On-Line Estimation of Diastereomeric Resolution with :
FBRM, Raman Spectroscopy, & ATR-FTIR

Sze Wing Wong

Advisor: **Christos Georgakis**
Gregory D. Botsaris

System Research Institute &
Chemical and Biological Engineering Dept,
Tufts University

Outline

- **Motivation & Objective**
- Background Information
  - System Background
  - Experimental Setup
- Estimation Model
  - Combination of On-Line Measurements
- On-Line Monitoring of Key Parameters in Diastereomeric Resolution

Motivation – Improve Previous Model

- Raman Capable of Quantifying % Diastereomer
- Slurry Density Improves Prediction Accuracy*
  - Model from Prior Data Set
  - 4 Different Models
  - Requires Sampling Loop or Weight Analysis

$$\left(\frac{133.93-19.62}{19.62}\right)\times 100\% = 583\% \qquad \left(\frac{10.67-9.99}{9.99}\right)\times 100\% = 7\%$$

RMS of % Error of PLS Models from TG-LD or TG-HD

	Model w/ S	Model w/ T, S	Model w/ D, S	Model w/ T, D, S
Exp 1 (LD)	22.79	13.01	13.14	8.99
Exp 2 (HD)	11.37	12.42	13.23	12.99
Exp 3 (HD)	10.49	10.67	10.53	9.99
Exp 4 (LD)	133.93	122.68	20.67	19.62

* Wong et. al. (2006). AIChE, San Francisco, CA, Section #668c

Objective

- **Alternative On-Line Measurement(s) to Replace Slurry Density**
  - FBRM – Particle counter
  - IR – Infer Slurry Density Information from Solute Concentration
- **How much Information Can be Extracted from Raw Raman, IR, and FBRM data?**

Outline

- Motivation & Objective
- **Background Information**
  - System Background
  - Experimental Setup
- Estimation Model
  - Combination of On-Line Measurements
- On-Line Monitoring of Key Parameters in Diastereomeric Resolution

