Process Intensification: Application in Pharmaceutical Manufacturing

Marianthi Ierapetritou

Department of Chemical and Biochemical Engineering
Current state of Pharmaceutical Industry

• **Process development challenges**
  – Time pressure for clinical supply delivery
  – Uncertainty throughout development process
  – Unique physical properties of APIs
  – Sequential scale-up of batch processes

• **Economic challenges**
  – Up to 27% of revenues spent on manufacturing costs
  – Increased global competition, generics manufacturers

• **Regulatory concerns**
  – Quality by Design (QbD) – companies need increased process understanding
  – Inherent variability in performance and sampling

*Need efficient and robust manufacturing strategies in order to remain competitive*

Buchholz, S. Chem Eng Process  2010, 49 (10), 993-995
Shah, N. Comput Chem Eng 2004, 28 (6-7), 929-941
Pharmaceutical industry is innovative in development of new drugs **BUT manufacturing is primitive** compared to other chemical industries.

Given a **new** formulation/product:

- Production predominantly in **BATCH** mode
- A batch is produced → samples are tested → batch **FAILS/PASSES**

Sample only a few

Is quality for rest of the tablets the patient will take **captured**?
Batch vs. Continuous Processes

- **Intermediate steps** in batch, not continuous

**Batch Process**

- POWDER → INTERMEDIATES → TABLETS
- Lots of down
- Time

**Continuous Process**

- POWDER → TABLETS
- No intermediate blends or steps

**Continuous Process**

- Washing and Prep
- Comill
- V – Blender
- Intermediate Blend
- Loss-in-Weight Feeders
- Materials fed constantly
- Mixer
- Feed Frame & Tablet Press
Continuous manufacturing has no lag times in production, while batch has delays due to washing, blending and comilling between batches (no product being made)

**Increased productivity** (No. tablets in X time)

- **Set Up**
- **Start Operation**
- **Filling Feed Frame**
- **Washing, down time blending, comilling**
- **Tablet Press**

**Remember:** Product is only manufactured when its at the tablet press
Batch vs. Continuous Processes

- **Batch**
  - **Cons**
    - Productivity is low
      - Down time
    - Set process design and operation
    - Powder exposure during process
    - Scale-up necessary
      - Time and new equipment
    - Harder to control
      - Wasted batches
      - Within-batch variability
    - Multiple operators required
  - **Pros**
    - Many products are produced in smaller quantities
    - Existing ‘know-how’ and fillings

- **Continuous**
  - **Cons**
    - Novel method
      - Few regulatory fillings
    - Requires engineering understanding
  - **Pros**
    - High Productivity
      - No down time in process
    - Set design, but varying parameters
      - Flexible operation
    - Enclosed powders = no exposure
    - Less scale-up studies
      - Extended operation = scale up
    - Better control
      - No failed batches
    - Automated process = less operators
    - Smaller footprint and equipment
Advantages of Continuous Processes

Some of the major advantages of continuous systems include:

– **Increased productivity**
  • Eliminate down time during operation

– **Fewer scale-up studies**
  • Parallelization, increased throughput
  • Reduced time to market

– **Small and compact equipment**
  • Reduced capital cost and utilities requirement
  • Small area footprint

– **Ability to implement control strategies**
  • Real-time feedback control, Model Predictive Control (MPC)
  • Enhanced process robustness

– **Advanced computational tools – process systems engineering**

Process Systems Engineering

- **Process modeling capabilities**
  - Supplement experimental work during process development\(^1\)
  - Design and test control strategies
  - Flowsheet models\(^2\), Discrete element method\(^3\)

- **Process analysis**
  - Sensitivity analysis\(^4\) – identify critical process parameters, control variables
  - Flexibility and feasibility analysis\(^5\) – design space and process robustness

- **Process optimization\(^6\)**
  - Determine optimal process design and operating conditions subject to product quality and process operating constraints

---

Challenges for a flowsheet model for solids

Critical material properties
- Lack of universal set throughout processes and industries
- Inherent variability in powder material properties & distributed parameters
- Lack of technology for monitoring desired material properties online

Critical process operating variables
- Lack of correlation between operating variables and material properties
- No control strategies

Modeling work
- Majority: Discrete Element Method (DEM) simulations → computationally expensive
- In recent literature: plethora of experimental studies & empirical correlations of certain inputs/outputs & specific materials
- Dynamic reduced order models are needed:
  - First-principle based
  - Population balance models
  - Data based models

