

© 2022

Pooja Rajendra Bhalode

ALL RIGHTS RESERVED

MULTI-SCALE MODELING AND ANALYSIS FOR
PHARMACEUTICAL MANUFACTURING

BY

POOJA RAJENDRA BHALODE

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Chemical and Biochemical Engineering

Written under the direction of

Marianthi G. Ierapetritou and Fernando J. Muzzio

And approved by

New Brunswick, New Jersey

MAY, 2022

ABSTRACT OF THE DISSERTATION

MUTLI-SCALE MODELING AND ANALYSIS OF PHARMACEUTICAL MANUFACTURING

by POOJA RAJENDRA BHALODE

Dissertation Directors:

Marianthi G. Ierapetritou and Fernando J. Muzzio

With the recent technological advancements in digitization and Industry 4.0 practices, the pharmaceutical industry is moving towards integrating these advancements within the drug manufacturing. This integration is termed as 'Pharma 4.0'. A crucial component of this integration involves the development of digital twins of the manufacturing lines. Regulatory agencies are collaborating with academic and industrial partners to develop digital twins and move towards Pharma 4.0, with the aim to develop robust, flexible, and agile manufacturing lines. However, despite these efforts, there are some crucial aspects that need to be addressed to develop accurate and integrated digital twins. These include and are not limited to the accuracy of model prediction, computational time of process simulation, and handling and pre-treatment of experimental data from the manufacturing line. The work presented in this thesis focuses on these aspects, with the research goal being to develop computationally efficient multi-scale process models that incorporate detailed particulate dynamics for obtaining accurate process prediction along with statistical analysis and data pretreatment of experimentally obtained datasets. The research goal is aimed towards developing digital twin of the manufacturing lines and the respective objectives are outlined in chapter 2.

The first aim of this thesis focuses on developing high-fidelity particulate simulation of pharmaceutical unit operations to understand detailed particle-level mechanics. This is demonstrated for a powder feeder unit, being located at the top of the manufacturing lines and a key contributor to the observed variability in powder flow. Chapter 3 presents the simulation of powder feeder using discrete element modeling (DEM), along with calibration of DEM parameters.

The second aim of the thesis is outlined in chapters 4 and 5, where the detailed particulate information is used to develop computationally efficient process models for digital twin applications, to improve the overall prediction of manufacturing lines. This is demonstrated for a continuous powder blender. This aim also includes a detailed review on mixing indices which are important to quantify the levels of mixing or blend uniformity within the blender system.

Lastly, the third aim of the thesis focuses on the physical component of digital twin involving data pre-treatment and analysis. This aim includes two aspects, outlined in chapter 6. First aspect involves the statistical data handling and denoising strategies for experimental residence time distribution (RTD) of manufacturing lines, given the high degree of variability in the datasets, which can significantly affect the process understanding and lead to erroneous conclusions. The second aspect of this aim focuses on optimizing the efforts required to obtain RTDs experimentally. This is investigated by identifying the important features of RTD from a quality assurance perspective, followed by directing efforts to capture them accurately. The proposed approach thus optimizes the required experimental efforts while ensuring regulation compliant manufacturing of drug products.

ACKNOWLEDGEMENTS

PhD has been an incredible journey that has transformed my life in various aspects, providing the opportunities to grow at a professional, personal, and individual level. I have had great luck in meeting the most amazing individuals over the past 5 years. The guidance and company of these individuals have made me the person I am today. It is truly said that the company you surround yourself with is the company you become, and I have been blessed to have had such an incredible company. I would like to take this opportunity to express my sincere and heartfelt gratitude towards them.

First and foremost, I would like to express my gratitude for my PhD advisors - Prof. Marianthi G. Ierapetritou and Prof. Fernando J. Muzzio, for not only being my advisors but also being an unwavering support personally and professionally. I have been incredibly lucky to have been advised by such dynamic individuals. Your passion, perseverance, and innate drive to work hard has and will always motivate and inspire me to keep moving forward in life and be the best version of myself. I would like to express my deepest and sincere thanks for giving me the opportunity to be a part of your research group and mentoring me along all these years. Your encouragement and trust provided me with the motivation and drive to continue working hard. It also provides me with a sense of confidence within myself and my abilities to not only be a good researcher as well as an all-round individual. I would really like to thank you for providing me with a safe space to share my deepest feelings, to feel a part of the research family, and grow as an individual. I cannot thank you both enough for all the guidance and support you have provided me along the years. I would also like to thank my thesis committee advisors – Prof. Rohit Ramachandran and Dr. Amanda Rogers for their thorough and valuable feedback on my PhD projects. I really appreciate your guidance on these projects. It has truly been a privilege to have had a chance to meet and work with

you. I would also like to thank our funding sources – U. S. Food and Drug Administration and National Science Foundation.

I would also like to extend my deepest gratitude to all the group mates that I had the pleasure to learn and work alongside with at Rutgers and University of Delaware. I would like to thank all the lab-mates from Ierapetritou and Muzzio research groups (Sebastian, Zilong, Nirupa, Atharv, Ou, Yi Tao, James, Yingjie, Huayu, Tianyi, Sonia, Thamer, Golshid, Gerardo, Ravendra, Shashwat, Jingzhe, Quishi, Chaoying, Jayanth, Ching-Mei, Robert, Dare, Katerina, and Borja) for making the group such a friendly place. Thank you, Sebastian, Nirupa, Ou, Atharv, for your valuable help during my PhD. We all have developed such fond memories over the years of learning together. I would also like to thank the alumni members of the groups (Fani, Atul, Amanda, Yifan) for guiding me, along with our Graduate Program Directors – Prof. Charles Roth and Prof. Fuat Celik for directing me during my PhD.

I would like to thank my advisors from Carnegie Mellon University – Prof. Lorenz Biegler and Dr. Mark Daichendt for inspiring me, helping me understand the basics of mathematical modeling and process systems engineering while providing me a safe place to grow. Thank you Dr. Daichendt for guiding me to pursue a PhD, your support has indeed motivated me take the plunge.

Lastly, I would like to express my deepest gratitude to my parents, my siblings, my husband, and my in-laws for the unconditional support, inspiration, love, and care that you all have showered me with all these years. I also feel very fortunate to have met and be mentored by Birjoo and Latika, in the company of whom, I have and will keep learning immeasurably. It has been an incredible 5 years in becoming another doctor (although not a medical professional) in the family tree, with the opportunity to get married to the person who has been an unwavering support and motivation in my life. This journey would not have been the same without the unconditional love, encouragement, and support from all of you.

TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION	ii
ACKNOWLEDGEMENTS	iv
TABLE OF FIGURES	ix
TABLE OF TABLES	xii
1. Introduction.....	1
1.1. Pharmaceutical industry	1
1.2. Pharmaceutical manufacturing.....	2
1.3. Progress of pharmaceutical manufacturing in the past decades.....	3
1.4. Manufacturing routes in CM.....	5
1.5. Current trends in pharmaceutical manufacturing.....	8
1.6. Background of process modeling in pharmaceutical manufacturing.....	11
1.6.1. High-fidelity mechanistic models.....	11
1.6.2. Multi-dimensional mechanistic models (PBM).....	12
1.6.3. Phenomenological or semi-empirical models.....	13
1.6.4. Empirical or data-driven surrogate models.....	13
1.6.5. Hybrid models.....	14
1.6.6. Flowsheet models	14
1.7. Current challenges in the pharmaceutical manufacturing.....	16
1.7.1. Capturing mechanistic insight in the process models to improve process predictability	16
1.7.2. Developing computationally efficient process models.....	17
1.7.3. Statistical analyses for systemic information in manufacturing lines.....	18
2. Research Objectives.....	21
2.1. Overall research goal.....	21
2.2. Specific research aims.....	21
2.2.1. Aim 1: Developing high-fidelity models for obtaining particle-level knowledge for pharmaceutical unit operations	21
2.2.2. Aim 2: Process modeling of unit operations using DEM	22
2.2.3. Aim 3: Data pre-treatment and analysis of manufacturing lines	23
3. Aim 1: Developing high-fidelity DEM simulation for obtaining particle-level knowledge with calibration for pharmaceutical unit operations	25
3.1. Introduction	25
3.2. Materials and methods	31
3.2.1. Discrete element modeling (DEM).....	31
3.2.2. Material bulk calibration using Drawdown test.....	33
3.2.3. DEM simulation of powder feeder	35

3.2.4.	Design of experiments (DOE)	37
3.3.	Results	41
3.3.1.	Material calibration.....	41
3.3.1.1.	Results for DOE of the Drawdown test.....	41
3.3.1.2.	Calibration based on the bulk measurements	43
3.3.2.	Simulation of feeder unit	45
3.3.2.1.	Simulation results	45
3.3.2.2.	Feeder refill studies	47
3.3.2.3.	Results for DOE of feeder simulation	52
3.3.2.4.	Developing correlations based on the DOE	53
3.4.	Discussion	56
3.4.1.	Material calibration results	56
3.4.2.	Verification studies for feeder simulation	58
3.5.	Conclusions	61
4.	Aim 2: Process modeling of unit operations using DEM – identifying mixing indices for continuous blending operations.....	64
4.1.	Introduction	64
4.2.	DEM simulation of blender unit	68
4.3.	Indices for mixing evaluation.....	70
4.3.1.	Review of mixing indices	70
4.3.2.	Classification of mixing indices	81
4.3.3.	Detailed comparison of mixing indices	83
4.4.	Results and discussion – DEM simulation and mixing evaluation	87
4.4.1.	DEM simulation and data extraction	87
4.4.2.	Evaluation of mixing indices	89
4.4.3.	Comparison of indices based on the T95 mixing times	91
4.4.4.	Selection of appropriate mixing index.....	94
4.5.	Summary and conclusions.....	96
5.	Aim 2: Unit operation process modeling using DEM – compartment modeling	99
5.1.	Introduction	99
5.2.	Materials and methods	103
5.2.1.	Discrete element modeling (DEM).....	103
5.2.2.	Multi-zonal compartment modeling	106
5.2.3.	Method for development of compartment model	109
5.3.	Results and discussions	112
5.3.1.	Results for radial compartmentalization	113
5.3.2.	Results for axial compartmentalization	115
5.3.3.	Validation of compartmentalization methodology	117
5.3.4.	Applications of developed compartment model	118
5.4.	Conclusions and future work	122
6.	Aim 3: Data pre-treatment and analysis of manufacturing lines	124
6.1.	Statistical data pre-treatment and denoising for residence time distribution profiles	124

6.1.1.	Introduction	124
6.1.2.	Experimental studies	130
6.1.3.	Methods.....	133
6.1.3.1.	Signal processing – signal and noise	133
6.1.3.2.	Noise characterization for RTD studies	134
6.1.3.3.	Denoising methods	136
6.1.3.4.	Comparison of denoising methods	144
6.1.4.	Results and discussions	147
6.1.4.1.	Denoising methods on noisy RTD datasets.....	147
6.1.4.2.	Comparison of denoising methods based on metric evaluation	148
6.1.4.3.	Impact of denoising on applications of RTD	152
6.1.5.	Conclusions	156
6.2.	Analysis of RTD using truncation.....	157
6.2.1.	Introduction	157
6.2.2.	Methods.....	162
6.2.2.1.	Experimental details	162
6.2.2.2.	Predictive identification of OOS product.....	163
6.2.2.3.	Evaluation of important features of RTD using truncation	165
6.2.3.	Results	168
6.2.4.	Discussions.....	171
6.2.5.	Conclusions	178
7.	Conclusions and future work.....	180
7.1.	Conclusions	180
7.2.	Future work	182
8.	Appendices.....	185
8.1.	Aim 1 – Developing high-fidelity DEM simulations for pharmaceutical unit operations with DEM calibration.....	185
8.1.1.	Calibration results for bulk measurements	185
8.1.2.	Evaluation of feed factor.....	187
8.2.	Aim 2: Unit operation process modeling using DEM – mixing indices	192
8.3.	Aim 2: Unit operation process modeling using DEM – compartment modeling	195
8.3.1.	Details of DEM simulation	195
8.3.2.	Radial mixing profiles using RSD	196
9.	References.....	197

TABLE OF FIGURES

Figure 1.1: Different routes of continuous pharmaceutical manufacturing - direct compaction, wet granulation, spray drying, extrusion and roller compaction or dry granulation (Modified from Escotet-Espinoza et al.[44])	8
Figure 1.1: Proposed framework for calibration and application of DEM for powder feeder	30
Figure 1.2: Drawdown test for material calibration	35
Figure 1.3: Feeder geometry in DEM	36
Figure 1.4: Correlation plot for DSD DOE	39
Figure 1.5: Model fit for shear angle using main effects	44
Figure 1.6: Model fit for shear angle using main and interaction effects	44
Figure 1.7: Profile of feed factor vs feeder weight	47
Figure 1.8: Profile of flowrate vs time with feeder refills	48
Figure 1.9: Image of the left shows refill particles of type 2 (highlighted as orange particles) being added to the feeder unit at 4.13 sec whereas image on the right indicates particles mixing during feeder operation	49
Figure 1.10: Total number of particles present in the feeder unit. Particles added at time = 0 seconds indicated in yellow whereas, particles added during refill at time = 4.13 seconds are indicated in orange.	50
Figure 1.11: Fraction of particles at the feeder outlet	51
Figure 1.12: Feed factor vs feeder weight curve-fit for the feed factor parameter	53
Figure 1.13: Contour plots of bulk measurements obtained from the drawdown test	58
Figure 1.14: Parity plot for FF_{max} including verification runs	59
Figure 1.15: Parity plot for FF_{min} including verification runs	60
Figure 1.16: Parity plot for β including verification runs	60
Figure 2.1: DEM simulation and PSM of blender (Printed from Gao et al. [41] with permission from authors)	69
Figure 2.2: PSM geometry for blender simulation - A) PSM dimensions, B) Components of PSM	69
Figure 2.3: Classification of mixing indices	81
Figure 2.4: Feature-based comparison of mixing indices	85
Figure 2.5: DEM simulation of blender using PSM - A) initial state of un-mixed powder system, B) Mixed state of powder system with chosen gridding for mixing evaluation	88
Figure 2.6: Mixing indices for evaluation of mixing in the periodic section with subplots for different indices.	90
Figure 2.7: T_{95} mixing times for indices at different blade speeds	93
Figure 2.8: Logical diagram for selection of appropriate mixing index	95
Figure 3.1: Discrete element modeling methodology	104

Figure 3.2: Geometric details of the periodic section used in this study (Taken with permission from Bhalode and Ierapetritou, 2020)	105
Figure 3.3: Simulation of periodic section using DEM - A) periodic section at t=0 sec, B) periodic section at t=20 sec	106
Figure 3.4: Proposed breakdown of periodic section - A) selection of periodic sections from DEM simulation of continuous blender, B) breakdown of periodic section for axial and radial compartments, C) compartment model representation for axial and radial compartments.	111
Figure 3.5: Distribution of filled and empty regions in the 1st slice of periodic section, corresponding to $x = 0$	113
Figure 3.6: Radial compartmentalization for slice corresponding to $x = 0$ within the periodic section - A) identified compartments for $x = 0$ along Y direction, B) identified compartments along Z direction, C) compartment map for Y direction, D) compartment map for Z direction, E) overlaid compartment map for radial direction (including Y and Z directions)	115
Figure 3.7: Axial compartmentalization for the periodic section - A) zone map for slice corresponding to $x = 0$, B) zone map for slice corresponding to $x = 1$, C) inter-connected zone map for axial direction	116
Figure 3.8: Validation of axial compartmentalization - RTD profile for DEM plotted for both particle types (P1_DEM, P2_DEM) and compartment model (RT_CM)	118
Figure 3.9: Relative standard deviation (RSD) profile for the periodic section as a function of time	119
Figure 3.10: Axial mixing for entire continuous blender obtained by convolution of 8 periodic sections	121
Figure 3.11: Extended RSD profile for the continuous powder blender	122
Figure 4.1: Comparison of different tracers for the same bulk material using the NIR spectra and the obtained tracer concentration profiles (Images taken with permission from Escotet et al. [48])	127
Figure 4.2: Experimental setup including feeder, blender, NIR probe, and vibratory feeder	131
Figure 4.3: Tracer concentration profiles for different tracer amounts with repeated runs	132
Figure 4.4: Noise present in tracer concentration profiles with different tracer amounts shown in each sub-plots, where colors (orange, green, and blue) are used to show the noise in repeated concentration profiles	136
Figure 4.5: Effects of selected smoothing window size on tracer concentration profiles for sliding average (SA) method	138
Figure 4.6: Effects of SV parameters on tracer concentration profile - A) effect of polynomial order, B) effect of smoothing window size for polynomial fit	140
Figure 4.7: Signal denoising using FFT - A) original noisy signal in time domain, B) frequency domain of noisy signal, C) denoising of noisy signal in frequency domain, D) denoising signal in time domain	141
Figure 4.8: Comparison of discrete wavelet families on tracer concentration profiles - A) original and denoised concentration profiles, B) original concentration profile, C) denoised profile using symlets, D) denoised profile using daubechies, E) denoised profiles using coiflets	144