The Chemical System

(R,S)-A + ***D*** ⟶ ***S-D*** + ***R-D***

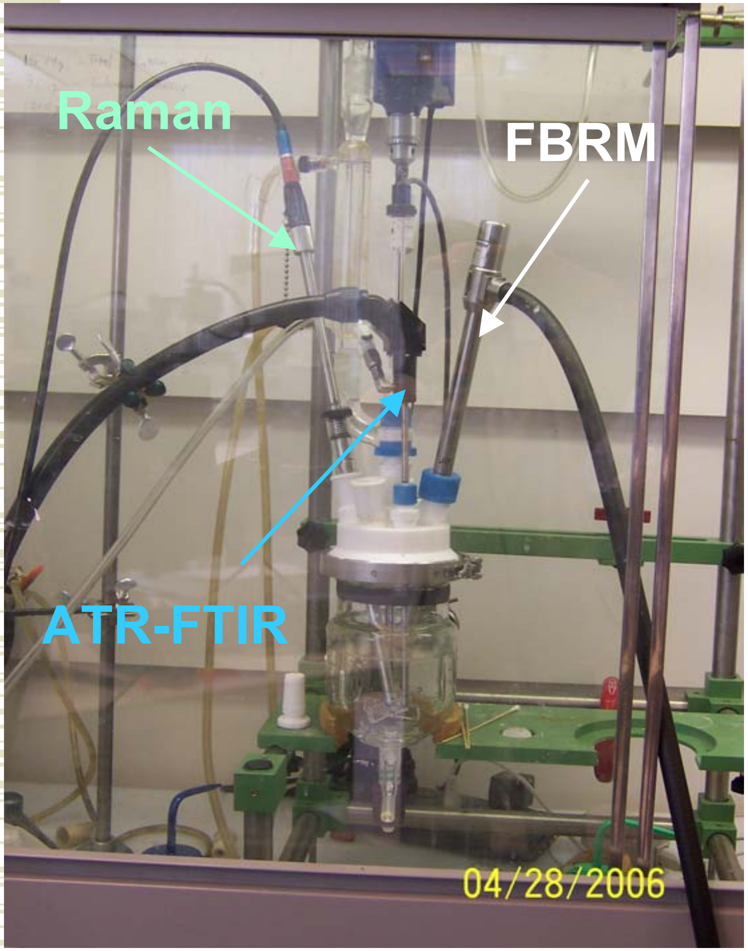
(R,S)-A → API

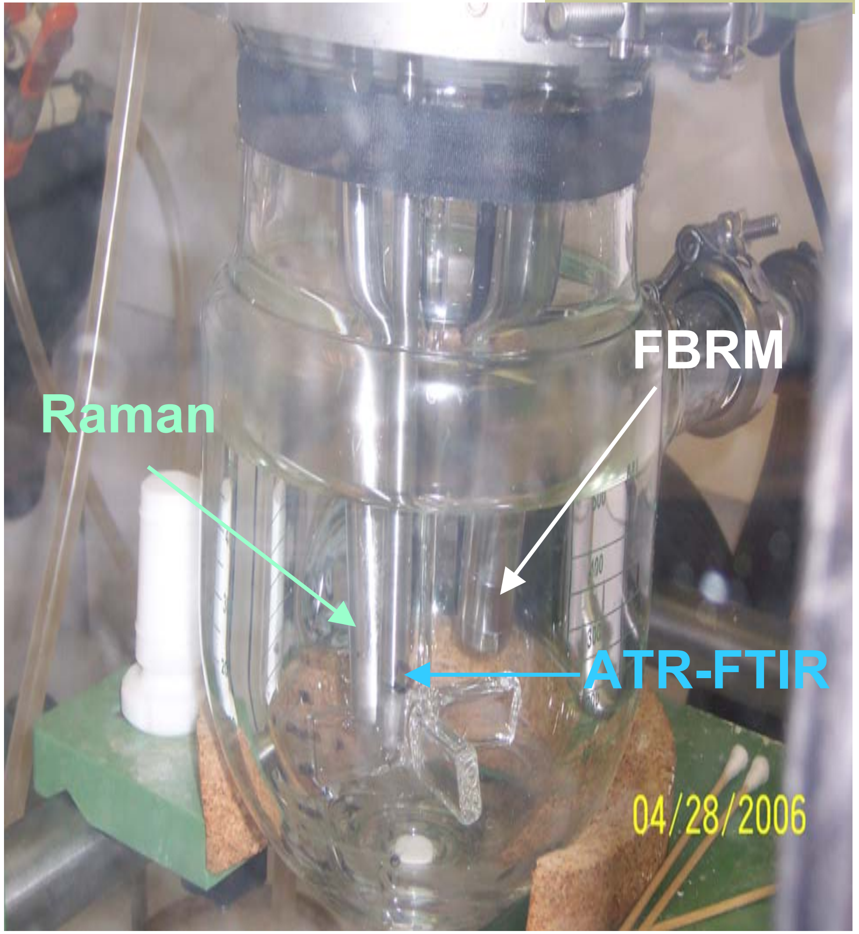
D → Chiral Acid

S-D → Separation Product

- Single Enantiomer as Final Drug Form (***Sepracor*** Compound)
- Diastereomers Resolution to Separate Desired Enantiomer
  - Diastereomer Have Different Physical Properties
  - Cooling Crystallization to Separate Desired Diastereomer

Experimental Setup





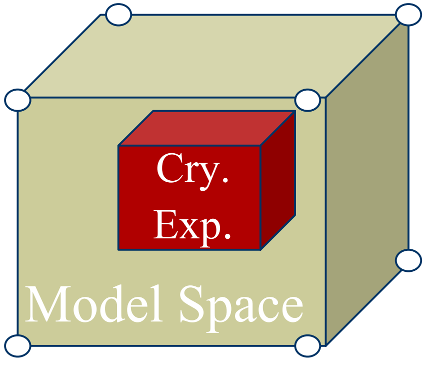
Outline

- Motivation & Objective
- Background Information
  - System Background
  - Experimental Setup
- **Estimation Model**
  - Combination of On-Line Measurements
- On-Line Monitoring of Key Parameters in Diastereomeric Resolution

Experimental Condition

- DOE Varies:
  - Cooling Rate, Seeding Amount, Agitation Rate
- Analyze Samples with HPLC
  - Obtain Charility & Solute Concentration
- PLS Model