Werther, J., et al. Computers & Chemical Engineering 23(11-12) 1773-1782
Boukouvala, F. et al. Computers & Chemical Engineering. 42(11) 30-47
Ideal Development of a Pharmaceutical Process

Given a Formulation

**Characterize** Powders using Universal Tests

**Evaluate** Unit Operation Models using Powder Properties

Create the Design Space for Individual Unit Ops

**Design the Process**

**Build** a Flowsheet Model

Develop the Design Space of the Overall Process and **Optimize**

**Determine** Operating Parameters and Control Strategies

**Validate** using Experimental Data

**Asses** Process Sensitivities and Risks
Integrated Process Models

API Feeder → Excipient Feeder → Blender → Granulator → Dryer → Tablet Press → Coater → Dissolution

Optional feeders for more ingredients (i.e. lubricant)

Optional Recirculation tank

Wet Granulation

Continuous flexible multipurpose platform
- Process simulation
- Sensitivity analysis
- Design space evaluation
- Optimization

NSF Engineering Research Center for Structured Organic Particulate Systems (C-SOPS)
Unit Operation Models: Direct Compaction

**FEEDERS:**
- **Model:** Delay Differential Equation
  
  \[ \text{rpm} \quad d_{50} \rho F_{\text{set}} \quad d_{50} \rho F_{\text{out}} \]

**MIXER:**
- **Model:** Population Balance model
  
  \[ \text{rpm} \quad d_{50} \rho F_{\text{in}, C_{i}} \quad d_{50} \rho F_{\text{out}, C_{i}, \text{RSD}} \]

**HOPPER:**
- **Model:** Mass flow buffer tank model
  
  \[ H, D_{\text{outlet}} \quad d_{50} \rho F_{\text{in}, C_{i}, \text{RSD}} \quad d_{50} \rho F_{\text{out}, C_{i}, \text{RSD}} \]

**TABLET PRESS:**
- **Model:** Heckel equation & feed frame residence time model
  
  \[ P, \text{rpm} \quad \epsilon, \sigma, C_{i}, F_{\text{out}} \quad d_{50} \rho F_{\text{in}, C_{i}, \text{RSD}} \]

**DISSOLUTION:**
- **Model:** Fick’s second Law
  
  \[ h, r \quad \epsilon, \sigma, C_{i}, F_{\text{out}} \quad t_{\text{diss}} \]

- Individual unit operation models consist of a series of equations meant to describe process physics and dynamics
- Unit operation equations can be combined sequentially to represent entire manufacturing processes
Latent Variable ROM based on DEM

- In CFD simulation: Discretize into finite elements → solve set of equations for specific element → calculate continuous variable values (T,P).
- In a DEM simulation → discrete elements (particles) → How do we extract information???
  - Discretize geometry and extract **average** information about number of particles inside each compartment. Consider “unreliable means” as **missing**

**BUT how to discretize?**
- ✓ Dense enough to capture spatial variation
- ✓ Coarse enough to have large number of particles inside each element

- Few number of particles: Consider as missing data (impute)
- Very few or no particles: Set equal to zero
- Large enough number of particles: Use average value
Discrete Element Reduced-Order Modeling Methodology

1. DOE – parameter variation
2. Discretize process geometry
3. Extract state data ($Z$)
4. Obtain response data ($Y$)
5. Pre-process state and response data

Input Space

$X \in \mathbb{R}^{N \times m}$

State Space

$Z \in \mathbb{R}^{N \times p}$

Output Space

$Y \in \mathbb{R}^{N \times m}$

6. Reduce dimensionality of state data (PCA)

7. Develop a mapping between input space ($X$) and reduced state space ($PCA scores$)

$X \rightarrow T$

8. Develop a mapping between input parameters ($X$) and output space ($Y$)
Steady State Case Study

Average $u_z$ of particles (m/s)

<table>
<thead>
<tr>
<th>DE-ROM</th>
<th>DEM</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MSE</th>
<th>28%</th>
<th>19%</th>
<th>21%</th>
</tr>
</thead>
</table>

Prediction error for RSD (1 case): 1%

NSF Engineering Research Center for Structured Organic Particulate Systems (C-SOPS)
Dynamic Case Study

Velocity predicted by ROM

Velocity obtained from DEM

Predicted $u_x$ vs. $u_x$ obtained from DEM 23 seconds after change from 160 to 250 rpms