Figure 4.9: Denoised tracer concentration profiles using denoising methods - A) sliding or moving average, B) SV filtering, C) FFT, and D) wavelet transform	148
Figure 4.10: Comparison of denoising methods demonstrated for 1st and 2nd experimental run	149
Figure 4.11: Material tracking using original and denoised profiles with lot boundaries defined based on 0.5% and 99.5% of $F(t)$	154
Figure 4.12: Funnel plots developed for A) original concentration profile, B) denoised concentration profile, C) both profiles overlaid together	155
Figure 5.1: RTD studies, A) experimental setup, B) tracer concentration profiles with repeated runs.	163
Figure 5.2: Steps involved in OOS analysis - A) obtaining tracer concentration profile, B) evaluating RTD $E(t)$ profile, C) quantifying OOS for input profiles when convoluted with $E(t)$, D) developing funnel plot analysis	165
Figure 6.15: Case studies for truncation of RTD profile - A) truncation of tail region, B) truncation along time, truncation along Y axis – focusing on C) low limits of detection, and D) high levels of truncation.	166
Figure 6.4: Funnel plots for different case studies - A) truncation of tail regions, B) truncation along time, truncation along Y axis – C) low limits of detection and D) high levels of truncation.	171
Figure 5.5: Workflow for determining truncation criteria using funnel plots	173
Figure 5.6: $E(t)$ profiles for different levels of truncation (a) averaged for all repeats and (b) regressed using axial dispersion model	174
Figure 5.7: Original and truncated tracer concentration profiles	176
Figure 5.8: Overlaid funnel plots for truncated and original tracer concentration profiles	177
Figure 6.9: The original and regressed $E(t)$ profiles using PAT-based truncation method	177
Figure 7.1: Analysis of variance and parameter estimates of main effects for shear angle	185
Figure 7.2: Analysis of variance and parameter estimates of main and interaction effects for shear angle	186
Figure 7.3: Main and interaction effects for mass in lower container of the drawdown test	186
Figure 7.4: Main and interaction effects for bulk measurement of angle of repose	187
Figure 7.5: Main and interaction effects for bulk measurement of maximum flow rate	187
Figure 7.6: Parity plot for FF_{max} with effects summary	189
Figure 7.7: Parity plot for FF_{min} with effects summary	190
Figure 7.8: Parity plot for β with effects summary	191
Figure 7.9: Prediction profiles for FF_{max}	191
Figure 7.10: Prediction profiles for FF_{min}	192
Figure 7.11: Prediction profiles for β	192
Figure 7.12: RSD profile obtained from DEM simulation of periodic section	196

TABLE OF TABLES

Table 1.1: DEM material properties for parametric study -----	37
Table 1.2: DSD for DEM material parameters -----	40
Table 1.3: Bulk measurements using DEM drawdown test-----	42
Table 1.4: Estimation of DEM material parameters for a given set of bulk measurements -----	57
Table 1.5: DEM parameter settings for model verification runs -----	58
Table 2.1: Mixing indices with corresponding applications -----	86
Table 2.2: T95 mixing times for indices with different blade speeds -----	92
Table 4.1: Metrics for comparison of different denoising methods on RTD studies-----	149
Table 6.2: Regressed model parameters for AD model with the student's t-test comparison for different truncation levels -----	174
Table 6.2 Regressed AD model parameters for the original and PAT-based truncation -----	178
Table 7.1: Regressed feed factor parameters for the selected DOE-----	188
Table 7.2: Results of model fit for blender operation with a blade speed of 100 RPM-----	192
Table 7.3: Results of model fit for blender operation with a blade speed of 75 RPM -----	193
Table 7.4: Results of model fit for blender operation with a blade speed of 50 RPM -----	194
Table 7.5: Details of DEM model parameters used for periodic section -----	195

CHAPTER I

1. Introduction

1.1. Pharmaceutical industry

The global pharmaceutical industry accounting for research, development, production, and distribution of medications has been growing significantly over the past two decades, with revenues reaching to 1.27 trillion U.S. dollars in 2020 [1, 2]. Amongst all the geographies worldwide, North America holds a major share of pharmaceutical revenue, approximately equal to 49% [1]. A huge fraction of this revenue is allocated for research and development (R&D) of new drugs in the pharmaceutical industry [3]. Pharmaceutical R&D includes all steps involved in development of new drugs, starting from initial research of disease, selection, and formulation of drug substance for the disease, formulation of drug product, pilot plant testing, testing over pre-clinical and post-clinical trials, manufacturing operations, and distribution of end drug product. During these stages, regulatory agencies such as U.S. Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), and European Medicines Agency (EMA) play a crucial role in regulating and approving the drug product to be distributed in the markets. Currently, the industry allocates roughly 200 billion U.S. dollars globally to the R&D sector [3]. These numbers have steadily incremented from a lower value of 137 billion U.S. dollars in 2012 [3]. There is a significant emphasis of R&D in the pharmaceutical industry given various factors such as the drive to manufacture innovative drugs, time-limitation of drug patents, threat of production of generic or biosimilar drugs that hamper market competitiveness, eventually affecting revenue of the drug product. These factors have led to a focus on R&D of new drug substances and specialty drugs along with emphasis on optimal ways to manufacture these drug products at a global scale. CDER approved 53 new drugs in 2020, with an average cost of 2.6 billion U.S. dollars required for developing a new pharmaceutical drug [4]. This trajectory of overall expenditure on R&D in the

pharmaceutical industry is predicted to be on the rise in the coming years and is estimated to amount to more than 250 billion U.S. dollars by 2026 [2, 3, 5]. To manage these expenses, pharmaceutical companies are coming up with innovative strategies to aid manufacturing and reduce overall costs. Some of these strategies include outsourcing large parts of R&D to other companies in partnership with the pharma companies focusing on manufacturing aspect of drug development and use of big data and machine learning strategies in clinical research to accelerate the discovery and identification of suitable drug substances as well as to aid in to develop robust manufacturing lines [6]. The introduction section focuses on providing a state-of-the-art literature review of the current practices being implemented and researched in the pharmaceutical manufacturing, aimed towards identifying the gaps and challenges currently encountered in the industry.

1.2. Pharmaceutical manufacturing

Pharmaceutical manufacturing is primarily divided into two main processes based on the desired end drug commodity – drug substance and drug product. Manufacturing of drug substance involves development of the active pharmaceutical ingredient (API) via synthesis, either in liquid or solid (granular) form, also known as upstream manufacturing. Following the development of API, the drug substance is combined with more pharmacologically inactive materials (known as excipients) to develop a drug product, aimed to ensure ease of administration of API for the patient, desired pharmacological availability of the drug substance within the patient's body, and stability of the drug product during transportation to the desired geographical region. Manufacturing of drug product, also known as downstream manufacturing, is further classified into various types (such as oral solid drug products or tablets, injectables, spray, ointment) based on the intended method of administration. Among these types, oral solid drug products or tablets comprise most of the drug products, roughly around 60%, and it's manufacturing commonly involves equipment capable of handling granular or powder flow [7]. Oral solid drug products have been traditionally developed

using the batch manufacturing route in the pharmaceutical industry. Batch manufacturing involves flow of material in disconnected batches where large number of laborers are required to move material from one unit to another in batches and ensure appropriate and timely discharge and cleaning of each unit operation used in the process. Due to a high degree of discontinuity in the manufacturing process, batch manufacturing often does not result in an efficient process for drug manufacturing given the number of disconnected processing steps, required interference of human efforts, difficulty in process automation and, need for scale up for satisfying the increasing demand of drug products [8-10]. Alternative to batch manufacturing, continuous mode of manufacturing involves manufacturing of drug products using an integrated series of unit operations, where the input material is continuously fed within the process and transferred from the upstream unit operations to downstream with continuous production of drug products [11]. Given the inherent process integrity within continuous manufacturing (CM), it has the potential to integrate process robustness within the manufacturing itself [8-10, 12-19]. CM technology has been applied for production of drug substance as well as drug products and integrated lines are labelled as end-to-end (E2E) manufacturing [20-24].

1.3. Progress of pharmaceutical manufacturing in the past decades

Over the past two decades, the pharmaceutical industry has been on the path of modernizing drug product manufacturing to develop efficient, agile, and flexible routes for producing high quality drugs [8]. The revolutionary initiative can be traced back to the pivotal concept of Quality-by-Design (QbD) initially developed by Dr. Juran in 1992 [25]. Janet Woodcock adopted the concept of QbD to the pharmaceutical industry, defining a high-quality product to be free of contaminants and manufactured with minimal regulatory oversight while delivering the desired therapeutic effect [26]. It has also been defined as the vision for U. S. FDA's 'Pharmaceutical Quality for 21st Century' initiative [27]. The main motivation behind adopting QbD is to ensure that product quality can be designed within the manufacturing process itself and ensured without intensive product

testing, thus encouraging systematic scientific and risk-based approaches towards process design and manufacturing of pharmaceutical drug products. The concept of QbD further drove the shift from a batch mode of manufacturing to continuous, as it allowed QbD to be integrated within the process itself [8, 11, 28]. To motivate the adoption of QbD, FDA issued guidelines [29] for the pharmaceutical industry to adopt a science and risk-based approach for drug product manufacturing, which led to the foundation of adopting QbD within the pharmaceutical sector. ICH devised multiple standards, such as ICH Q8 (R2)[30, 31], ICH Q9[32], ICH Q10[33], ICH Q11[34], ICH IWG[35], ICH Q13[36], aimed towards providing high-level recommendations for QbD for CM within the pharmaceutical industry, known as pharmaceutical QbD. CM thus has the potential to integrate process knowledge and robustness within the manufacturing lines. Wahling (2021) provides an insightful review on continuous manufacturing of oral solid dosage drug products, outlining the technical and regulatory advantages, tax incentives, regulatory aspects, entire manufacturing end-to-end processes, current trends and industrial collaborations with academia and government agencies to progress the adoption of CM technologies in the pharmaceutical industry [12]. At the time of writing of this thesis, there exists 7 approved drug products manufactured using a CM route, approved by regulatory agencies such as EMA, FDA and PMDA [12, 13, 28]. A recent market research report indicated the market size of CM to grow at a potential rate of 8.8%, reaching 3.8 billion U.S. dollars in 2024 from the value of 2.3 billion U.S. dollars in 2018 [5]. The reason behind such a high increment rate is the advantage that CM provides over the batch manufacturing route. These advantages include faster transition from development phase to commercial production phase, product benefits such as improved product quality, longer shelf life, enhanced product reliability, cost benefits such as lower production costs, low labor cost, smaller footprint, improved utilization of resources. It further provides avenues for patient benefits such as suitable for personalized medicines and alternative manufacturing routes, and societal benefits such as reduced risk of exposure to API, reduced waste and environmental impact, higher safety of operators, and process intensification [7, 8, 10, 12, 13, 20, 37-42].

Despite all these advantages, it has been challenging for the pharmaceutical industry to adopt CM over batch manufacturing. One of the key challenges is that the existing equipment and facilities are originally oriented towards batch mode of manufacturing and thus require additional capital investment for adoption of CM [12, 42]. Further, there exists a possibility that the existing equipment is depreciated so the incentive to invest the additional capital reduces [42]. Another challenge involves the lack of expertise and experience of plant operators with limited experience of drug product submissions using CM [12]. Industries are apprehensive about switching from batch to continuous mode of manufacturing given the differences in the drug approval requirements and approvals required for switch from batch to continuous, dearth of globally recognized CM approval process, need for expertise in on-line or in-line process monitoring using PAT, verification, control, and QbD quality assurance [37]. To overcome these challenges and allow a smooth transition from batch to continuous, regulatory agencies have outlined detailed guidelines over the years and regulatory incentives such as tax incentive, expedited approvals for products developed using continuous as compared to batch [8]. They have also encouraged active collaborations between industry, academia, and regulatory organizations to aid the industry in moving towards adopting CM practices that ensure consistency of product quality with reduced manufacturing costs [12, 43].

1.4. Manufacturing routes in CM

Continuous manufacturing of oral solid dosage drug products involves different routes of manufacturing based on the material and powder flow properties of the ingredient components. These routes include direct compaction, dry granulation, wet granulation, spray drying, and extrusion [7, 10, 12, 13, 41, 44]. Direct compaction involves direct compressing of ingredient powder blend into a tablet and include unit operations ranging from powder feeders to ensure consistent powder flow of ingredient powders, powder blender to ensure uniform mixing of powders and obtain desired blend uniformity, and tablet press with feed frame to transition the

powder blend into tableting cavities, to be compacted to develop tablets. Direct compaction is used for ingredient powders in the tablet formulation with good flowability. For formulations with ingredient powders having poor flowability such as cohesive powders and powders with non-uniform particle size distribution (PSD), granulation routes including dry and wet granulation, are employed[45]. In dry granulation, powder feeders and blenders are used to obtain a homogenous powder blend (similar to direct compaction) [46, 47]. Following this stage, the powder blend is supplied as an input to a roller compactor where powder is compacted into ribbons and these ribbons are uniformly broken down into granules of equal sizes using a milling unit at the exit. These granules are then compressed into tablets using a tablet press. Alternatively, wet granulation route uses a granulating liquid to obtain more uniform PSD and improve powder flowability [48-52]. This route includes powder feeders and blenders to obtain a homogeneous powder blend, which is provided as an input to the granulator. Here, a wetting liquid (typically water and a high-density polymer) is added to the system to improve the PSD via agglomeration[52, 53]. The wetting liquid is subsequently removed in the next unit operation – dryer and the PSD is reduced to the desired size range using a screen mill. The uniform milled powder blend is finally then compressed into tablets in the tablet press. Spray drying route is used for situations where the solubility of the formulation needs to be improved. Thus, this route involves the powder blend (after blender) being sent to homogenizing unit and spray dried using an air-drying system [54, 55]. The homogenized dried powder blend is then sent to a milling unit to be milled followed by a tablet press to compress and make tablets. Last route involves extrusion of powder blend using hot melt extrusion units such as single screw extruder and twin-screw extruder [56-58]. This route includes powder feeders being fed to an extruder unit wherein, the formulation ingredients are mixed with lubricant, melted, broken down into pellets to be transferred to a tablet press unit to formulate tablets or a capsule filling machine to develop capsules. All these routes for continuous manufacturing are outlined in Figure 1.1. Out of these different routes of manufacturing, direct compaction is the simplest (with minimum number of unit operations) and most energy efficient route of manufacturing as it does

not include any heating or cooling components, and wet granulation and spray drying tends to be the most energy intensive route with use of drying components [59]. Thus, direct compaction is commonly favored for situations where the ingredient powders inherently have good flowability. The manufacturing routes are also chosen in combination with the stage focusing on developing the formulation of the drug product where desired mechanical flow, pharmacokinetic, pharmacodynamic properties and the biological availability of the drug product inside the human body can be modified [44].

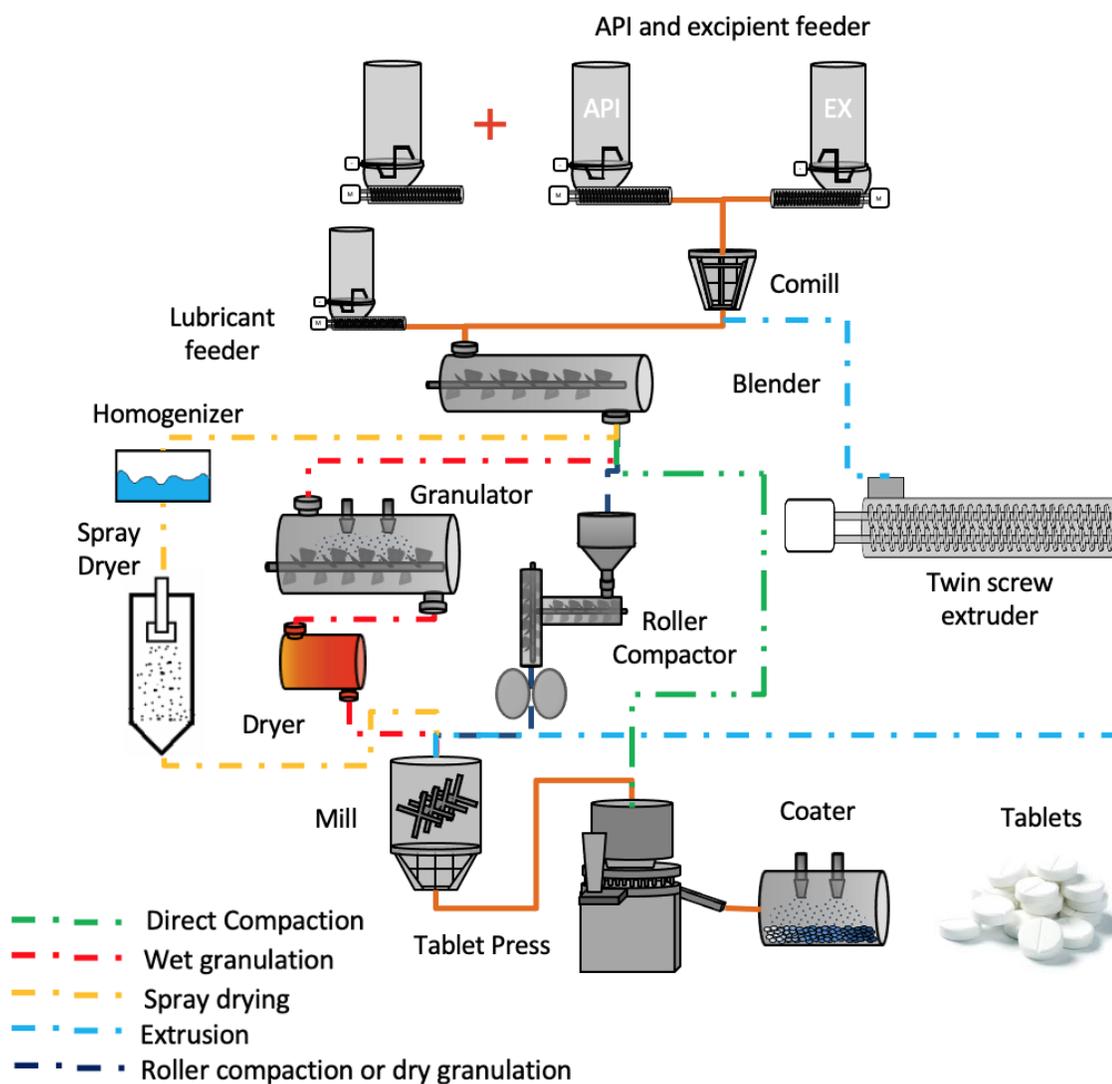


Figure 1.1: Different routes of continuous pharmaceutical manufacturing - direct compaction, wet granulation, spray drying, extrusion and roller compaction or dry granulation (Modified from Escotet-Espinoza et al.[44])

1.5. Current trends in pharmaceutical manufacturing

Given the increasing level of digitization, digital advancements are currently being explored in the manufacturing sector to improve innovation, overall profitability, and productivity. Such advancements have given rise to applications of artificial intelligence, Internet of Things, and Industry 4.0 with technical and financial backings from academic and government organizations to

bring about the paradigm shift of digitization in the manufacturing sector [60-64]. This transformation is also being investigated for the pharmaceutical manufacturing industry with integration of various smart technologies such as cloud computing (CC), cyber-physical systems (CPS), and big data analytics with machine learning applications (BDA) within the drug manufacturing lines to transform the current routes of manufacturing towards smart manufacturing [65-67]. With these technologies, important manufacturing activities such as remote sensing, real-time testing, process monitoring, and process visualization would become more accessible and easier to integrate within the current manufacturing platforms, aimed towards developing digital twins of the manufacturing lines to enable robust manufacturing for the pharmaceutical industry [65, 68-70]. Digital twins, first coined by Michael Grieves in 2002 [71], is essentially a digital construct of a physical system that operates on its own and is fully integrated with the physical system to ensure real-time representation of the physical system in the digital construct. Chen et al (2021) provides a detailed review of digital twins for pharmaceutical and biopharmaceutical manufacturing wherein, the authors show a schematic representation of the digital twin for these manufacturing lines [65]. The digital twin consists of two components - a virtual or digital component and a physical component (manufacturing plant) that are seamlessly interconnected with each other such that real-time system behavior observed in the physical component is captured in the virtual component and process predictions with control strategies are devised in the virtual component in real-time, to be implemented in the physical component. Digital twins thus utilize smart technologies such as real-time process data acquisition using advanced process monitoring strategies, integration of various sources of process data, cloud storage and computing, quick predictions of system behavior using predictive process flowsheets, identification of real-time process control in situations of process disturbances, instantaneous transfer of information from the virtual component to the physical plant, and implementation of control strategies to ensure robust manufacturing practices. To advance the adoption of these technologies and enable industries to manufacture regulation compliant drug products, regulatory agencies such as U.S. FDA and ICH

are collaborating with industries and academic institutions to realize their vision of the pharmaceutical industry for the 21st century. This move of process digitization towards smart manufacturing is coined as 'Pharma 4.0' [61, 62, 72, 73].