  - 84 Samples – Build Model
  - 35 Samples – Test Model

Experiment #	Seed (wt %)	Cooling Rate (^{0}C/min)	Agitation
1	2	0.1	550
2	4	0.1	350
3	2	1	350
4	**4**	**1**	**550**
5	3	0.5	450
6	**4**	**0.5**	**300**
7	2	Nonlinear Cooling Profile 1	350
8	2	Nonlinear Cooling Profile 2	250
9	**6**	**Nonlinear Cooling Profile 2**	**350**
10	6	Nonlinear Cooling Profile 1	250



Data Fusion

- Normalized Different Measurements Inside the Data Matrix
  - Autoscale Temperature, Slurry Density, and Percent Composition of S-D diastereomer

$$\hat{T}_i = \frac{T_i - T_{avg}}{\sigma_T} \qquad \hat{D}_i = \frac{D_i - D_{avg}}{\sigma_D} \qquad \hat{P}_i = \frac{P_i - P_{avg}}{\sigma_P}$$

  - Normalize Raman Spectra with Under Curve Area
    - *Raman Spectra – Take 1st Derivative for Baseline Correction*

$$F_i = \int f_i(w)dw \qquad F_{avg} = \frac{\sum_{i=1}^{n} F_i}{n} \qquad \tilde{S}_i = \frac{f_i(w) - F_{avg}}{\sigma_F}$$

