model.

Summary

- A hybrid system for design space estimation is described where lubricant and glidant effects are integrated into the tableting process.
- The system can be useful when a mechanistic/data-driven model alone is insufficient to describe the process reliability.
- The mechanistic model can help in the glidants integration whereas RSM approach can be useful for integrating lubricant effects into the continuous tableting process.

Future work

- When do you need hybrid models?
- Improve the quality of data-driven alternatives with SUR model-based DoE.
- Develop computationally efficient Bayesian estimation methods.

PURDUE UNIVERSITY | Davidson School of Chemical Engineering

Thank You

cp3

CryPTSYS

NSF

PURDUE UNIVERSITY | Davidson School of Chemical Engineering