To develop a true digital twin of the pharmaceutical manufacturing lines with a seamless transfer of information between the two components, there are critical challenges that need to be addressed. These are outlined in detail by Chen et. al (2021) and include the heterogeneity of sources from which process knowledge is obtained from the physical plant, development of predictive models that can be incorporated within the real-time, and accurate prediction framework of virtual component, and maintenance issues such as cyber-security [61]. Heterogeneity of sources generating data in the physical plant arises from different process analytical tools and the inbuilt incompatibility of the software within these tools. Several manufacturers are working towards ensuring this by utilizing the commonly used OPC interface. Recent efforts from academic institutions towards addressing this challenge involve the development of an inter-connected I4.0 framework for data collection, transfer, and storage for enabling seamless connection with the virtual component [66]. Another significant challenge involves the need for development of predictive models that allow quick and accurate predictions in the virtual component for system analyses such as risk assessment, process control, diversion of material, system health monitoring, feasibility analysis, and optimization [65]. The predictive models of unit operations in the manufacturing lines need to be developed balancing the level of accuracy retained in the models and the computational time required for process simulation. The more accurate the predictive models are, longer is the associated computational time, making them incompatible with flowsheet models. On the other hand, the predictive models that provide quick simulations lack the physical insight into the process dynamics making it hard to adapt to a new unforeseen situation in the plant operation. Thus, it is important to develop integrated process models that combine the mechanistic knowledge and enable process simulations within reasonable computational times such that they

are better equipped with handling variabilities and process fluctuations. The capability of ensuring reasonable computational time is especially crucial as the simulation time required for overall flowsheet increases incrementally with each unit operation present in the manufacturing line and can severely impact the real-time capabilities of the digital twin framework.

1.6. Background of process modeling in pharmaceutical manufacturing

Process models are essentially centric to the virtual component of the digital twin framework and useful to simulate and replicate the system dynamics as observed in the real physical system [65]. There exist various categories of process models based on the extent of mechanistic knowledge and data-driven learnings incorporated within the process model [7, 44, 74-76]. Mechanistic knowledge corresponds to the process information derived from first-principles or fundamental equations of motion. Data-driven models, on the other hand, correspond to empirical or surrogate models developed using machine learning techniques using the available datasets, to develop correlations between the input parameters with the desired output parameter. Such models do not incorporate or provide any mechanistic insight and are purely based on the available datasets. Between these two extremes, various process models have also been developed that lie within this spectrum and are outlined below. Readers are encouraged to explore recent review articles for in-depth description and application of these models in the pharmaceutical industry [41, 76, 77].

1.6.1. High-fidelity mechanistic models

High-fidelity mechanistic models are models based on first-principle or fundamental equations. These models thus encompass detailed process physics observed in the system, using mechanistic equations to describe the movement of granular flow. Some examples of these models include discrete element modeling (DEM) and computational fluid dynamics (CFD) based on the discrete and continuum models. DEM has been widely used for modeling of pharmaceutical powder flow within unit operations to simulate free-flowing and cohesive powders [78, 79]. The workflow of

DEM is based on Newton's laws of motion. Different contact models are used to simulate particle interactions in the system that further dictate the particle forces acting on the particles and motion of particles. Equations are written for each particle interaction in the system for each time step and thus such models can get quite expensive for simulation of realistic unit operations [80-82]. Alternatively, researchers have also explored mesh-based computational fluid dynamics (CFD) methods such as finite element method (FEM) and finite volume method (FVM) for simulation of pharmaceutical unit operations [83-93]. However, these methods can become challenging for simulation of free-surface fluids such as partially filled systems. Authors investigated sophisticated mesh refinements to address this challenge and simulate free-surface fluids [94]. Recently, smoothed particle hydrodynamics (SPH) has been shown to model particulate systems where, SPH fluid elements are simulated as Lagrangian elements [56, 95]. As SPH is mesh-free, partially filled systems can be simulated in similar computational times as filled systems, which otherwise become challenging using CFD and FEM approaches. Though these high-fidelity models can provide detailed mechanistic understanding of particle dynamics from a particle scale, they lead to huge computational times in the order of few hours making them difficult to be integrated within process flowsheet models.

1.6.2. Multi-dimensional mechanistic models (PBM)

Multi-dimensional mechanistic models that incorporate simulation of multiple dimensions in combination with the particulate flow. Population balance model (PBM) is an example of such models which combine external properties (particle locations, collision) with the internal properties of particles (particle size, attrition, moisture, porosity). PBM models have been widely used for particulate simulation of unit operations involving change of particle sizes such as granulators, dryers, and milling units [48, 49, 53, 90, 93, 96-99]. These models being computationally cheaper than DEM result in a smaller simulation time than DEM. The reduction in simulation time is achieved by grouping discrete systems and related mechanistic equations into groups and

considering their averaged values. Thus, discretized PBM models are developed with model parameters regressed using experimental datasets [100].

1.6.3. Phenomenological or semi-empirical models

Phenomenological or semi-empirical models are intermediate between mechanistic and data-driven models [44]. Developed based on the assumptions in the first-principle models, these models (commonly known as engineering models) are computationally quick to simulate for particulate systems. An example of such models includes low-dimensional PBM model, involving mass and energy balances to provide system behavior combined with empirical models to correlate process variables and process conditions to the desired process outputs. Residence time distribution (RTD) is an example of low-dimensional PBM models. RTD basically quantifies the probability distribution of time that a material spends or resides in the system under study [101]. RTD has been extensively used for understanding powder mixing and developing predictive models of unit operations and process flowsheets along with flowsheet applications such as quality assurance, material diversion, system health monitoring, and process control [102].

1.6.4. Empirical or data-driven surrogate models

Empirical models commonly known as data-driven or surrogate models are models developed purely to capture the correlation between process inputs including processing parameters and material properties to the process outputs. These models being purely statistical models, do not capture the underlying physics or mechanistic knowledge in the models and thus are low dimensional in nature and provide quick simulation times. Some examples of the methods used to develop empirical models include multivariate regression, principal component analysis, partial least squares, response surface, black-box modeling, artificial neural network, gaussian process regression, support vector machines, and kriging [103]. As these models are purely based on the

available data of inputs and outputs, these models are recommended to be strictly used within the defined domain and not for extrapolation outside of the explored input space. Such models have been extensively used to develop process models for unit operations with quick computation times to predict output parameters for a defined set of input variables.

1.6.5. Hybrid models

Hybrid modeling strategy essentially combines the empirical or data-driven modeling with the mechanistic models derived from first principles. Thus, hybrid modeling thus integrates the process knowledge based on the first principles as well as the available datasets. Chen et al. (2020) provides a detailed review of hybrid modeling in the pharmaceutical industry outlining the different structures for constructing a hybrid model [104]. The different structures include series, parallel and combined structure. Hybrid modeling has been very widely used in the pharmaceutical manufacturing sector as well as other sectors such as energy and chemicals, to utilize data for unknown process dynamics integrated with the known mechanistic knowledge [81, 84, 105-110]. Hybrid models thus utilize available process data to ensure sufficient degree of retention of process information in the predictive model along with reasonable simulation times and is further used for identification of plant-process mismatch and correction in real-time [104].

1.6.6. Flowsheet models

A manufacturing process has multiple units or equipment connected to each other in sequence with the input material being added at the inlet of the first unit and products obtained at the outlet of the last unit in sequence. Similar flow of unit operations is used for manufacturing of oral solid drug products as outlined above in Figure 1, wherein, each unit operation is connected to the next in sequence and there is a flow of material between the unit operations. Continuous manufacturing involves continuous flow of material whereas, batch manufacturing involves flow of material in

batches [8]. To develop an integrated process model for replicating the material flow behavior, process models for individual unit operations are also integrated together such that the model inputs are defined based on the outputs of the previous unit operation and the model outputs then define the inputs of the subsequent unit operation model. This connected network of unit operations resembles a flow of material as observed in the actual manufacturing process and is known as the flowsheet model. Flowsheet models are mathematical representations of series of unit operations connected in sequence that resemble the actual manufacturing process and have been widely applied for process analysis such as scenario, sensitivity, and feasibility analysis as well as process optimization and risk assessment [75, 111-119].

Such models can be developed using standard modeling languages such as MATLAB or Python and commonly used flowsheet software such as gPROMS, Aspen Plus and SimaPro [23, 75, 117, 119-123]. Since flowsheet models include various unit operation models integrated together, it is important to consider the overall computational time required by each unit operation model. Thus, it can become challenging to incorporate high-fidelity models that are computationally intensive within flowsheet models leading to a simulation time in the order of days. To ensure reasonable simulation times of the process flowsheets, commonly low dimensional models are recommended for process flowsheets such as empirical or surrogate models and phenomenological or semi-empirical models. Research efforts have focused on integration of multi-dimensional models such as PBM within process flowsheets to incorporate a higher degree of process knowledge with a simulation time of few hours [124]. However, currently flowsheet models are incapable of incorporating high-fidelity models that provide detailed mechanistic insight while ensuring reasonable simulation times. Thus, there is a need to balance the two aspects – incorporation of mechanistic knowledge within the flowsheet models to ensure accurate predictions and reasonable simulation times for its application for system analyses. Ensuring reasonable simulation time is quite important especially in digital twin applications where process predictions need to be made

in real-time in the virtual component along with recommendations for process correction to be supplied to the physical component for corrective action. This transfer must be quick enough in the digital twin for ensuring real-time process monitoring and control to avoid production of non-regulation compliant drug product during manufacturing.

1.7. Current challenges in the pharmaceutical manufacturing

With the increasing encouragement from regulatory agencies and focus on development of digital twin applications for smart manufacturing of pharmaceutical drug products in collaboration with academic institutions, it is becoming important to address some of the inherent challenges outlined below.

1.7.1. Capturing mechanistic insight in the process models to improve process predictability

The unit operation models that are commonly developed for process flowsheets and flowsheet applications including system analyses and optimization utilize low-dimensional models to ensure quick simulation times. This is especially important for flowsheet applications where various unit operation models are integrated together sequentially which further increases the overall simulation time of the process flowsheet. However, the models severely lack information about the actual process physics observed within individual unit operations. As the details of powder dynamics is missing, the models essentially are limited to the variable space explored (which is normally devoid of process disturbances or uncertainties in manufacturing) and cannot confidently be extrapolated for uncertain situations. Furthermore, due to lack of physical insight into the operation of individual unit operations, the unforeseen aspects such as presence of dead zones developed dynamically within the unit operations can potentially hamper the level of accuracy of the developed predictions. These variabilities are further amplified with various powder properties of pharmaceutical powders

based on the external conditions such as humidity, pressure, and temperature. Such predictions when used for the correction of plant operations (as in the case of smart manufacturing for digital twin applications) can lead to erroneous conclusions and corrective actions, which can potentially be detrimental for production of regulation compliant drug products in the manufacturing lines. Thus, it is important to investigate and incorporate the details of particle-scale dynamics using high-fidelity simulations, such as DEM and CFD, within process models aimed towards developing accurate predictive models for unit operations. An additional challenge while developing such simulations for understanding powder flow and dynamics for various unit operations is to ensure that an accurate representation of experimentally used material is developed within the simulation environment such that the bulk flow of material is captured correctly. This challenge, commonly known as calibration, has been attempted by several authors using bulk calibration tests aimed to correctly simulate the bulk flow of material as observed experimentally [82, 125-142]. However, the calibration approaches presented in literature cannot address the problem of solution multiplicity where multiple solutions are present [139-141, 143, 144]. Thus, it is important to identify a unique solution for the calibration, followed by developed calibrated high-fidelity simulations to accurately understand powder flow dynamics observed in the unit operations. This would eventually help develop mechanistically informed process models for accurate process prediction.

1.7.2. Developing computationally efficient process models

Despite use of high-fidelity models such as DEM and CFD to understand particle scale dynamics during simulation of unit operations, these models cannot be directly integrated within the flowsheet model due to the huge computational expense associated with such simulations ranging in the order of hours to days. Such models when integrated within the flowsheet models would substantially increase the flowsheet simulation time, making it incompatible for flowsheet and digital twin applications. Thus, there is a need to transform the high-fidelity models into

computationally efficient process models while retaining the intended degree of process information. These models need to be developed such that they retain the systemic (pertaining to the overall system) and local (pertaining to the particle scale) information systematically in a computationally efficient framework [145]. Systemic or global information commonly corresponds to the global information relating the inputs to the outputs and includes overall mass and energy balances, component mass balances, residence time distribution (RTD) for tracing. A common example of systemic model is representing unit operation by a network of ideal flow systems (continuous stirred tanks and plug flow systems) such that the bulk flow profile of the unit operation is replicated by the network. As systemic models mainly focus on capturing the bulk flow profile of the unit operation, local particle scale information such as velocity and concentration profiles, and phase fractions are not captured. Local models on the other hand, thus focus on capturing the detailed local information about the process from mechanistic simulations such as CFD and DEM. Normally, systemic information is used to obtain the degree of axial mixing in the system whereas, the local particle information is used to quantify the degree of radial mixing, especially for unit operations such as powder blenders. There exist various mixing indices that aim to quantify the degree of mixedness of particles in high-fidelity simulations [146, 147]. However, it can become confusing to identify the most appropriate mixing index based on the intended application. Thus, there is a need to compare the different mixing indices and identify the most appropriate index for local models to be used for developing process models of unit operations. Furthermore, there is a need for systematic integration of the systemic and local models, to develop computationally efficient process models. Such models would thus be applicable for quick process simulations with accurate predictions for integration within a digital twin framework.

1.7.3. Statistical analyses for systemic information in manufacturing lines

The unit operation models are integrated together to develop flowsheet models, which act as virtual component of the digital twin to replicate the process behavior of the entire manufacturing line.

The virtual component can further be used for applications such as system analysis, techno-economic analysis, and optimization [75, 112, 117, 122, 148-150]. However, to develop digital twin, it is important to ensure the systemic information of process flowsheet match that of the manufacturing lines obtained experimentally. This is especially important for experimental component of the digital twin, as the experiments can get affected by unseen external noise and variability and can appear differently from the virtual component despite accurate characterization. External noise plays a significant role in the experimental measurements, coming from various sources such as environment, humidity, temperature effects, human disturbances, process disturbances, and external factors leading to disturbances during the experimental operation. These are specifically important and can severely impact the systemic knowledge of the unit operation under investigation. One such example involves experiments to obtain the residence time distribution (RTD) of the system under study. Along with the experimental disturbances, RTD being measured using process analytical technologies (PAT) further gets impacted based on the accuracy of the developed PAT models. As the PAT measurements are relative to each other, a variability or noise in the system can impact the overall RTD profile and subsequently affect the systemic information of the manufacturing line under study [151, 152]. Literature on RTD applications in the pharmaceutical industry currently lacks well-defined guidelines on handling such variabilities or noise captured in RTD profiles, obtained experimentally. Furthermore, obtaining RTD experimentally can become quite intensive in terms of the time and resources. This is especially pertinent for capturing the tail and baseline regions of RTD accurately, as there exists a high degree of variability or noise. Combined with the inherent noise present in the system, it becomes challenging and time intensive to obtain true RTDs of the system. Thus, there is a need to investigate if the important aspects of RTD are well captured, combined with the noise handling techniques, and subsequently identify ways to optimize the time and efforts required in obtaining RTD experimentally.

The above-mentioned challenges in the pharmaceutical manufacturing have motivated the research directions of the thesis and the research objective with specific aims are outlined in Chapter 2.