Combination of On-Line Measurements

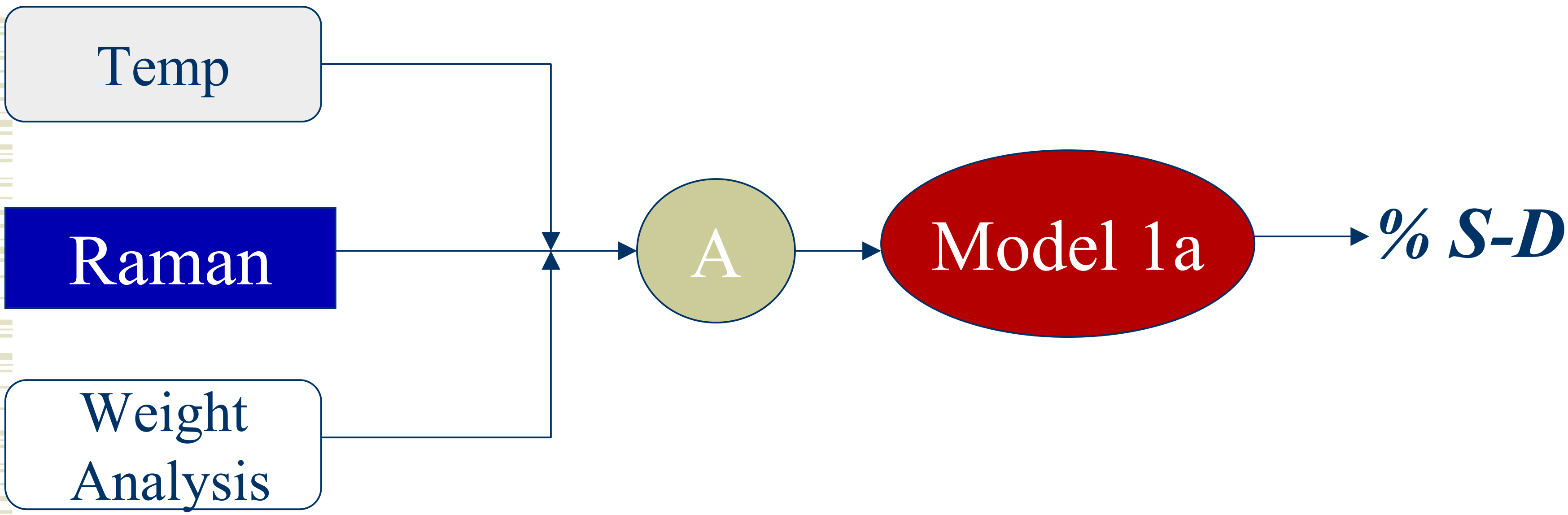
- Data Matrix = [Temp, **Density**, & Raman]
  1. ***Data Set A***: Density Obtained from Weight Analysis
  2. ***Data Set B:*** Density Obtained **ATR-FTIR**
  3. ***Data Set C:*** Density Obtained from **FBRM** & **Raman**
- Model Comparison
  - With Relative % Error, RMS % E & RMSE

$$\%E = \left| \frac{\left(y_{i\,estimate} - y_{i\,\mathrm{experiment}}\right)}{y_{i\,\mathrm{experiment}}} \right| \times 100\%$$

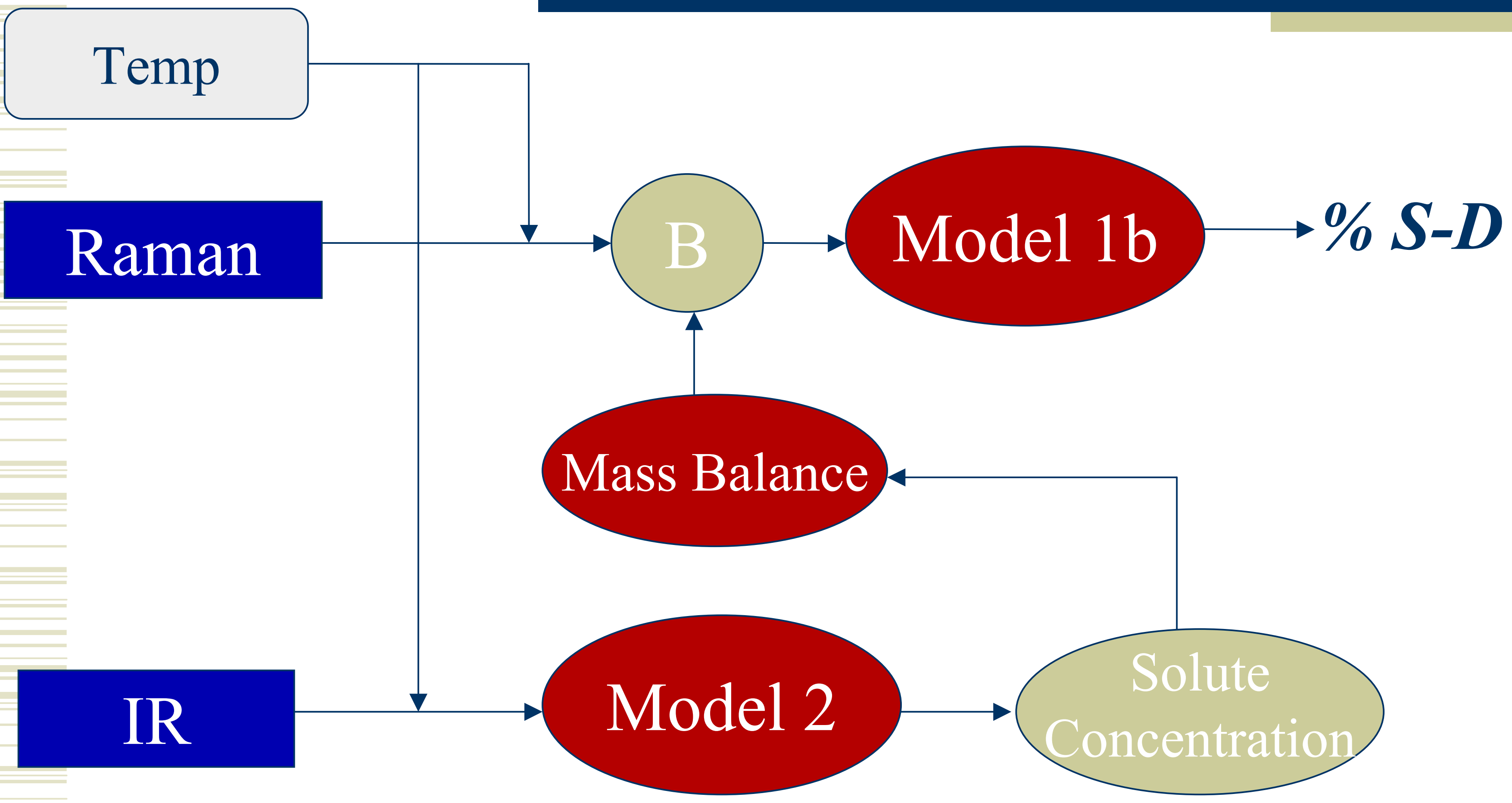
$$RMS\ of\ \%E = \sqrt{\frac{\sum_{i=1}^{n} (\%Error)^2}{n}}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(y_{i\,estimate} - y_{i\,experiment}\right)^2}{n}}$$