- Velocity and RSD predictions have good accuracy
- Velocity prediction can be used directly in PBM model
- Prediction of RSD can be used for surrogate-based modeling or sensitivity analysis applications

<table>
<thead>
<tr>
<th></th>
<th>$U_x$</th>
<th>$U_y$</th>
<th>$U_z$</th>
<th>RSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>% MSE</td>
<td>0.55%</td>
<td>0.96%</td>
<td>1.13%</td>
<td>1.07%</td>
</tr>
</tbody>
</table>
DESIGN SPACE

How much **uncertainty** can a process tolerate?
"Flexibility of a process is defined as its capability to maintain feasible operation for a range of uncertain conditions that may be encountered during operation" \(^1,^2\)

Vast literature on formulation of optimization problems which to find max acceptable deviations under uncertainty

\[
\begin{align*}
\max / \min \ & P(d, z, x, \theta) \\
\text{s.t.} \ & h(d, z, x, \theta) = 0 \\
\ & g(d, z, x, \theta) \leq 0
\end{align*}
\]

\(^3\) Lepore, J., & Spavins, J. Journal of Pharmaceutical Innovation, 2008

\(^4\) Boukouvala et. al, Journal of Pharmaceutical Innovation, 2010

1 Halemane et al. (1983), AIChE Journal.
2 Floudas et al. (2001), IECR
Black-box Process Feasibility

Goal is to locate boundaries of feasible operation:
- When multiple constraints are present
- Closed form expression for constraints may not be available
- When discrete designs are possible

Feasibility function – process is feasible when $u \leq 0$

$$\psi(d, \theta) = \min_{u,z} u$$

$$s.t. g_j(d, z, \theta) \leq u, \quad j \in J$$

---

**Initial Samples (DOE)** → **Determine Expected Improvement $E[I(x)]$** → **$E[I(x)] < tol?$** → **Yes** → **Feasible region boundaries identified** → **No** → **Additional samples**

$$E[I(x)]_{feas} = s \Phi \left( \frac{0 - y_{pred}}{s} \right)$$

Probability of $u = 0$: boundary

Model uncertainty

---

Initial sampling → Refined sampling → Predicted feasible region

---

1. Jones et al. 1998
2. Boukouvala et al., Computers and Chemical Engineering, (36), 2012

---

NSF Engineering Research Center for Structured Organic Particulate Systems (C-SOPS)
Design Space of Continuous blender

We define feasibility:

- Since we want to minimize the output RSD, we set a maximum threshold value that can be tolerated ($RSD_{\text{max}}$).
  - If the predicted output is lower, it's feasible.
  - If the predicted output is higher, it's infeasible.

Where:

- $i$: design
- $m$: total number of designs
- $n$: number of input variables
- $\beta_j$: RSM coefficients
- $z_j$: input variables
- $x_i$: response surface produced for each design

Introduce binary variables for each design ($m$) and form a MINLP problem:

$$\begin{align*}
\text{min/max} & \quad \sum_{i=1}^{m} y_i x_i \\
\text{s.t.} & \quad \sum_{i=1}^{m} y_i = 1 \\
& \quad y_i \in \{0,1\}^m \\
& \quad z_j^{(i)\text{lo}} \leq z_j^{(i)} \leq z_j^{(i)\text{up}} \quad i = 1...m, \; j = 1...k
\end{align*}$$

Where:

- $i$: design
- $m$: total number of designs
- $n$: number of input variables
- $\beta_j$: RSM coefficients
- $z_j$: input variables
- $x_i$: response surface produced for each design

Constraint to make sure only one design is chosen:

$$x_i = \beta_0 + \sum_{j}^{n} \beta_j z_j + \sum_{j<k}^{n} \beta_{jk} z_j z_k + \sum_{j}^{n} \beta_{jj} z_j^2$$


NSF Engineering Research Center for Structured Organic Particulate Systems (C-SOPS)
Inverse problem:
Based on desired properties, what should the design of the flowsheet be?
Surrogate Based Optimization: Proposed Methodology

- Combination of global search (initially) with local search (final stage)
- Incorporation of a black-box feasibility stage to identify form of feasible region
- Final local trust-region approach by allowing multiple starting points if clusters of promising feasible points is identified
- Alleviation of noise effects through a stochastic kriging model
  - heteroscedastic variance case

Initially sample entire region

Refine in boundary regions to map feasible region

\[
E[I(x)]_{feas} = s \phi \left( \frac{0 - y_{pred}}{s} \right)
\]