CHAPTER II

2. Research Objectives

2.1. Overall research goal

The research presented in this thesis is motivated by the current directions for pharmaceutical manufacturing and the respective challenges as discussed in section 1.7. Thus, the overall research goal is *‘to develop multi-scale process models that incorporate detailed mechanistic insight and are computationally efficient for developing process flowsheets along with statistical analysis and data pre-treatment of experimentally obtained process knowledge, aimed towards developing a digital twin to incorporate smart manufacturing technologies.’* This goal is accomplished based on a series of research aims outlined below.

2.2. Specific research aims

2.2.1. Aim 1: Developing high-fidelity models for obtaining particle-level knowledge for pharmaceutical unit operations

Aim 1 focuses on developing high-fidelity models to understand particle scale interactions and particle dynamics observed in pharmaceutical unit operations.

To improve the process predictability of unit operation models for flowsheet applications, it is important to incorporate detailed understanding of the particulate mechanics observed in the system under study. However, the available process models for flowsheet applications being mainly limited to empirical or semi-empirical models, fail to capture the desired level of detail concerning particle mechanics [75]. To address this concern, high-fidelity simulations need to be developed for pharmaceutical unit operations. Such simulations thus are useful to develop accurate predictive models and are especially useful for understanding process behavior during uncommon processing

conditions such as process or external disturbances. An important challenge to develop detailed high-fidelity simulations such as DEM, involves ensuring the powder properties are well represented within the simulation environment to capture the true powder behavior as observed experimentally. Powder calibration though has been investigated widely in the field, identifying a unique solution to represent the powder system has been a challenge [129, 130]. This is important as the unique solution enables capturing similar powder behavior in situation of process disturbances.

This research aim focuses on developing a unique solution for powder calibration, followed by investigation of powder mechanics for a pharmaceutical unit operation - powder feeder. Feeders are chosen as they are mainly the source of process fluctuations in powder flow and thus obtaining mechanistic understanding of powder flowability can aid in controlling the variabilities in powder flowrate [153]. This research aim thus attends to powder flow from a particle scale perspective and is outlined in Chapter 3.

2.2.2.Aim 2: Process modeling of unit operations using DEM

Aim 2 focuses on developing computationally efficient process models using the high-fidelity models developed in Aim 1 to be incorporated within process flowsheets for digital twin applications.

Though high-fidelity simulations can provide detailed particulate insight into the powder mechanics, they lead to huge computational costs in the range of hours to days. This limits the application of such models in process flowsheets, where quick process assessments are desired to devise corrective actions or control strategies for the overall process operation. Thus, a synergistic approach is required which ensures quick process simulation with retention of sufficient degree of mechanistic knowledge about the particulate system. Thus, this aim focuses on developing computationally efficient process models using high-fidelity simulations such as DEM. To develop such models for integration within process flowsheet, it is crucial for them to incorporate systemic

or global information about the process as well as local (or particle scale) information about the powder mechanics. Thus, this aim focuses on developing of compartment models integrating systemic and local level knowledge. Such modeling approach is demonstrated for development of process model for continuous powder blender which is widely used in all manufacturing routes for CM production. The first part of this aim deals with developing mixing indices to characterize the systemic information – blend uniformity based on local information of particle positions in the blender. A detailed literature review of mixing indices is performed and guidelines for selection of appropriate index is outlined based on end application. Following this, the compartment model is constructed for continuous powder blender which ensure reasonable computational cost along with retention of important process knowledge. Thus, this aim focuses on model development at a unit operation level and is outlined in Chapters 4 and 5.

2.2.3. Aim 3: Data pre-treatment and analysis of manufacturing lines

Aim 3 focuses on experimental data pre-treatment and analysis of manufacturing lines. This aim is divided into two sub-aims outlined below. First sub-aim focuses on data handling and analysis of systemic process knowledge obtained experimentally from the manufacturing lines. Systemic RTD obtained experimentally commonly encounters noise fluctuations and variabilities in the RTD profile given environmental, human, and external factors. As it is almost impossible to eliminate all these factors to avoid noise in the captured RTD profiles, there is a need to devise appropriate noise characterization and noise handling strategies to address the additional noise added in the system. This is important as noise present in the RTD profiles used for RTD-based applications of manufacturing line can impact the process characterization and lead to erroneous conclusions of the overall system dynamics. Such datasets when used in digital twin applications can further lead to ill-defined process control and corrective measures for the physical plant, gravely affecting the production of drug products. Thus, the first part of the aim deals with devising appropriate noise handling and denoising strategies for RTD datasets whereas, the second part addresses the aspect

of optimizing experimental efforts to obtain the RTD. This is performed by understanding the impact of different aspects of RTD (by truncation of RTD) on end RTD-based application – quality assurance. Based on this understanding, the efforts to characterize the RTD experimentally can be directed to focus on capturing the important features of RTD more prominently, thus optimizing the overall experimental efforts. This would also reduce the effects of environmental and external noise being captured in the RTDs and prevent erroneous conclusions in RTD-based applications such as quality assurance. These sub-aims are outlined in Chapter 6.

CHAPTER III

3. Aim 1: Developing high-fidelity DEM simulation for obtaining particle-level knowledge with calibration for pharmaceutical unit operations

Acknowledgement of publication status:

Entire sections of this chapter have been published in a scientific journal written by the author of this thesis with the title: Discrete element modeling for continuous powder feeding operation: Calibration and system analysis. This article was accepted in the International Journal of Pharmaceutics on 9th May 2020.

3.1. Introduction

Due to the increasing attention towards regulation compliant continuous production in the pharmaceutical industry [42], there has been a surge of research efforts focusing on efficient modeling of continuous pharmaceutical unit operations [7, 16, 17]. In order to provide accurate predictions for implementation in accurate ‘digital-twin’ framework, the models need to be constructed based on detailed process understanding and knowledge of process operation [10, 46, 75, 116, 117, 120, 154-156] including the effects of powder material properties on process operation. The inclusion of material properties is of extreme importance as it has been observed that pharmaceutical powders can have erratic fluctuations in powder flowability like rat-holing and dynamic powder behavior like sticking to the walls of the unit. These fluctuations can disrupt the downstream tablet content uniformity, if not considered and can seriously undermine the predictive ability of the developed process models.

Process models that have been developed so far, do not incorporate a complete material property dependence accompanied with mechanistic understanding of observed powder behavior. In absence of this understanding, the developed process models might fail to predict accurate powder behavior for a new powder material. To address the above-mentioned issue, this article focuses on the detailed process modeling of feeder units to gain insight into the effects of material properties on process operation.

Continuous pharmaceutical manufacturing lines consist of feeder units placed at the top, to obtain a consistent powder flow for the downstream unit operations. To ensure this with minimum fluctuations, a loss-in-weight feeding operation is used [157]. The loss-in-weight feeder consist of a hopper and a twin-screw section with a bridge-breaking impellor near the entrance of screws [158, 159]. Weighing scales kept below the feeder unit allows for continuous monitoring of the rate of change of feeder weight. As the powder leaves the unit, the difference in feeder weight amounts to the output flowrate. This measurement is sent to a PID controller to adjust the screw speed of the feeder unit. Before the start of the operation, the feeder unit is filled with powder. During the operation, screw speed is controlled using the PID controller, and powder is dispensed from the feeder outlet. If the powder bed height reduces beyond a certain set level, a refill system is used to supply more powder into the feeder system. Based on the refill and screw speed of the feeder unit, there are two operating modes – volumetric and gravimetric which are used to control powder flow. Volumetric mode works as an open loop, where screw speed is set to a certain value and the flowrate is not controlled by adjusting the screw speed. Gravimetric mode is a closed loop operation wherein the screw speed is adjusted to control the flowrate at the feeder outlet. The feeder flowrate ($F(t)$) is evaluated as a function of time (t) from screw speed ($w(t)$) and feed factor ($FF(t)$) as shown in Eq. (1).

$$F(t) = w(t) * FF(t) \quad (1)$$

Feed factor corresponds to the amount of powder present in one screw pitch which gets dispensed by the feeder in one screw revolution [157, 159]. Thus, it depends on the compressibility of the powders at the screw section since the larger the compressibility, the more powder gets compressed within the screws and the higher is the observed feed factor at the outlet. Since powder compressibility affects the relative bulk density of powder, feed factor also relates to the relative bulk density of the powder under compression in the screw section. Based on this relationship, Escotet-Espinoza et al. developed and verified the feed factor model based on Heckel's equation experimentally, as shown in Eq. (2) [44, 75, 160].

$$FF(t) = FF_{max}(t) - (FF_{max}(t) - FF_{min}(t)) * \exp(-\beta W(t)) \quad (2)$$

where, the physical meaning associated with the feed factor model parameters (FF_{max} , FF_{min} and β) can be derived from Heckel's equation. FF_{max} corresponds to the maximum relative bulk density within the screws, attained with maximum compression when the feeder hopper is full of powder, whereas FF_{min} corresponds to the minimum relative bulk density or powder compressibility within the screws, when the feeder is almost empty i.e., feeder weight approaches zero. Lowest value of feed factor is not physically observed during feeder experiments since, the refill system is set in place to add more powder to the feeder unit when the fill height of powder goes below a certain set level. Also, during experiments when the feeder unit is almost empty, there is still some amount of powder present in the hopper and screw section which does not exit the unit even when the flowrate is zero. Thus, feeder weight needs to be extrapolated to zero, to obtain the feed factor parameter $-FF_{min}$. β on the other hand, is related to the relative difference between FF_{max} and FF_{min} . In other words, FF_{max} and FF_{min} can be interpreted as maximum and

minimum quantity of powder present in one screw pitch of the feeder unit whereas, β corresponds to the rate at which amount of powder present in the screw pitch decreases with respect to decrease in feeder weight. Escotet-Espinoza (2018) performed experimental studies to determine how powder properties affect the feed factor parameters and developed regression models for the feed factor parameters as a function of process and powder material properties [44]. To obtain a detailed understanding of the powder behavior during the feeder operation, in this article we aim to explore the mechanics of powder flow using particle-scale simulations and obtain physical insight into the correlations of powder properties using the feed factor parameters.

Particle-scale simulations like Discrete Element Modeling (DEM) have been widely used to obtain first-principle understanding of powder behavior within process operation. These simulations focus on particle level mechanics including particle-particle and particle-wall interactions influencing the forces acting on each particle through contact models, which eventually dictate the particle trajectories. First proposed by Cundall and Strack in 1980 [161], DEM provides an insight into particle movement and powder mixing within the unit operation, thus reducing the need for expensive experimental studies. The reason behind the choice of DEM methodology is that DEM allows the user to create desired unit geometry and define particle material parameters. Once completed, the simulation allows for extraction of detailed powder properties like velocity and internal stresses. Thus, application of DEM can provide a detailed insight into analysis of powder behavior which would not be feasible using an experimental approach. Another advantage of using DEM over experimental approaches is that there is no loss of actual powder, which can be expensive especially for pharmaceutical APIs. Despite the above advantages, one of the drawbacks of using high-fidelity models like DEM is the associated computational requirements. Focusing on particle-level mechanics can get computationally intensive as the simulation needs to evaluate all the particle interactions and forces for all particles at each time step. Different research articles [49, 79, 96, 162-170] have described DEM simulations ranging from a few hours to few days extending

almost to a month using high computational resources, simulating anywhere from 10,000 to a few million particles.

Material calibration can greatly improve the prediction accuracy of the DEM simulation as it allows to evaluate input DEM material parameters which replicate experimentally observed nature of powder. It involves linking experimentally measurable material properties to input DEM material parameters like static friction and rolling friction. There are multiple authors [126-130, 132, 135, 136, 138, 170-175] who have attempted to improve material calibration, but it is still a developing area. Overall, there are mainly two approaches for material calibration [131] namely the direct measurement and bulk measurement. Direct measurement [176, 177] uses actual measurements for DEM parameters like static friction, rolling friction, shear modulus and coefficient of restitution. These measurements are obtained from specialized experimental tests like direct shear and particle impact and can get very expensive. Bulk measurement, on the other hand, is a preferred option [131, 135, 138, 140, 174] wherein simple bulk powder measurements like bulk density and angle of repose are obtained experimentally. Similar geometry setup is created in DEM and the bulk calibrations are matched to the experimentally obtained values by adjusting the DEM material parameters. Thus, DEM material parameters corresponding to the actual bulk properties are calibrated. Although there is a substantial volume of research in this area, the issue of solution multiplicity still persists wherein, multiple material properties are related to few bulk measurements [82]. These bulk measurements when used to estimate multiple DEM material parameters, can lead to multiple solutions as the corresponding system of equations is under-determined. In other words, the problem of evaluating a unique solution [82, 140] for material calibration still needs to be fully addressed. Considering all the above concerns, this article proposes application of DEM with material calibration for studying powder behavior during process dynamics of feeder unit as shown in Figure 3.1.

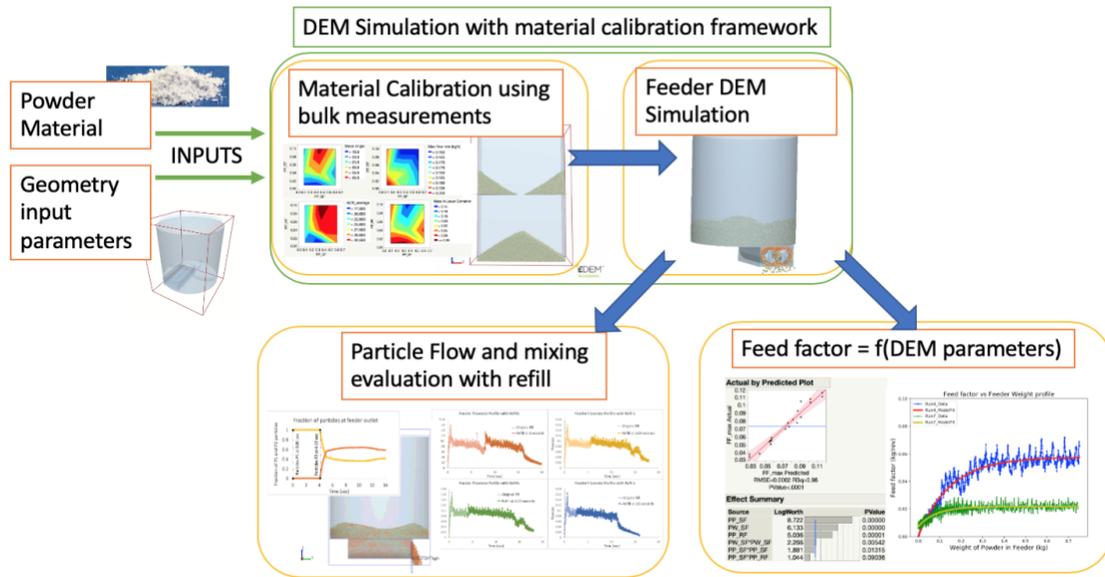


Figure 3.1: Proposed framework for calibration and application of DEM for powder feeder

The chapter is arranged as follows. Section 1 provides a detailed background and introduction to the problem statement. Section 2 highlights the methods implemented in this article including Discrete element modeling (DEM) applied for material calibration and feeder simulation with relevant information regarding the simulation setup within DEM. To understand effects of powder behavior on process dynamics, authors implement design of experiments for DEM simulations for material calibration and feeder simulation. Details of DOE used in this article is provided in section 2. Section 3 illustrates the results based on the simulation results for material calibration and feeder simulation. Material calibration is performed using a bulk measurement test to correctly represent powder properties within DEM. To provide insights into the feeder dynamics, refill studies are also simulated. Following this, the DOE results are highlighted to analyze correlations of material properties with the process operation. Following this, section 4 provides a discussion of the obtained results with implementation of calibration results for identifying a unique set of material parameters. For feeder unit, additional verification runs are performed and analyzed to confirm the previously obtained correlations of material properties and process operation. This concludes with

section 5 providing a summary and conclusion of the proposed work along with identification of possible avenues for enhancement of proposed work.

3.2. Materials and methods

3.2.1. Discrete element modeling (DEM)

DEM methodology has been widely used for particle-scale process understanding in pharmaceutical industry [178]. It involves evaluation of particle forces acting on each particle because of its interaction with other particles and the unit geometry. Normal and tangential forces are evaluated along with corresponding stresses for all particle which dictates the particle motion. Normal force on the particle contributes to the forward movement of the particle whereas tangential forces contribute to the particle rotation. The equations of particle motion need to be solved for all particles at each time step. For a DEM simulation, the required geometry is set up along with the selection of equipment and material parameters. User defined particle properties including particle size and shape, and number of particles to be added to the system are selected. Once the simulation starts, the particles are added to the geometry which interact with each other as well as the equipment geometry. During such interactions, equation of motion governs the motion of individual particle (i) when in contact with another particle or equipment (j) as shown in Eq. (3).

$$\frac{m_i dv_i}{dt} = \sum (F_{ij}^n + F_{ij}^t) + m_i g \quad (3)$$

As illustrated in Eq. (3), normal (F_{ij}^n) and tangential interaction forces (F_{ij}^t) affect the velocity (v_i) of particle with mass (m_i) where g corresponds to the gravitational constant ($g = 9.81 \text{ m/s}^2$). The normal and tangential forces are dependent on the contact models chosen in the DEM simulation. In this study, Hertz-Mindlin contact model is selected where the forces are calculated based on Hertzian contact theory [179] and Mindlin-Deresiewicz model [180, 181]. Since this

model provides accurate calculation of particle forces, it is used as default contact model. The equations associated with evaluation of normal and tangential forces are shown below using Eq. (4-7) which are functions of normal (δ_n) and tangential particle overlap (δ_t). Tangential force is evaluated based on the tangential stiffness (S_t) of the particle.

$$F_n = \frac{4}{3} E^* \sqrt{R^*} \delta_n^{\frac{3}{2}} \quad (4)$$

$$F_t = -S_t \delta_t \quad (5)$$

$$E^* = 2G^*(1 + \nu^*) \quad (6)$$

$$S_t = 8G^* \sqrt{R^*} \delta_n \quad (7)$$

In the above equations, equivalent Young's modulus (E^*), particle radius (R^*), shear modulus (G^*) and poisson's ratio (ν^*) are evaluated as a weighted average based on the particles in contact. Along with the translational motion, the rotational motion is governed by radial torque (τ_{ij}^r) exerted on the particle (i) during an interaction with particle (j) as shown in Eq. (8) where moment of inertia (I_i) of the particle (i) is used to evaluate angular velocity (w_i).

$$I_i * \frac{dw_i}{dt} = \sum (R_i * F_{ij}^t - \tau_{ij}^r) \quad (8)$$

These equations are solved for each particle interaction within the DEM simulation at each time step. The time step is selected based on the Rayleigh time step (T_R) as shown in Eq. (9) where multiple particle properties including particle radius (R), particle density (ρ), shear modulus (G)

and Poisson's ratio (ν) govern T_R . The time step (Δt) is recommended to be 20% of Rayleigh time step (T_R) to maintain accuracy of the DEM simulation. If the selected time step is significantly higher than the recommended value, the DEM simulation can miss some interactions and lead to inaccurate results. Hence, the time step chosen for all the DEM simulations performed in this article is set to be 20% of T_R .

$$\Delta t < T_R = \frac{\pi R \sqrt{\rho G}}{0.01631 \nu + 0.8766} \quad (9)$$

Commercial DEM software EDEM[®] (DEM Solutions Inc.) incorporating all the above-mentioned equations (Eqs. 3-9) is used for performing the DEM simulations mentioned in this article.