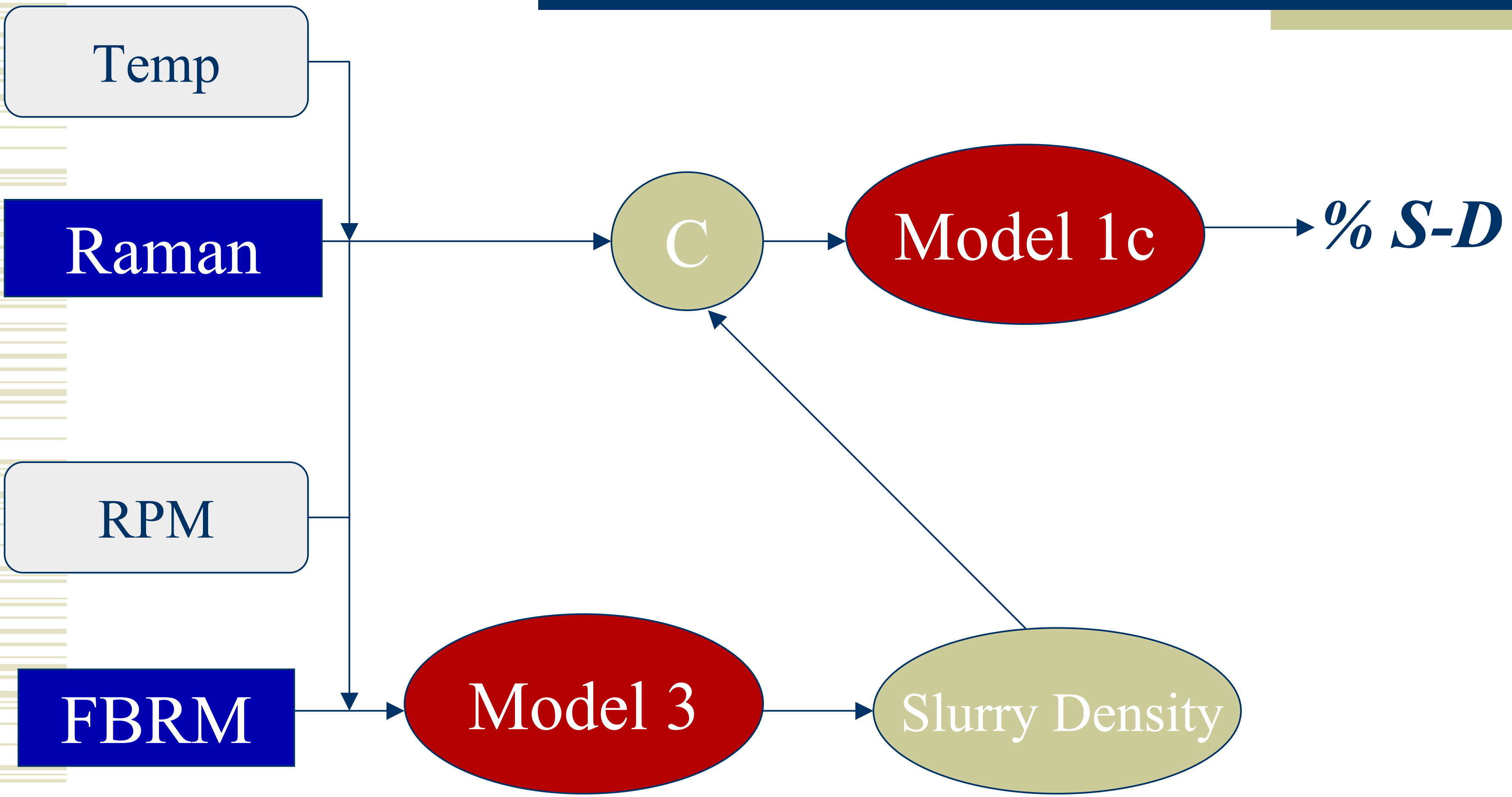
Data Set A



Data Set B

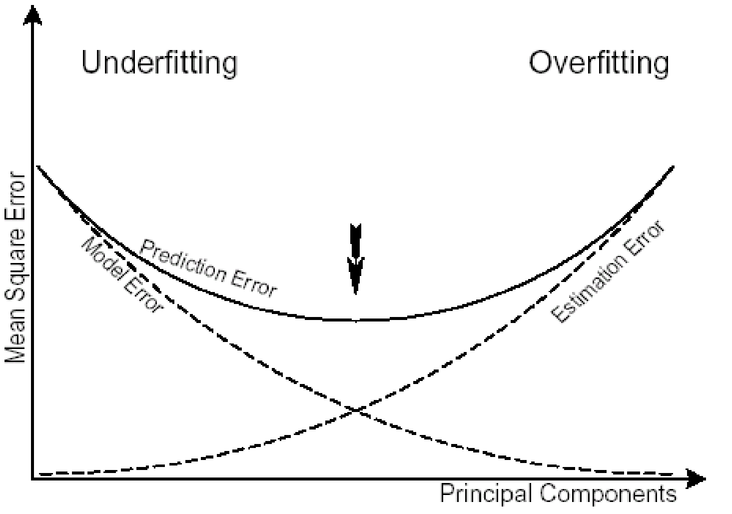

Temp
Raman
IR
B
Model 1b
% S-D
Mass Balance
Model 2
Solute Concentration

Data Set C



PLS Estimation Model

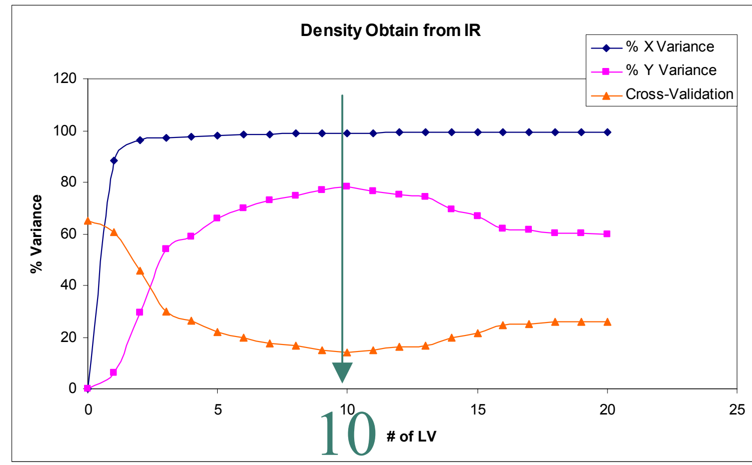
- Number of Optimal Latent Variables
  - Cross-Validation
    - Iterate Combinations of Subgroups to Build & Test Models
  - % X & Y Variance
    - % Variance Covered by X & Y Variables Levels Off
  - Testing Group