Process Optimization

- **OBJECTIVE:** minimize cost of a 1 day operation of continuous direct compaction
- **DECISION VARIABLES:** process capacities, operating conditions, throughput, refill strategy
- **SUBJECT TO:** Process capacity bound constraints, Product quality constraints, Minimum production requirement
- Leads to an optimization of an expensive-to-evaluate model, with complex constraints and uncertainty: **SURROGATE SIMULATION-BASED OPTIMIZATION**

**Step 1:** Formulate objective and constraints

\[

c_{	ext{equipment}} + c_{\text{operation}} + c_{\text{mixture}}
\]

\[
F_{\text{total}} \leq F_{\text{total}} \leq F_{\text{total}}^{\text{up}}
\]

\[
\text{rpm}_{\text{mixture}} \leq \text{rpm}_{\text{mixture}} \leq \text{rpm}_{\text{mixture}}^{\text{up}}
\]

\[
0.99 \times 0.0092 \leq C_{\text{MgSt}} \leq 1.01 \times 0.0092
\]

\[
0.1 \leq RL \leq 0.6
\]

\[
t_{\text{hopper}} \leq t_{\text{hopper}} \leq t^{\text{up}}_{\text{hopper}}
\]

\[
t_{\text{mixture}} \leq t_{\text{mixture}} \leq t^{\text{up}}_{\text{mixture}}
\]

\[
P_{\text{total}} \leq P_{\text{total}} \leq P^{\text{up}}_{\text{total}}
\]

\[
\text{hardness} \leq \text{hardness} \leq \text{hardness}^{\text{up}}
\]

\[
e^{\text{lo}} \leq e \leq e^{\text{up}}
\]

\[
t_{\text{diss}} \leq t_{\text{diss}} \leq t^{\text{up}}_{\text{diss}}
\]

\[
\text{Tablet prod min} \geq \text{Tablet prod up}
\]

**Step 2:** Flowsheet simulations for different conditions based on DOE

**Step 3:** Build surrogate model and optimize. Approximate uncertainty

**Optimal cost:** $153,892

<table>
<thead>
<tr>
<th>Variable</th>
<th>Optimum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{\text{total}}$ (kg/h)</td>
<td>54</td>
</tr>
<tr>
<td>Mixer rpm</td>
<td>102</td>
</tr>
<tr>
<td>$C_{\text{MgSt}}$ (w/w)</td>
<td>0.0092</td>
</tr>
<tr>
<td>RL(%)</td>
<td>54</td>
</tr>
<tr>
<td>$V_{\text{hopper}}$ (m³)</td>
<td>0.03</td>
</tr>
<tr>
<td>$P_{\text{comp}}$ (Pa)</td>
<td>1042</td>
</tr>
</tbody>
</table>

Adaptive sampling

Output Space

Response surface

Uncertainty surface

Step 1: Formulate objective and constraints

Step 2: Flowsheet simulations for different conditions based on DOE

Step 3: Build surrogate model and optimize. Approximate uncertainty

NSF Engineering Research Center for Structured Organic Particulates
Conclusions and future goals

• As the industry is moving to advanced manufacturing solutions, process intensification will be in the center of attention.

• There is a need for predictive models for optimization of process design and operations

• Reduced order modeling techniques are needed, due to the complexity of models necessary for complex pharmaceutical processes

• Technologies are transferrable to other powder processing industries such as food, consumer goods.

• As flowsheet models are being used, flowsheet synthesis framework will be developed to design process for any new formulation
**Motivation:** Exhausting petroleum resources have prompted the development of sustainable biorefinery to produce biofuel and bio-chemicals from biomass feedstocks.

**Objectives:**
- Perform techno-economic analysis on the productions of biobased chemicals and estimate the minimum cost of the products
- Apply life cycle assessment to evaluate the environmental impacts
- Implement process synthesis and optimization to achieve an optimal process diagram

**Accomplishments:**

### Techno-Economic Analysis
- Experimental Data
- Process Design
- Discounted Cash Flow

### Life Cycle Assessment
- Aspen plus simulation
- Inventory
- LCA software
- Energy flow
- HEN
- Aspen Energy Analyzer

### Process Synthesis and Optimization
- Biomass feedstock
- Catalytic conversion
- Bioprocess
- Separation units
- Bioproducts
- Black-Box optimization and synthesis
Acknowledgements:

Funding provided by the ERC (NSF-0504497, NSF-ECC 0540855)