3.2.2. Material bulk calibration using Drawdown test

Material calibration is a widely studied area in the field of particle-scale simulations, to obtain accurate material parameters corresponding to the actual powder system [130]. Direct and bulk measurements are two approaches used for material calibration. Direct measurements include powder measurements at the particle level. These measurements include static and rolling friction, shear and Young's modulus, which are directly related to the DEM parameters. The drawback for this approach is the need for specialized and expensive experimental setup for particle-level measurements. Authors have demonstrated material calibration using direct measurement tests by replicating FT4 [173, 182-184] and shear cell experiments [171, 185] within DEM framework. Bulk measurements, on the other hand are powder properties measured for a bulk mass of powder. These include bulk density, angle of repose, shear angle which are indirectly related to particle-level measurements. Since these measurements are more generic in comparison to direct measurements, this approach can make use of simple experimental setups like hourglass test, ledge test [132], conical pile [138], sandpile test [82, 140, 174, 175] and drawdown test. Amongst direct

and bulk measurement approaches, bulk measurements are preferred as they are easier, cheaper, and computationally efficient due to the simplified experimental set up. Further, bulk measurements focus on bulk powder interactions which are more prominent in the system to be studied using DEM. For these reasons, we use bulk measurement approach in this study, focusing on calibration using a drawdown test.

Drawdown test is applied in this study for its simplicity and ease of application. This test permits evaluation of multiple bulk measurements using a single test, allowing for a robust calibration [140, 141]. The drawdown test is performed as shown in Figure 3.2, where the drawdown system consists of a frame separated into two compartments with an opening in the middle. Particle factory for particle generation is created within the top compartment as shown in Figure 1. A plate is used to close this opening, which prevents the particles from falling directly into the bottom compartment at the start of the operation. For this study, 50,000 particles of 1 mm are added to the system with dimensions X – 224.4 mm, Y – 51 mm, and Z – 306 mm. The drawdown test is operated as follows. Initially, particles are generated and allowed to settle using the particle factory for 1 second. After particles have settled in the top compartment, the plate blocking the opening is removed and particles fall into the bottom compartment. Particles settle in the bottom compartment and system attains a steady state after 10 seconds. Following this, bulk measurements like mass in lower container, angle of repose, shear angle and maximum flowrate are measured. Powder flowrate is monitored during the entire operation and maximum flowrate is determined based accordingly, whereas other bulk measurements are measured after 10 seconds. DEM simulation runs presented in this article are completed using commercial software (EDEM[®] 2019) utilizing an Intel Xeon[®] E5-2650 v4 2.2 GHz (2 processors) with 128 GBs of RAM. On an average, the computational time required for simulation of drawdown test is 3.5 hours with 6 cores.

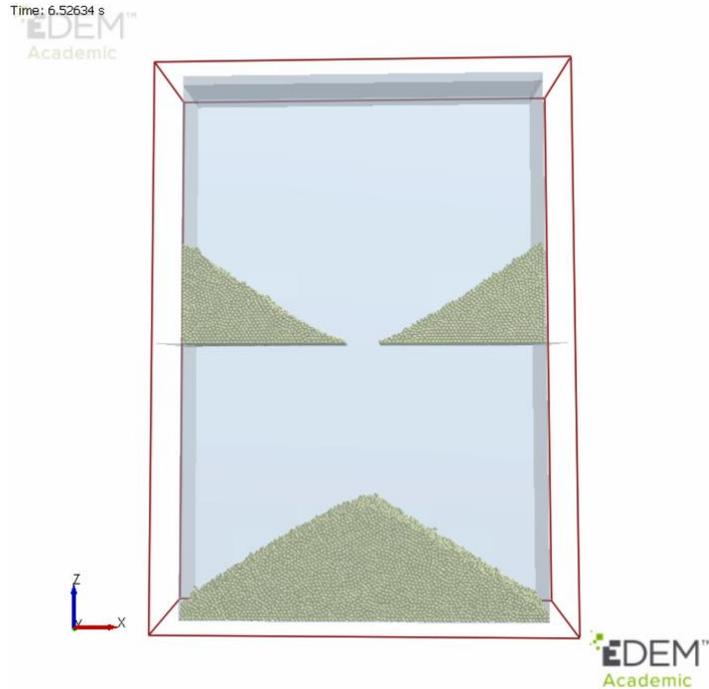


Figure 3.2: Drawdown test for material calibration

Thus, the drawdown test is used to measure 4 bulk properties focusing on real material properties including mass in lower container, angle of repose, shear angle and maximum flowrate. Mass in lower container corresponds to the overall powder mass obtained after steady state is reached. Angle of repose is the angle formed by the powder in the top compartment whereas, shear angle is the angle formed by powder heap in the bottom compartment. Flowrate at the opening is tracked during the entire simulation and maximum flowrate is evaluated. Application of these bulk measurements towards material calibration and addressing the problem of solution multiplicity is highlighted in section 3.1.

3.2.3. DEM simulation of powder feeder

The feeder unit geometry is constructed as shown in Figure 3.3 [186] and imported within the DEM framework. The geometry consists of a virtual plate for particle factory positioned at the top, hopper section, screw section consisting of twin screws and a bridge-breaking impellor to sweep powder

into the screw section. The virtual plate for particle factory is constructed to generate particles within the system. Spherical, cohesion-less particles with standard rolling friction models are selected in this study. Particles of radius 1.0 mm and density of 1500 kg/m^3 [187] are added to the system for 1 second and allowed to settle. In order to obtain a complete operating profile of the feeder unit, 150,000 particles are added to the system. The feeder unit is simulated for 20 seconds within which the feeder runs out of powder. Impellor speed and screw speed are maintained constant at 25 rpm and 100 rpm for all simulation runs. The overall dimensions of the feeder unit are 20.4 cm, 20.4cm and 26.3 cm in X, Y and Z directions. The feeder equipment is assumed to be made of steel [187] and the equipment material properties are assigned accordingly. Though the overall simulation time depends on the material properties selected within DEM and the number of computational cores used, the feeder DEM simulation presented in this article approximately requires 17 hours with 8 computing cores.

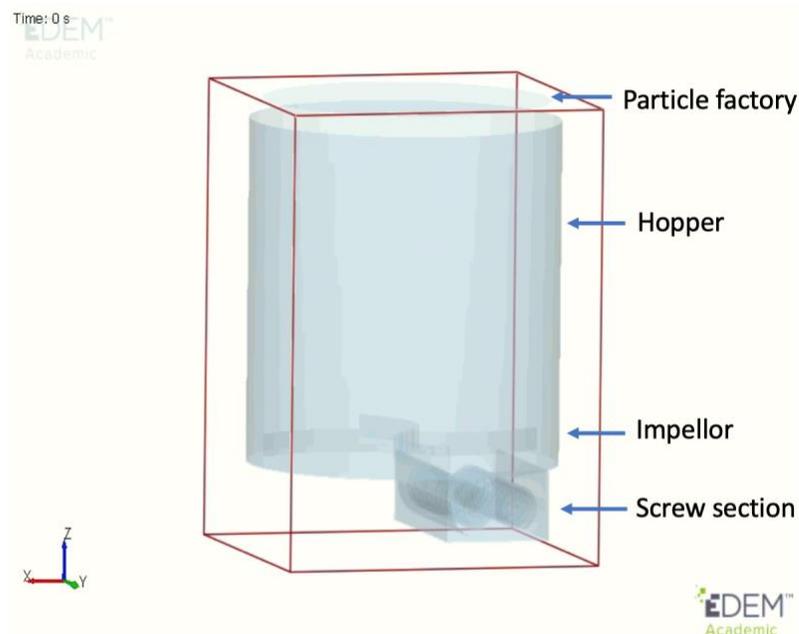


Figure 3.3: Feeder geometry in DEM

3.2.4. Design of experiments (DOE)

To develop material calibration studies and understand effects of material properties on feeder operation, a parametric study is performed. Different material properties are considered for the parametric study including properties affecting particle-particle (PP) and particle-wall (PW) interactions, shown in Table 3.1. We select a three-level parametric design where the lower, middle and upper bounds are selected based on previous literature [49, 168, 175, 187-190].

Table 3.1: DEM material properties for parametric study

Material property	Nomenclature	Bounds			Units
		Lower	Middle	Upper	
Coefficient of Static Friction	PP_SF	0.1	0.3	0.6	-
Coefficient of Rolling Friction	PP_RF	0.01	0.04	0.1	-
Coefficient of restitution	PP_COR	0.1	0.4	0.7	-
Coefficient of Static Friction	PW_SF	0.1	0.3	0.6	-
Coefficient of Rolling Friction	PW_RF	0.01	0.04	0.1	-
Coefficient of restitution	PW_COR	0.01	0.1	0.2	-
Shear Modulus (Pa)	SM	1E6	5E6	1E7	Pa
Poisson's ratio	PR	0.2	0.35	0.5	-

Authors have demonstrated that other material properties like shear modulus (SM) have insignificant effect on powder behavior within the DEM framework [82, 175]. Furthermore, SM being associated with evaluation of Rayleigh time step (Eq. (9)), higher values of SM can lead to huge increments in the overall simulation time. Hence, the parameter bounds considered for SM are [1E6, 1E7] as shown in Table 1 which are more compact than those used in literature [82, 175].

Since the proposed DEM simulations are computationally intensive, appropriate design of experiments (DOE) needs to be chosen to ensure that we can evaluate the main effects of these parameters as well as their interactions while maintaining reduced number of runs. DOE based on Definitive Screening Design (DSD) is chosen, requiring low number of simulations runs for evaluating the following attributes [191-193]:

- a) Main effects independent of interactions.
- b) Two-factor interactions not completely confounded with each other.
- c) Quadratic effects not confounded with other two-factor interactions.

These attributes of resolution 5 designs can be validated using the correlation plot shown in Figure 3.4 where, correlations between main and corresponding two factor interactions are evaluated using the correlation matrix $|\mathcal{r}|$. If two factors are correlated, it is difficult to identify which of those factors affect the outcome. Hence, these designs are chosen as they allow for lower correlation between factors, given the low number of runs.

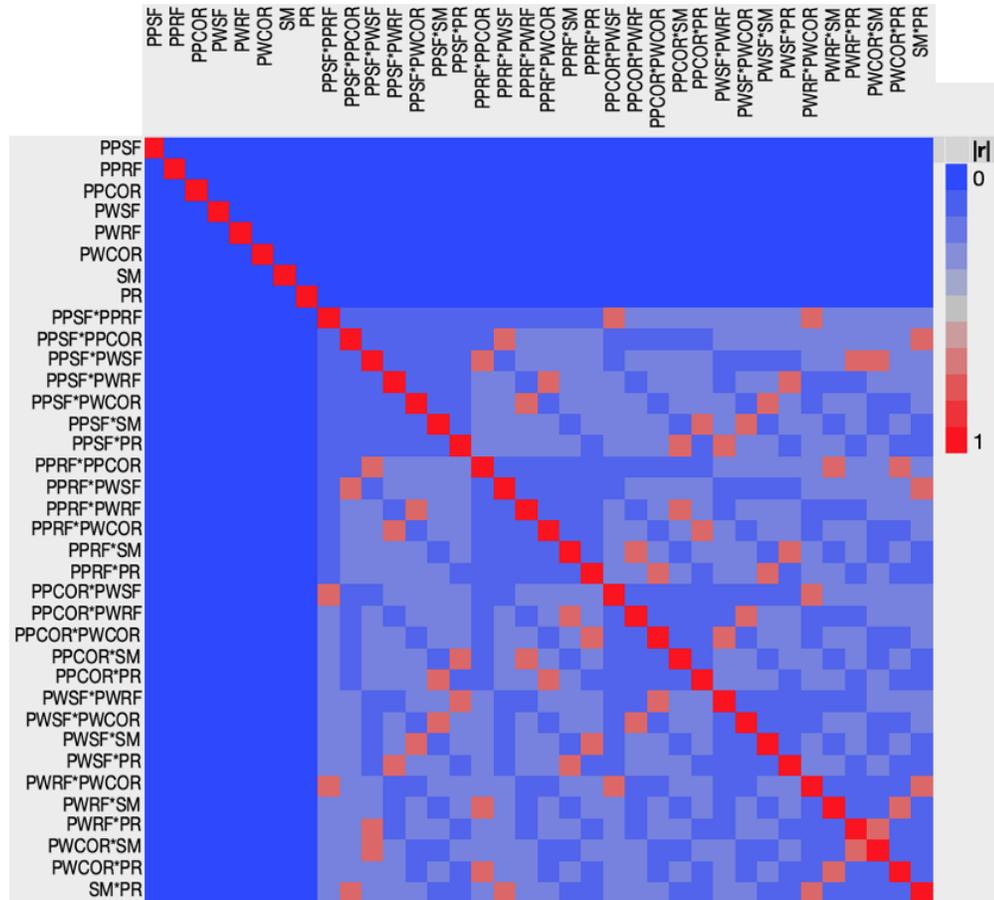


Figure 3.4: Correlation plot for DSD DOE

There are certain limitations of DSD embedded within the design, including the heredity and sparsity assumption. Heredity assumption focuses on the heredity of the observed significant interaction. It states that specific interaction is observed to be important if and only if the associated parents are also important. Sparsity assumption on the other hand, focuses on three or less parameters being identified as important. Since the validity of these assumptions cannot be confirmed at the start of the study, Jones and Nachtsheim, 2017 recommend incorporating 4 additional runs within the DOE [191-193]. Thus, the overall number of runs equals $2^n + 1 + 4$ where, 2^n correspond to the fold-over pairs associated with each parameter (n) in the DOE and one center run. Jones and Nachtsheim, 2017 have proven that this modified design can be applied successfully in the presence of the above limitations [191-193]. Thus, this design is adopted for

the parametric study presented in this article and the randomized DOE consisting of 8 material properties is shown in Table 3.2.

Table 3.2: DSD for DEM material parameters

Run	PP_SF	PP_RF	PP_COR	PW_SF	PW_RF	PW_COR	SM	PR
1	0.6	0.1	0.7	0.1	0.1	0.01	1E+07	0.2
2	0.1	0.1	0.7	0.35	0.01	0.01	1E+06	0.2
3	0.1	0.01	0.7	0.1	0.1	0.01	5E+06	0.5
4	0.35	0.01	0.1	0.1	0.01	0.01	1E+06	0.2
5	0.1	0.1	0.4	0.6	0.01	0.01	1E+07	0.5
6	0.1	0.01	0.1	0.6	0.01	0.2	1E+06	0.5
7	0.6	0.1	0.1	0.6	0.1	0.01	1E+06	0.35
8	0.1	0.1	0.1	0.1	0.04	0.2	1E+07	0.2
9	0.6	0.04	0.1	0.1	0.01	0.01	1E+07	0.5
10	0.6	0.01	0.7	0.6	0.01	0.1	1E+07	0.2
11	0.1	0.01	0.1	0.6	0.1	0.01	1E+07	0.2
12	0.35	0.04	0.4	0.35	0.04	0.1	5E+06	0.35
13	0.6	0.1	0.1	0.6	0.01	0.2	5E+06	0.2
14	0.6	0.01	0.7	0.6	0.04	0.01	1E+06	0.5
15	0.1	0.1	0.1	0.1	0.1	0.1	1E+06	0.5
16	0.6	0.1	0.7	0.1	0.01	0.2	1E+06	0.5
17	0.6	0.01	0.4	0.1	0.1	0.2	1E+06	0.2
18	0.35	0.1	0.7	0.6	0.1	0.2	1E+07	0.5
19	0.6	0.01	0.1	0.35	0.1	0.2	1E+07	0.5
20	0.1	0.04	0.7	0.6	0.1	0.2	1E+06	0.2

21	0.1	0.01	0.7	0.1	0.01	0.2	1E+07	0.35
----	-----	------	-----	-----	------	-----	-------	------

3.3. Results

This section highlights the results obtained for different methods proposed in section 2. Material calibration is performed using the drawdown test and calibration results are obtained based on bulk measurements obtained from DSD DOE of drawdown test. These are highlighted in subsection 3.1. Section 3.2. discusses the results of the feeder simulation with refill studies, followed by DOE of feeder simulation to explore the correlations between the DEM material parameters and feeder operation.

3.3.1. Material calibration

Material calibration efforts using a drawdown test are proposed to identify a unique set of DEM material parameters which represents bulk nature of powder accurately within DEM framework. For this, DSD based runs are performed for drawdown test to obtain bulk measurements. This is followed by model development for bulk measurements. This system of equations is later implemented to address the question of solution multiplicity.