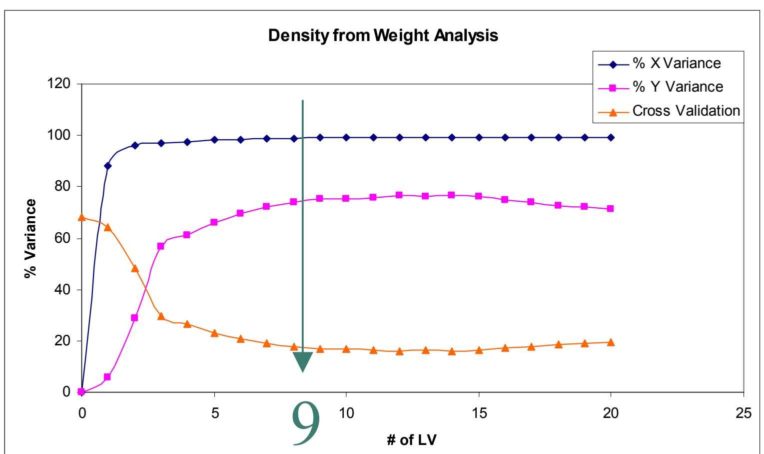


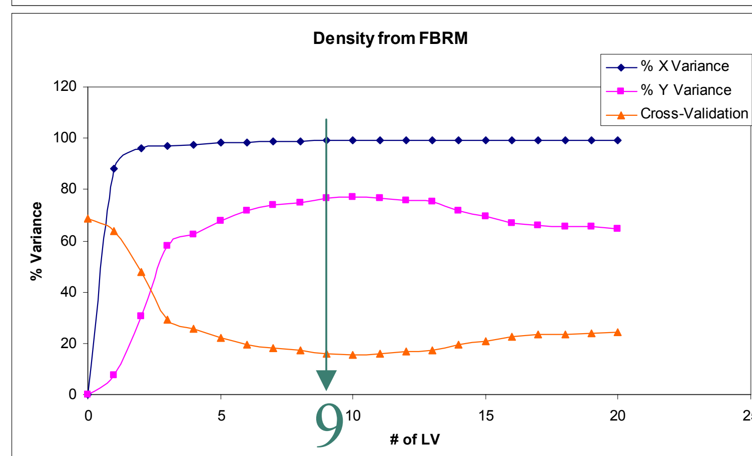
Reference: Multivariate Data Analysis by Esbensen

Selection of Optimal # of LV

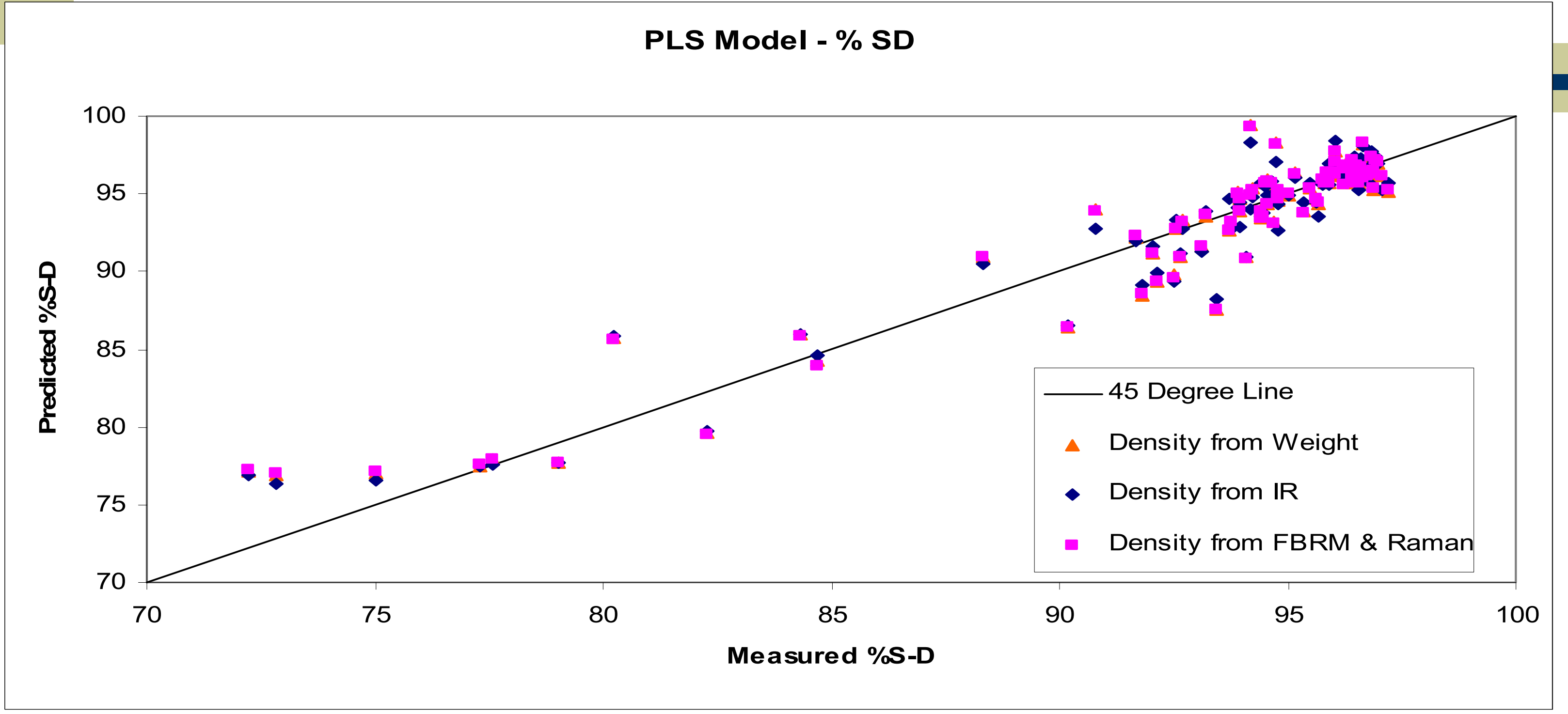
- Determine # of LV
  - Data Set A: 9
  - Data Set B: 10
  - Data Set C: 9







PLS Model Detail



- Good Agreement Between Models
  - Data Points Overlapped
- Good Correlation Between Data & % S-D
  - $R^2 = 0.9$

Result – **Dynamic Exp**

- Prediction Error from Testing Group
  - Total of 35 data points
    - From Different Experiments
- Model with Density Obtained from IR Gave Best Estimation