3.3.1.1. Results for DOE of the Drawdown test

The drawdown test is performed using DEM for all the runs in the DOE highlighted in section 2.4. These DEM simulations are processed using image analysis tools to obtain the bulk measurements, shown in Table 3.3. The measurements for angle of repose and shear angle are obtained by averaging the values obtained on both sides of the drawdown test.

Table 3.3: Bulk measurements using DEM drawdown test

Runs	AOR Angle (°)	Shear Angle (°)	Mass in lower container (kg)	Max flowrate (kg/s)
1	31.951	22.3	0.1718	0.1736
2	22.266	23.52	0.2224	0.1842
3	19.495	12.07	0.2649	0.2086
4	15.543	11.99	0.2679	0.1773
5	23.606	23.59	0.2180	0.1865
6	19.362	12.05	0.2750	0.2015
7	20.705	35.79	0.1935	0.1387
8	18.845	11.92	0.2639	0.1961
9	30.998	16.21	0.2476	0.1652
10	28.681	26.77	0.2116	0.1599
11	19.673	19.79	0.2324	0.1954
12	30.800	31.09	0.1854	0.1806
13	35.171	37.37	0.1499	0.1462
14	28.580	27.24	0.1613	0.1626
15	12.914	10.87	0.2680	0.2034
16	32.295	19.42	0.2423	0.1563
17	27.755	13.12	0.2525	0.1784
18	35.071	42.84	0.1303	0.1536
19	28.152	27.16	0.1994	0.1726
20	21.358	24.5	0.2113	0.1839
21	19.510	8.451	0.2829	0.2074

3.3.1.2. Calibration based on the bulk measurements

After evaluating the bulk measurements, we aim to develop predictive models for the bulk measurements. To correctly implement this, we check for correlations with DEM material properties using a two-step regression approach [191-193]. First step involves identification of important main effects. After the important main effects are identified, second step involves identification of important two-factor interaction effects. These steps are followed to ensure all effects are correctly captured without overfitting the data. Identification of important effects is completed by using Akaike information criteria (AIC) [194], Bayesian information criteria (BIC)[195] and p-values. These are used to identify important DEM properties and screen out the effects which are not correlated to the bulk measurements. This two-step process is demonstrated for shear angle as shown below. Initially, important main effects are identified using stepwise regression. Figure 3.5 illustrates the model fit for shear angle measurement using main effects, where particle static friction, rolling friction and particle-wall static friction are identified to be important. These results seem appropriate from a physical point of view as friction parameters – static friction and rolling friction control the powder-wall interaction affecting the formation of shear angle. In second step, the important main effects with the corresponding two-factor interactions are combined for model building. The model fit results including main and interaction effects are shown in Figure 3.6 and the corresponding response surface model (RSM) is shown in Eq (10) with $R^2 = 0.90$. The details of the model fit including parameter estimates and analysis of variance are provided in section 8.1.1.

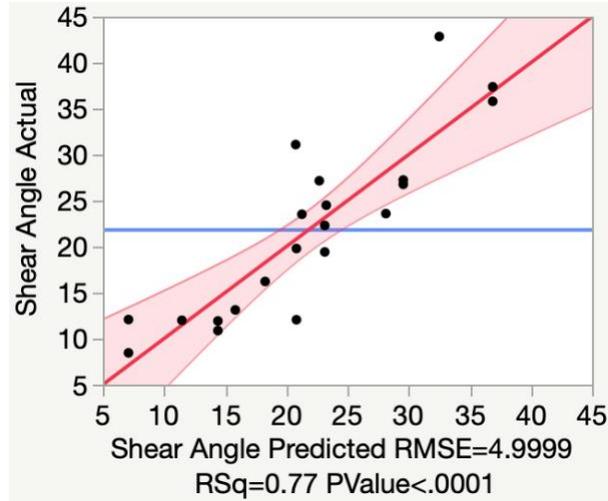


Figure 3.5: Model fit for shear angle using main effects

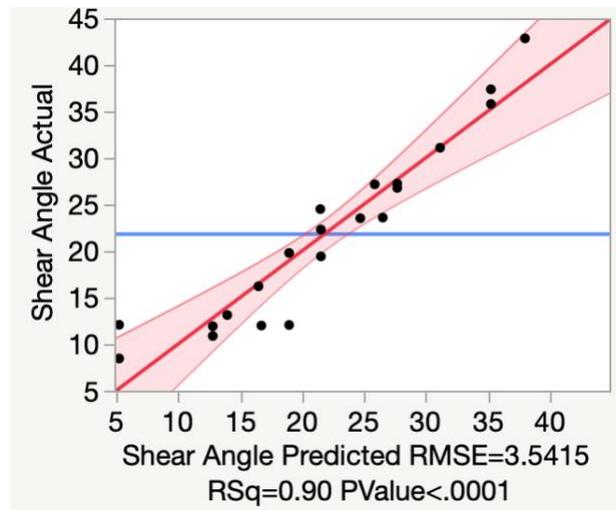


Figure 3.6: Model fit for shear angle using main and interaction effects

$$\begin{aligned} \text{Shear Angle} = & 12.07512 + 17.47 * PP_SF + 84.0688 * PP_RF + 27.4642 \\ & * PW_SF - 113.98 * (PP_SF - 0.35)^2 - 80.78 * (PW_SF - 0.35)^2 \end{aligned} \quad (10)$$

Following a similar methodology for other bulk measurements including the mass in lower container, angle of repose and max flowrate, RSM equations are shown in Eq. (11-13) with an R^2

value of 0.9, 0.72 and 0.92 correspondingly. Given the RSM equations resulted in statistically good model fits, complex models have not been implemented for model construction.

$$\text{Mass in Lower Container} \quad (11)$$

$$\begin{aligned} &= 0.29688 - 0.08876 * PPSF - 0.40432 * PPRF - 0.03561 \\ &* PPCOR - 0.11142 * PWSF - 0.2474 * PWRF + 13.378 \\ &* (PPRF - 0.0538)^2 - 1.787 * (PPRF - 0.05287) \\ &* (PPCOR - 0.4) + 0.257626 * (PPCOR - 0.4)^2 - 0.47643 \\ &* (PWSF - 0.35) * (PPSF - 0.35) - 0.7473 * (PPCOR - 0.4) \\ &* (PWRF - 0.05286) \end{aligned}$$

$$\begin{aligned} AOR &= 16.5772 + 19.39 * PPSF + 7.0079 * PPCOR + 4.7239E^{-7} * SM - 2.20842 \\ &* (SM - 5428571.4286)^2 \end{aligned} \quad (12)$$

$$\begin{aligned} \text{Max Flowrate} &= 0.21327 - 0.0697 * PPSF - 0.1527 * PPRF - 0.03064 * PWSF \\ &+ 0.143184 * (PPSF - 0.35)^2 \end{aligned} \quad (13)$$

This system of equations (Eq. (10-13)) for evaluation of bulk measurements can be combined to identify unique set of DEM parameters for a given set of bulk measurements. The details of the model fits are provided in section 8.1.1.

3.3.2. Simulation of feeder unit

3.3.2.1. Simulation results

The feeder simulation is set up and simulated as illustrated in section 3.2.3. The simulation time is set to 20 seconds which encompasses enough time for a feeder run out simulation. Once the

feeder simulation is completed, a few checks are performed to ensure the validity of the DEM simulation. First check includes particle overlaps not exceeding 4% along the normal and tangential directions [189]. Second check is based on mass balance of particles entering and leaving the feeder system. This is performed to ensure particles are not exiting the system from an unknown location. Once these checks are verified, the simulation results are obtained and processed as described below.

Using data analysis within DEM, the particles exiting the feeder unit are tracked along time to evaluate the feeder flowrate and weight of the feeder unit is measured based on the number of particles present in the unit along time. Following this, the feed factor is evaluated using Eq (1). Feeder weight considers the amount of active powder present in the system, i.e., the amount of powder which participates in the powder feeding. The stagnant powder present in the feeder unit is not accounted for as it does not contribute towards the flowrate of the unit. The stagnant powder corresponds to the powder left in the hopper section after feeder run out simulation. This further ensures that the feeder measurements obtained in this study are similar to those obtained from GEA scale in feeder experiments [44]. After obtaining all these variables, the feed factor vs feeder weight profile can be plotted as shown in Figure 3.7. This plot is used to regress the feed factor parameters as shown in Eq. (2) using SciPy curve-fit in Python.

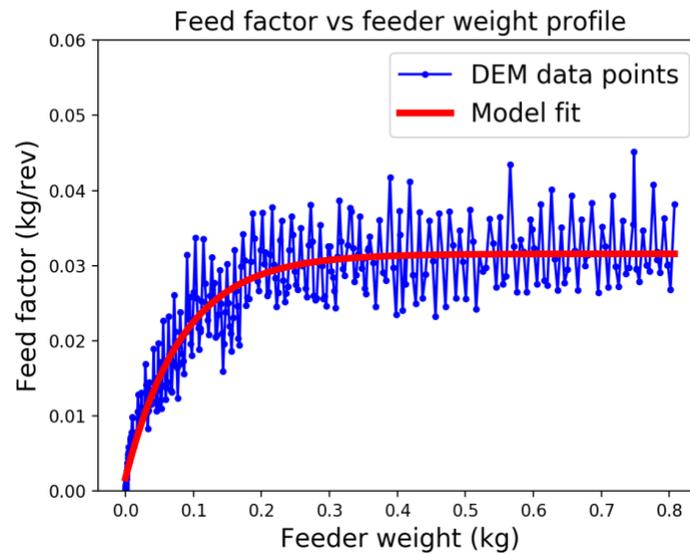


Figure 3.7: Profile of feed factor vs feeder weight

3.3.2.2. Feeder refill studies

DEM simulation of feeder unit mentioned in section 2.3 is modified to include refills and the effects of these refills are examined on the flowrate profiles of the feeder unit. Refills are added by generating two different types of particles are generated within the DEM framework. Particles of type 1 are added during the original fill of the feeder unit, whereas particles of type 2 are added during refill. Different feeder refill runs are simulated where refill particles of type 2 are added to the feeder unit at approximately 20%, 40%, 60% and 80% of fill-level of the hopper attained during the original fill of type 1 particles. These refill studies are performed as a sensitivity study to analyze effects of different refills on powder flow. It should be noted that these studies need to be modified to include more powder to develop a true representative of real powder feeder operation. The proposed aim is to explore powder flow and mixing within the presented studies and understand powder feeding from a conceptual point of view. In the presented study, the total number of particles is balanced such that it is large enough to simulate powder flow dynamics while maintaining the computational costs.

During the feeder DEM operation, when the powder fill-level within the hopper reaches the refill set value, the particle factory adds new particles of type 2 to the system which have the same material properties as the original particles. Figure 3.8 shows these refill runs where effects of refill on the flowrate of the feeder are observed as a function of time.

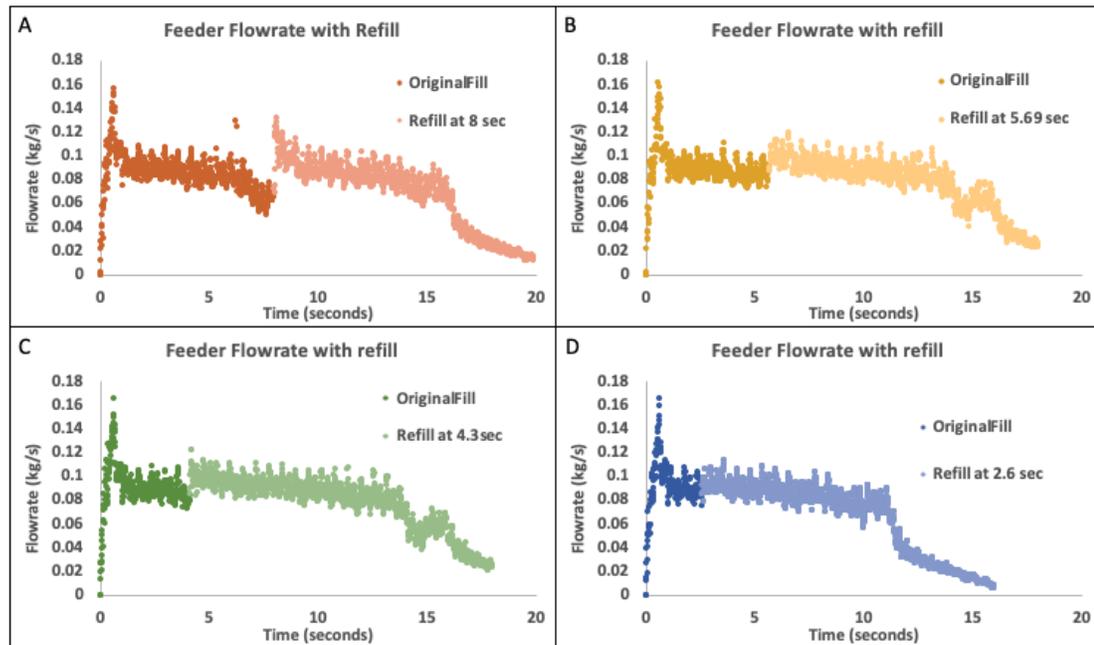


Figure 3.8: Profile of flowrate vs time with feeder refills

The refill at 8 seconds corresponding to 20% fill level is shown in Figure 3.8(A), 5.69 seconds corresponding to 40% shown in Figure 3.8(B), 4.13 seconds corresponding to 60% shown in Figure 3.8(C) and 2.6 seconds correspond to 80% shown in Figure 3.8(D). The fill-level corresponds to the fill-level attained by the original fill of particles rather than the complete hopper height. This is performed to explore refill situations whilst keeping the computational cost low. The slight differences in different profiles shown in Figure 3.8, can be accounted to the time of refill and the number of particles added during the refill. For 2.6 second refill run, 40,000 particles are added whereas for other runs, 100,000 particles are added to the system. Higher the number of particles added to the system, wider is the observed flowrate profile.

Following this, the modified DEM simulations of feeder units with refills are used to explore particle mixing within the unit. The effects of powder mixing are observed from the feeder flowrate profiles. This is an important aspect as feeders are meant for dispensing powders, not mixing. If mixing is identified within the unit, it needs to be taken into consideration through mixing models or residence time distribution models for accurate material tracking and lot-to-lot delineation. For mixing evaluation, two types of particles are created within the DEM framework and particles of type 1 are added during the initial fill of the feeder unit (indicated by yellow particles), whereas particles of type 2 are added during the refill (indicated by orange particles) as shown in Figure 3.9. These particles only differ in color to help visualize and evaluate mixing. This can further be illustrated as shown in Figure 3.10 where the total number of particles of type 1 and 2 are plotted along time axis with a refill of type 2 particles at 4.13 seconds.

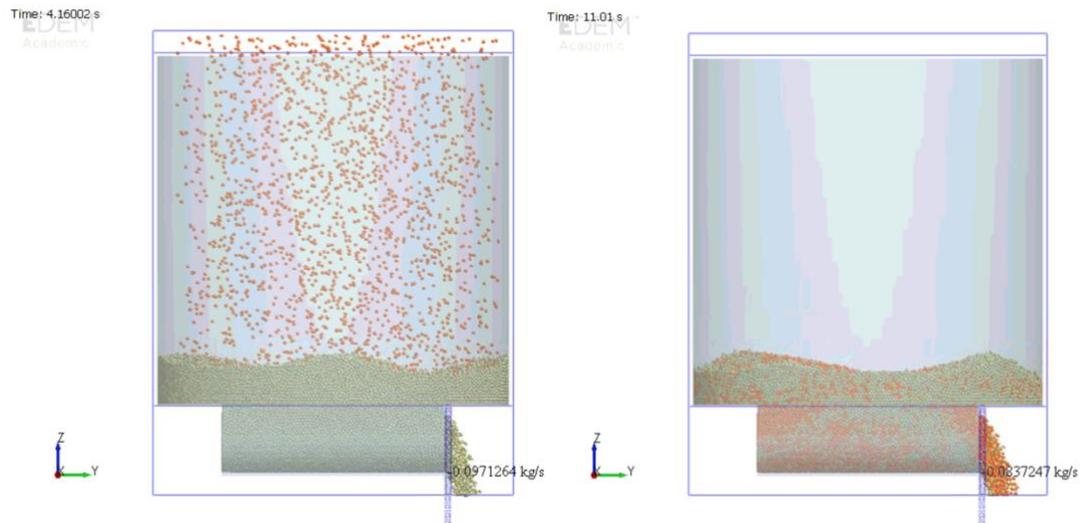


Figure 3.9: Image of the left shows refill particles of type 2 (highlighted as orange particles) being added to the feeder unit at 4.13 sec whereas image on the right indicates particles mixing during feeder operation

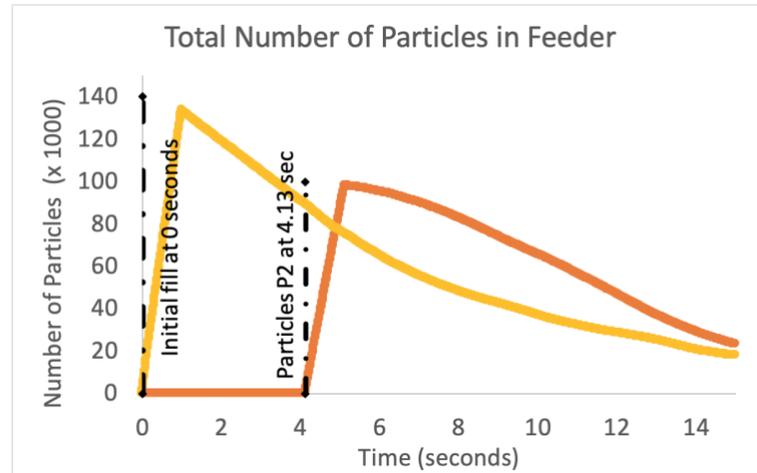


Figure 3.10: Total number of particles present in the feeder unit. Particles added at time = 0 seconds indicated in yellow whereas, particles added during refill at time = 4.13 seconds are indicated in orange.

Flowrate of the feeder unit is measured near the feeder outlet using a rectangular grid as shown in Figure 3.7 which measures the number of particles exiting the unit as a function of time. To evaluate particle mixing within the unit, fraction of each particle type observed at the feeder outlet is measured as a function of time. This is shown in Figure 3.11, where initially, fraction of particles of type 1 exiting the unit correspond to 1. This eventually reduces with addition of particles of type 2. After addition of type 2 particles, it is observed that both types of particles exit the feeder unit, instead of particle 1 exiting earlier.

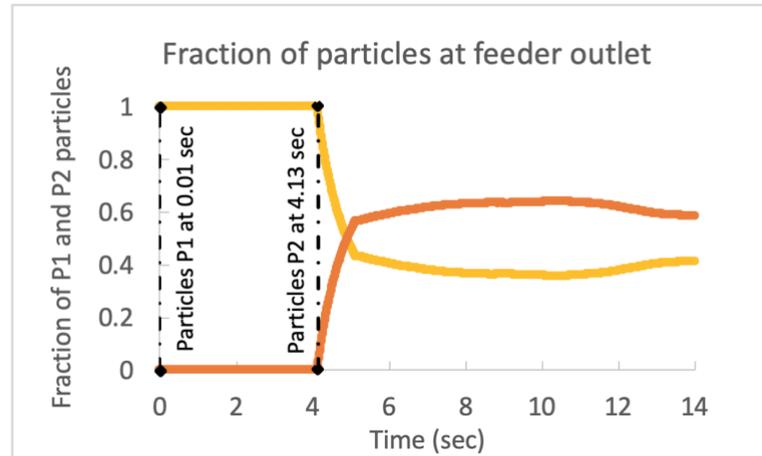


Figure 3.11: Fraction of particles at the feeder outlet

This illustrates the two types of particles getting mixed during the feeder operation and a mixture of particles being observed at the outlet. Though these studies are useful in identifying mixing within the DEM simulation, it is important to note that the experimental set up for feeder refills can show slightly different profiles than the DEM simulation. Experimental set up for feeder refills includes powder filling up majority of the hopper section and powder refills are initiated when powder level present in the hopper reduces below a set fill-level limit. This limit is based on the total hopper height of the feeder unit. As hopper section acts more like a plug flow, the experimentally observed mixing profiles of refill particles exhibit more delayed response. This is not included within the DEM simulation study presented here, as adding enough powder to fill up the hopper would immensely increase the computational time required for feeder simulations. Thus, the DEM feeder simulations presented here are portrayed as conceptual studies to help identify mixing within the feeder unit. Such mixing studies can further be combined with a plug flow regime and used to develop residence time distribution or mixing models for accurate material tracking and lot-to-lot delineation of the feeder unit.

3.3.2.3. Results for DOE of feeder simulation

Following the evaluation of feed factor vs feeder weight plot based on DEM simulations illustrated in section 3.2.1, this section focuses on a parametric study of DEM material parameters to explore their correlations with regressed feed factor parameters. Definitive screening designs (DSDs) is applied for the parametric study.

DEM simulations are performed based on the DSD shown in Table 3.2 and simulation runs are analyzed as illustrated in sub-section 3.2.1. Based on the feed factor vs feeder weight profiles, feed factor parameters are regressed for all the parametric runs (shown in section 8.1.2). The physical understanding associated with each feed factor parameter [44, 186] is outlined as follows, where FF_{max} is the maximum value of feed factor observed in the feed factor vs feeder weight plot. This value depends on the feeder hopper geometry and the type of screws selected. FF_{min} indicates the lowest value of feed factor obtained when the feeder weight is extrapolated to zero. Since this parameter is not observed during the feeder operation, there is a higher likelihood of uncertainty associated with evaluation of FF_{min} . Lastly, β depends on the rate of feed factor decay as shown feed factor vs feeder weight profiles. These parameters are obtained using curve-fit in python. To provide a comprehensive idea of the goodness of the fit, two simulation runs with the highest R^2 value (0.978 for Run 4) and lowest R^2 value (0.769 for Run 7) is shown in Figure 3.12.

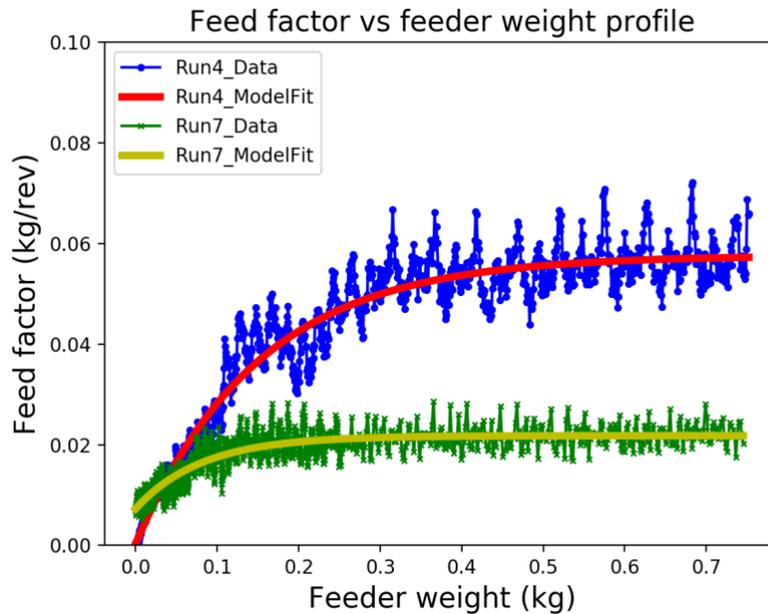


Figure 3.12: Feed factor vs feeder weight curve-fit for the feed factor parameter

3.3.2.4. Developing correlations based on the DOE

After obtaining the regressed feed factor parameters for all runs in the DSDs, this section focuses on the identification and evaluation of correlations between DEM material and feed factor parameters (FF_{max} , FF_{min} and β). This is explored using stepwise regression. Initially, main effects are identified using stepwise regression combined with Akaike information criteria (AIC) and Bayesian information criteria (BIC). Following this, important two factor interactions are also evaluated using least square regression. A response surface model is created for each feed factor parameter including important main and two-factor interaction effects. This analysis allows for a thorough assessment of correlations between DEM material and feed factor parameters. JMP® (SAS) software (SAS Institute Inc.) is used to perform the above analyses for all feed factor parameters.

Model results of least-square regression for all feed factor parameters - FF_{max} , FF_{min} and β are shown in the section 8.1.2, including an actual vs predicted plot with important main and interaction effects. The goodness of the obtained plots is verified by the R^2 values and ANOVA test, wherein a probability < 0.0001 indicates that the models are significant for all feed factor parameters. It is observed from Figure 8.6 that particle static friction (PP_SF), rolling friction (PP_RF) and particle wall static friction (PW_SF) are significant for FF_{max} . This makes physical sense because FF_{max} is observed when the maximum amount of powder is present in the feeder hopper. During this time, there is a significantly higher likelihood of bulk powder particles interacting with each other and the walls causing the static friction and rolling friction to mainly influence the particle flow within the feeder unit. Similarly, Figure 8.7 illustrates particle-wall static friction (PW_SF) and particle-particle static friction (PP_SF) to be important for FF_{min} . Evaluation of FF_{min} corresponds to the time when the feeder is almost out of powder. During this time, majority of powder present in the feeder exits the unit and remainder of powder is left in the screw section. Powder particles present in the screw section interact with the walls of the screw and other particles causing the above parameters to appear significant for FF_{min} . For coefficient β , it is observed that particle static friction (PP_SF), particle-wall static friction (PW_SF), particle coefficient of restitution (PP_COR), particle rolling friction (PP_RF) and particle coefficient of restitution (PP_COR) are significant as observed from Figure 8.8. As β relates to the profile of feed factor decay with respect to feeder weight, it depends on the curvature of the feed factor vs feeder weight profile. Correspondingly, it would depend on all the DEM material properties identified to be important for FF_{max} and FF_{min} . Further, during the feeder operation, powder particles are pushed by the impellor into the screw section, causing the coefficient of restitution to affect the β parameter, as corroborated from the results observed in Figure 8.8. Furthermore, a close insight into the relationships between individual DEM material properties and feed factor parameters can be obtained using prediction profile plots as shown in Figure 8.9-Figure 8.11

Following the regression analyses of all feed factor parameters, response surface models are constructed as shown below (Eqs. 14-16) to evaluate feed factor parameters for different values of DEM material properties.

$$\begin{aligned}
 FF_{max} = & 0.06856 - 0.050247 * PP_SF - 0.131 * PP_RF - 0.0303 * & (14) \\
 & PW_SF + 0.1176 * (PP_SF - 0.35)^2 - 0.1723 * \\
 & (PP_SF - 0.35) * (PP_RF - 0.05287) + 0.1182 * \\
 & (PW_SF - 0.35)^2
 \end{aligned}$$

$$\begin{aligned}
 FF_{min} = & -0.0033 + 0.00749 * PP_SF + 0.00792 * PW_SF + 0.019 & (15) \\
 & * (PW_SF - 0.3456) * (PP_SF - 0.3456) + 0.02398 \\
 & * (PW_SF - 0.3456)
 \end{aligned}$$

$$\begin{aligned}
 \beta = & 0.595 + 9.735 * PP_SF + 26.79 * PP_RF + 1.717 * PP_COR + 3.37 & (16) \\
 & * PW_SF - 1297.39 * (PP_RF - 0.0526)^2 + 113.7 \\
 & * (PP_COR - 0.3978) * (PP_RF - 0.0526) + 26.425 \\
 & * (PW_SF - 0.3456) * (PP_SF - 0.3456) + 51.215 \\
 & * (PW_SF - 0.345)^2
 \end{aligned}$$

Based on this analysis, it can be confirmed that the particle static friction (PP_SF) and particle-wall static friction (PW_SF) are significant for all feed factor parameters and overall feeder operation. Particle rolling friction (PP_RF) is observed to be important for FF_{max} corresponding to the initial phases of feeder operation when hoper is filled with powder whereas coefficient of restitution appears significant when powder falls from the hopper to the screw section due to impellor

rotations. Thus, it can be concluded that only a small subset of DEM material properties is important for the feeder operation. This finding is significantly important and helpful in concentrating material calibration efforts to few DEM material parameters which are important for the system in consideration. Other DEM parameters like shear modulus are confirmed to be insignificant as identified earlier by (Yan et al., 2015) for a powder discharge system. The proposed work assumes negligible backpressure for in-flowing powder from material already residing in the screws. Additional feeder simulations with longer screw section can be performed to confirm these results.

3.4. Discussion

This section focuses on application of the results obtained for material calibration for identification of unique set of DEM material properties for a given set of measurements based on bulk powder behavior, highlighted in 3.4.1. Section 3.4.2 provides verification studies for the results obtained in section 3.3.2 for predictive models developed for feed factor parameter using feeder DEM simulations.

3.4.1. Material calibration results

This section highlights application of material calibration, wherein, we can estimate the DEM material parameters for a given set of bulk measurements. Bulk measurements obtained from drawdown test have been identified to depend on 6 DEM material parameters. To illustrate the proposed application of material calibration, authors focus on estimating 4 DEM material parameters based on 4 bulk measurements from drawdown test. The DEM material parameters to be estimated are particle static friction (PP_SF), rolling friction (PP_RF), coefficient of restitution (PP_COR) and particle wall static friction (PW_SF). These parameters are chosen as they are comparatively difficult to obtain from direct measurements as compared to other parameters including shear modulus (SM) and particle coefficient of restitution (PP_COR). Thus, the other

parameters are obtained from the literature [187]. The predictive models for bulk measurements (Eq. 10-13) are used to estimate DEM material parameters for a chosen set of bulk measurements, as shown in Table 4 to illustrate material calibration. This system of equations is solved by minimizing the sum of squared errors using GAMS/BARON [196] with an objective function value equal to 4.2E-14. The obtained values of DEM material parameters are also shown in Table 3.4.

Table 3.4: Estimation of DEM material parameters for a given set of bulk measurements

Bulk measurements	Values	DEM material parameters	DEM Values
Angle of repose	30°	Particle static friction	0.4931
Mass in lower container	0.22 kg	Particle rolling friction	0.0383
Shear angle	25°	Particle coefficient of restitution	0.1851
Max flowrate	0.17 kg/s	Particle wall static friction	0.195

To obtain a visual understanding of the relationship of the bulk measurements with DEM parameters, contour plots are shown for the bulk measurements on static friction vs. rolling friction space, shown in Figure 3.13. These plots can be used to identify allowable values of static and rolling friction for specified values of bulk measurements. These further highlight the problem of unique solution in material calibration where multiple sets of DEM parameters can correspond to the same value of bulk measurements. The authors address this problem by developing a well-determined system of equations where the number of DEM material parameters equals the number of bulk measurements. This can be further improved by inclusion of all DEM material parameters with additional bulk measurements.

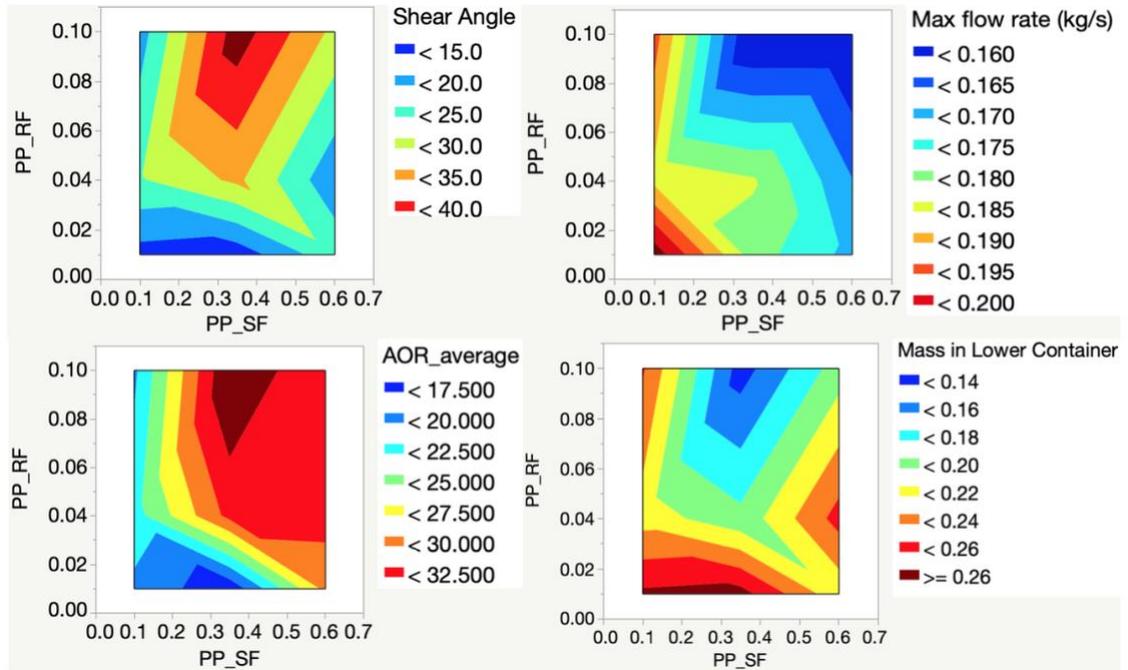


Figure 3.13: Contour plots of bulk measurements obtained from the drawdown test

3.4.2. Verification studies for feeder simulation

Following the predictive models developed for the feed factor parameters in section 3.3.2.4, verification runs are performed at different settings to corroborate the above models and to assess the lack of fit of the developed empirical models for feed factor parameters. These settings are selected such that they are different from those included in model construction. The model verification runs with parameter sets are shown in Table 3.5. The parameter values are chosen such that they lie within the parametric bounds of DEM parameters highlighted in Table 3.2 and do not match the ones selected earlier in the DOE. This way, the verification runs can be used to assess the developed feed factor models.

Table 3.5: DEM parameter settings for model verification runs

Parameters	1 st verification run	2 nd verification run
PP_SF	0.2	0.4

PP_RF	0.03	0.07
PP_COR	0.25	0.5
PW_SF	0.2	0.4
PW_RF	0.03	0.07
PW_COR	0.07	0.15
PR (Pa)	3E6	8E6
SM	0.3	0.4

Feeder DEM simulations and analyses are performed in a similar manner as illustrated in section 2.3. Regressed feed factor parameters are added to the overall datasets and previously developed predictive models are updated. The parity plots as shown below in Figure 3.14-1.16 where the 1st verification run is marked by circled orange ‘ Δ ’ and 2nd by circled green ‘ ∇ ’.