Model	# of LV	% X Variance	% Y Variance	Max % E	RMSE
Density from Weight Analysis	9	98.99	75.2	11.2%	4.84
Density from IR	**10**	**99.05**	**78.35**	**10.3%**	**4.79**
Density from FBRM	9	98.97	76.55	11.7%	4.8

Outline

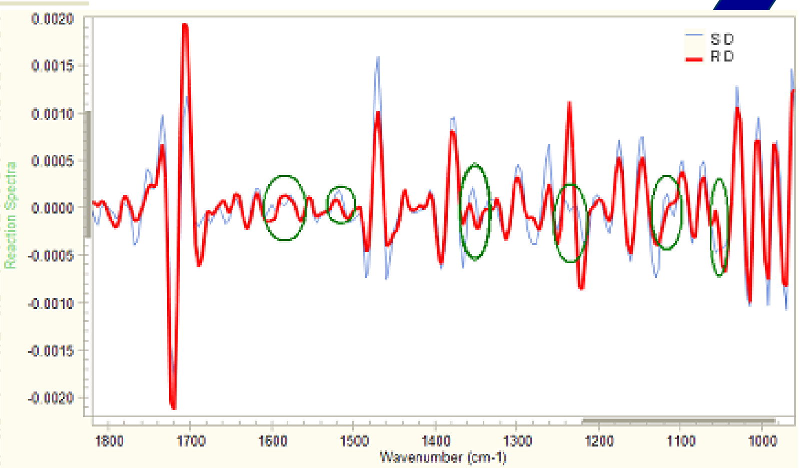
- Motivation & Objective
- Background Information
  - System Background
  - Experimental Setup
- Estimation Model
  - Combination of On-Line Measurements
- **On-Line Monitoring of Key Parameters in Diastereomeric Resolution**

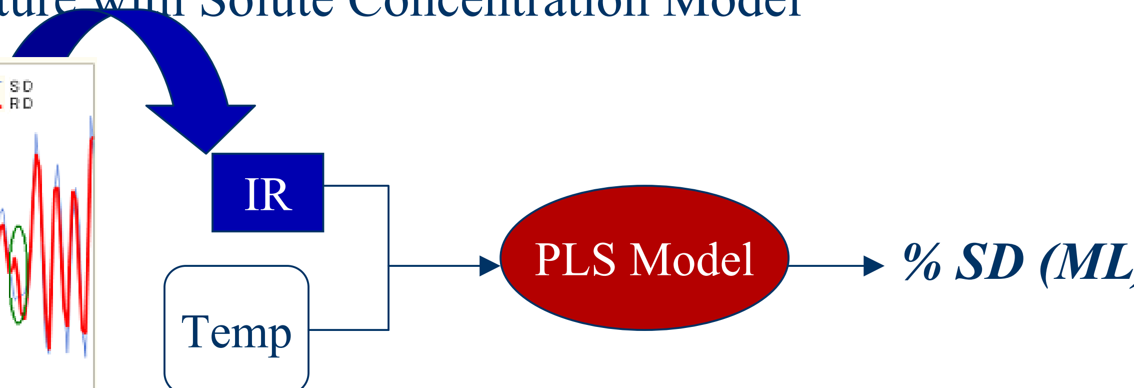
Key Parameters for…

- Nucleation & Growth Kinetic Estimation
  - Converted Information Thus Far
    1. % S-D (Crystals) ← Model with **Temp, IR, & Raman**
    2. Slurry Density ← Model with **FBRM & Raman**
    3. Solute Concentration of A ← Model with **Temp & IR**
  - Unknown Parameter
    1. Solute Concentration of S-D & R-D
    2. Particle Size

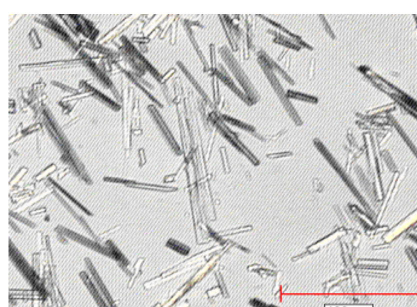
Obtaining Unknown Parameters

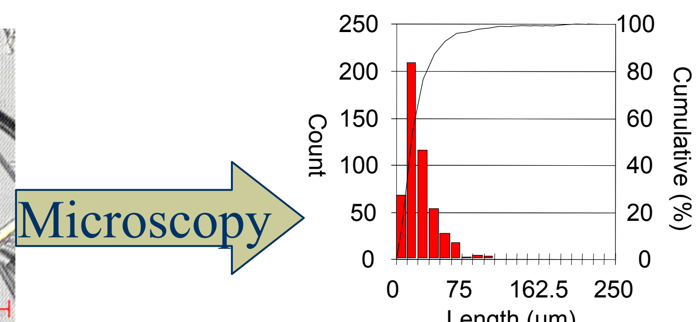
- IR Differentiates SD & RD in Solution
  - Same Model Structure with Solute Concentration Model



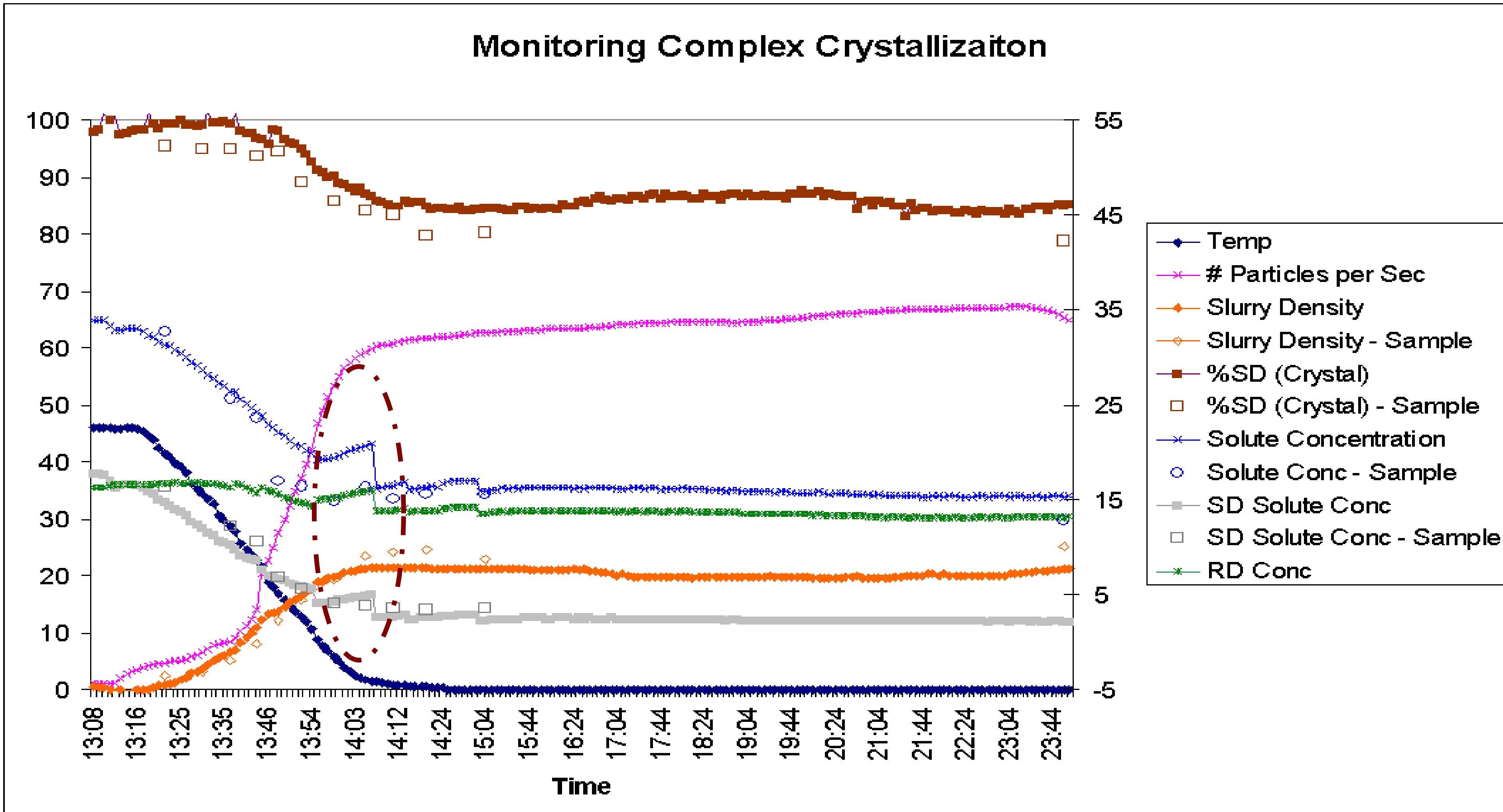


- Particle Size





Experimental Condition: 4wt% Seed, 1^0C/min Cooling, & 550 RPM



Conclusion

- Improve Previous Model
  - No Sampling Loop Require
- Convert Spectral Data into Information
  - Need Combination of Measurements
- Continuous Process Monitoring
  - Utilize the Acquired Information for Data Driven Modeling

Future Challenge:

- Modeling Complex Crystallization
  - Nucleation Kinetic
  - Growth Kinetic
- Optimization & Control

	Experiment #	Seed (wt %)	Cooling Rate (^{0}C/min)	Agitation
Nucleation	1	2	0.1	550
	2	4	0.1	350
	3	2	1	350
	4	4	1	550
	5	3	0.5	450
	6	4	0.5	300
Growth	7	2	Nonlinear Cooling Profile 1	350
	8	2	Nonlinear Cooling Profile 2	250
	9	6	Nonlinear Cooling Profile 2	350
	10	6	Nonlinear Cooling Profile 1	250

Acknowledgement

- **Sepracor:**
  - Kostas Saranteas
  - Roger Bakale
  - Patrick Mousaw
  - Mike Sizensky
- **Camo Inc (The Unscrambler)**
  - DongSheng Bu
- Financial Support From Sepracor Inc.