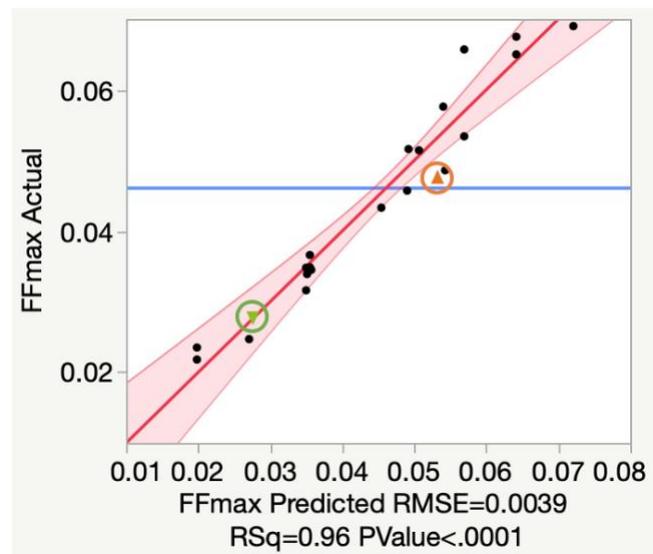


Figure 3.14: Parity plot for FF_{max} including verification runs

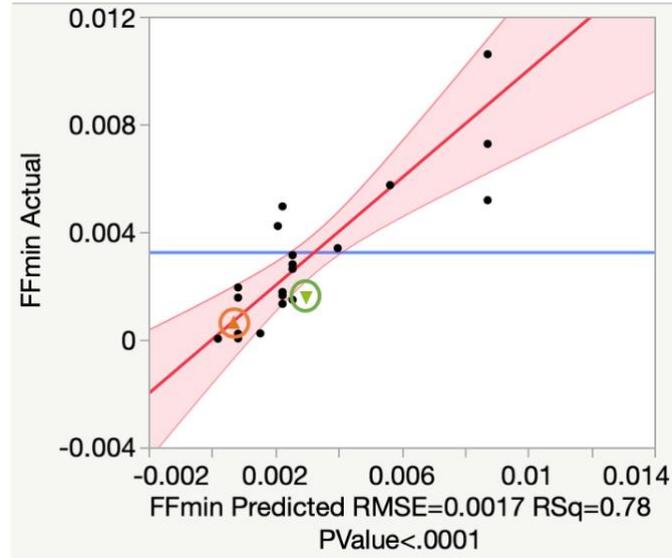


Figure 3.15: Parity plot for FF_{min} including verification runs

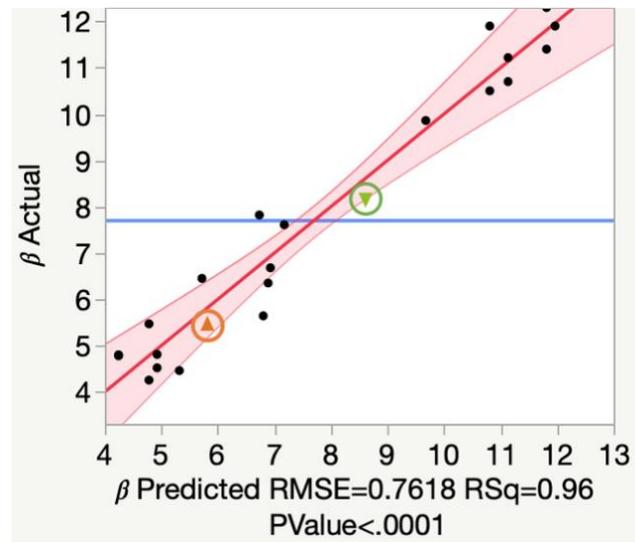


Figure 3.16: Parity plot for β including verification runs

Parity plots highlighted in Figure 3.14-1.16 show good model fits after incorporating the new verification points based on goodness of model fits and R^2 values. Slight differences observed in Figure 3.14 (Run 1- highlighted by orange) and Figure 3.15 (Run 2 – highlighted by green) between predicted and actual values for verification studies, can be attributed to the inherent variability in

the data used for model construction. Thus, the above verification runs confirm the predictability of feed factor models developed in section 4.2.2.

3.5. Conclusions

This article presents an integrated approach of material calibration and DEM simulation to represent bulk powder behavior within DEM framework for detailed process understanding and assessment of effects of DEM material properties on feeder operation. Material property related inputs of DEM simulations cannot be easily obtained as it involves measurement of particle scale properties. This is especially challenging and expensive for pharmaceutical powders given the small particle size. To address this issue, material calibration is implemented using a bulk calibration approach to evaluate the DEM material parameters which are provided as inputs to the DEM simulations. By developing a correlation between the DEM material parameters and experimentally measurable bulk properties, bulk calibration is used to obtain the DEM material parameters which best represent the true bulk nature of powder behavior. For bulk calibration approach, a drawdown test is simulated within DEM, followed by bulk measurements including angle of repose, shear angle, mass in lower container and maximum flowrate. Definitive screening designs (DSDs) are chosen for the parametric study needed for bulk calibration. After evaluating the DEM simulations runs in the parametric study, the observed bulk measurements are checked for correlations with the DEM material parameters and a response surface model is created for each bulk measurement from the drawdown test. It was identified that only few DEM material parameters are important for bulk measurements. The system of equations, thus developed, can be applied to identify DEM material parameters for a new powder material based on the bulk measurements observed in the drawdown experiment which are then, supplied as inputs to the DEM simulation of the process under study.

After material calibration, DEM simulations of feeder unit are performed for detailed process understanding, focusing on powder flow, and mixing within the unit. The flowrate and feed factor profiles obtained from DEM simulations, are used to explore particle movement for different refill and mixing studies. For powder mixing studies, two different types of particles are added during original and refill stages of DEM simulation, monitoring the fraction of each type of particles exiting the unit as a function of time. This demonstrates the mixing within the feeder unit as a mixture of particles is observed at the outlet. Since the presented study focuses on tank section mixing based on the level of powder present in the hopper, it can be integrated with hopper mixing models. to characterize the overall mixing levels within the feeder unit for accurate material tracking or lot-to-lot delineation.

DEM simulations of feeder unit are also applied to understand the effect of material properties on feeder operation. For this study, feed factor parameters are regressed from the feed factor vs feeder weight profiles. The simulation results obtained from the DSD parametric study are used to explore correlations between DEM material parameters and regressed feed factor parameters - FF_{max} , FF_{min} and β . These studies reveal that only a subset of DEM material properties is important for feed factor parameters, also observed for material calibration. This is an important finding as it allows to narrow down the list of material properties important for powder feeding and allows to focus on the reduced material subspace thus, accelerating the overall material calibration process.

In conclusion, this article focuses on material calibration and DEM simulation of feeder unit for enhancing process understanding and analyzing effects of DEM material properties on process operation. Material calibration efforts currently focus on a parametric study for reduction of total number of runs. This can be made more robust and accurate by performing space-filling design like Latin hypercube sampling (LHS). The material calibration process can also be automated (Orefice

and Khinast, 2019) for evaluation of bulk measurements from DEM simulations to minimize user efforts, followed by experimental validations. The feeder DEM studies mentioned in this article can further be improved by addition of larger number of particles with reduced particle size with experimental validation of feeder operation. This would allow the simulation to move closer to the actual pharmaceutical process where particle size is in microns. However, the smaller the particle size, the larger the number of particles required to simulate the process and hence, the larger the computational requirement associated with such DEM simulations. To incorporate the increased computational requirement, parallel computing GPUs can also be implemented within DEM framework. Integral techniques like smoothed particle hydrodynamics would also be helpful towards building a continuum field replicating powder behavior with low particle sizes [56, 95]. The contact models accounting for powder cohesion and non-sphericity of particles with particle size distributions can also be added to depict a wider range of pharmaceutical powders within the DEM framework applied for feeder unit.

CHAPTER IV

4. Aim 2: Process modeling of unit operations using DEM – identifying mixing indices for continuous blending operations

Acknowledgement of publication status:

Entire sections of this chapter have been published in a scientific journal written by the author of this thesis with the title: A review of existing indices in solid-based continuous blending operations. This article was accepted in the Powder Technology on 13th June 2020.

4.1. Introduction

U.S. FDA's QbD initiative [1] recommends a systematic approach towards pharmaceutical development and manufacturing based on risk-management. It focuses on continuous improvement through detailed understanding of effects of process design and operation on overall process performance. Thus, CM operation aims for process improvement through steady increment in process understanding [2]. All the unit operations contained within these lines are studied using the same philosophy, including blender units. Blending has been widely studied [3] and though there exists vast literature [4-6] focusing on blending or powder mixing, there still exists a huge gap in process understanding and its implementation towards process improvement. Thus, there exists a need for detailed understanding and evaluation of powder mixing within these units. Another reason which supports the above claim, is the focus of CM lines on maintaining content uniformity of tablet products. Content uniformity is measured at the end of the CM lines and depends on the blend uniformity evaluated at the blender outlet [7]. Blend uniformity refers to the homogeneity of

powder blend in a tablet manufacturing process [8]. If blend uniformity is identified to be off-limits with coefficient of variation higher than 5% as recommended [9], the operators can devise rectification measures before the powder blend exits the line and off-spec product tablets are formed. These rectification measures can be implemented in real-time to correct the uniformity of the powder blend and correspondingly the content uniformity of tablets, such that the end-product adheres to the regulatory requirements.

Blend homogeneity needs to be measured in radial and axial direction. Uniformity along the radial direction refers to the cross-sectional blend homogeneity whereas, uniformity along the axial direction refers to blend homogeneity along the length of the blender. Authors have successfully developed and illustrated experimental approaches to quantify axial mixing through the implementation of residence time distribution (RTD) [10,11] models. RTD models are mainly used to approximate the dampening coefficient of blender units given upstream variability in inlet flowrates. This is observed in actual CM lines where feeder flowrate variability is transferred to the blender units, and blender reduces this variability before transmitting it to downstream unit operations. Radial mixing on the other hand, has been evaluated for batch blender [12] units through slicing of the powder bed and calculated positioning of sampling locations [13]. However, for continuous blender units, this is cumbersome to evaluate, due to the difficulty and inaccessibility of sampling locations along the radial directions in continuous blender units. Alternative to the sampling approach are other methods like image analysis and near-infrared spectroscopy (NIR). Image analysis focuses on use of Positron emission particle tracking (PEPT) techniques [14] for non-invasive analysis of blend homogeneity. Such techniques require specialized equipment set up, making them quite expensive for process design and understanding. NIR techniques focus on placement of NIR probe [15] within the blender or at the outlet to measure mixing [7,16]. The correct placement of the probe becomes challenging as it needs to measure the true homogeneity without running into issues like probe coating [17] and powder inaccessibility.

Computational approaches can be used for analysis of radial mixing within the blender through the implementation of Discrete Element modeling (DEM). Gao et al. [18,19] focused on this idea to simulate blender dynamics within DEM platform with periodic section modeling (PSM) to analyze radial mixing levels as function of process design and operating conditions. PSM approach results in reduction of overall computational costs by simulation of blender section instead of complete blender. This is used for radial mixing through implementation of relative standard deviation (RSD) index [20,21]. Though radial mixing seems comparably easier to measure within DEM simulations, the challenge lies in appropriate selection of measurement metric for mixing evaluation. The main reason underlying this, is the vast plethora of proposed mixing indices available in literature. Computational metrics for mixing evaluations has been proposed since early 1950s, including variance reduction in 1953 [10] and Lacey's index in 1973 [22]. Since then, there have been many derivations and modifications of these indices, resulting in new indices, including Kramer's index [23], non-sampling mixing index [24], Lacey, Weidenbaum, and Bonilla index (LWB index) [23], along with development of new indices like Average height, Nearest neighbors, Neighbor distance [4], Particle scale [25], Segregation index [26], Mixing entropy [27], Degree of mixedness [28], Variance amongst bimodal bin counts [29], Variance reduction ratio [10], mixing segregation index, Index of Segregation [30], Co-ordination number [31], Siiria index [32] and more. Fan et al. [4] attempted at providing a thorough list of some of these indices, applied to various blending studies conducted independently. Due to this vast availability of indices for mixing, it becomes challenging to identify appropriate mixing index for specific applications. Furthermore, due to the application of different indices, it becomes strenuous to utilize the available literature using different metrics for measurement of powder mixing towards enhancement of process understanding of the blender unit. Wen et al. and Deen et al. [33,34] have previously attempted to compare these indices with their applications on different blending operations like

fluidized bed, blender and spouted bed. However, they only consider a small subset of the exhaustive list of indices mentioned above.

This chapter thus aims to address the issue of overabundance of mixing indices, by providing a detailed review, classification, and comparison of these indices. To highlight the variations between these indices, specifically focusing on methods used for mixing evaluation, a PSM section of blender is simulated using DEM based on previous literature [18]. Mixing indices are implemented for evaluation of mixing as a function of time for qualitative comparison. For quantitative comparison, a T_{95} mixing time is calculated for each index [34] which illustrates the time required to reach 95% of total mixing within the system. Effects of blender operating parameter – blade speed on the mixing times are also assessed for different indices. The rest of the chapter is organized as follows. Section 2 details the DEM methodology, followed by PSM blender simulation details. This section also addresses data extraction strategies for the DEM simulation. After defining the system of study, section 3 focuses on the review of existing mixing indices, enlisting respective details and equations used for mixing evaluation. It further includes classification and comparison of these indices for providing an overview picture of similarities and differences between these indices. Section 4 focuses on the application of the mixing indices on the PSM blender simulation wherein, the effects of operating parameter - blade speed are evaluated using the T_{95} mixing times of the blender system. This section further highlights the considerations required for appropriate selection of mixing index based on the powder system under study. Lastly, section 5 provides a summary and conclusion with additional work that will be pursued in the future.

4.2. DEM simulation of blender unit

In this work, EDEM[®] 2019 commercial software is used for DEM simulations. Initially, Gericke convective continuous blender unit geometry is created within EDEM. AUTOCAD geometry can either be imported or created within EDEM. The blender geometry is created as shown in Gao et al. [41]. The unit dimensions are selected according to the standard manufacturing size. Dynamic parts of the geometry including mixing blades are identified, and the rotational or translational velocities are defined accordingly. Additional virtual plates can be added near the inlet of the blender to represent particle factories for particle generation. Within EDEM, particles with user-defined properties are generated using the virtual particle factories. Particles representing cohesionless, mono-disperse spheres are generated for the simulations mentioned in this article. After particle generation, particle contacts are detected using Hertz-Mindlin [42,43] contact model. Details of the numerical equations involved in DEM simulations are provided in the EDEM documentation [44] and implemented by other articles [36,37]. Due to the large size of the unit, huge computational costs are associated with the overall blender simulation along with the inherently complex DEM simulations. Boukouvala et al. [20] reports the overall simulation time of 3-7 weeks for a single blender simulation for the same specifications. To avoid such high computation costs, a periodic section modeling (PSM) approach is implemented for blender modeling [18]. Periodic section modeling has been successfully implemented for blender systems [20,21] and shown in Figure 4.1. Using this approach, a section of blender is simulated using periodic boundaries instead of the complete blender. The periodic section is selected as a central section of the blender instead of the entry or exit sections which allows it to accurately represent majority of the blender section.

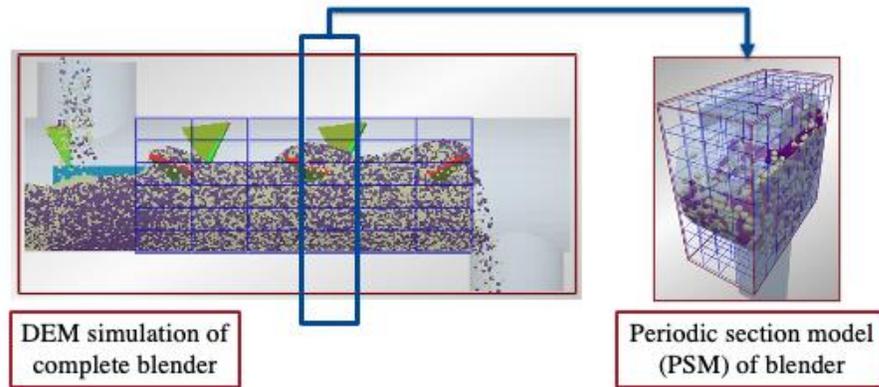


Figure 4.1: DEM simulation and PSM of blender (Printed from Gao et al. [41] with permission from authors)

Periodic boundaries are assigned on both sides of the periodic section where, if particles exit from the left due to particle dynamics, they re-appear on the right and continue interacting with the system dynamics. PSM can be extended for obtaining complete blender dynamics with the time scale in PSM translating to the length of blender axis. This correlates to the powder mixing along time in PSM to the powder mixing in the complete blender along the axial direction. Thus, periodic section simulated for the residence time of the blender, can approximately represent powder system at the exit of the complete blender simulation.

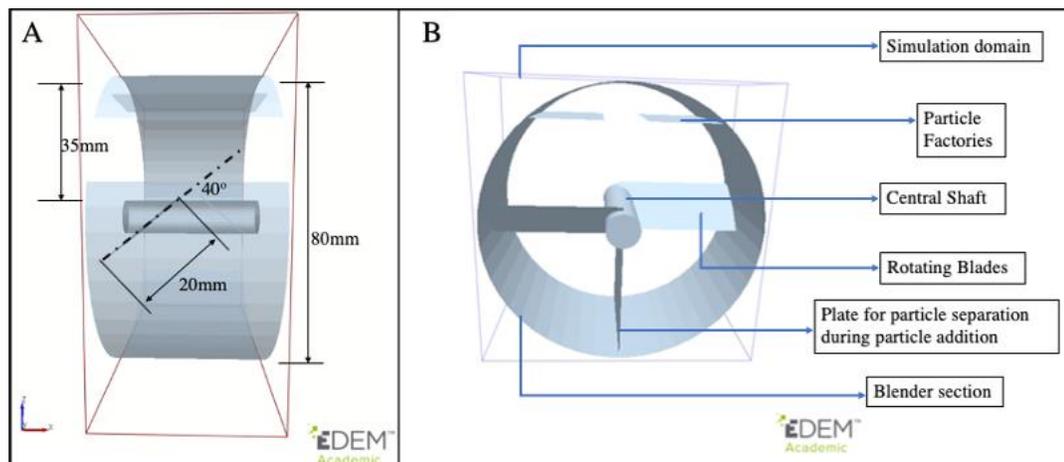


Figure 4.2: PSM geometry for blender simulation - A) PSM dimensions, B) Components of PSM

PSM for blender, is constructed and implemented as shown in Figure 4.2. The PSM geometry dimensions are defined as 40x80x80 mm in X, Y and Z directions where, particles of size 1.5 mm are added. 3,000 particles are added for each particle type, resulting in the total number of particles to be 6,000 within the periodic section. DEM material properties for these particles are chosen based on Gao et. al [41]. The number of particles is chosen such that it achieves a fill-level of approximately 50% within the blender. The geometry consists of the blender section with rotating blades attached to the central shaft. Two types of particles are added to the system using separate particle factories located at the top section. These particles are defined to have same properties, differing only in color (red and blue) to assist in mixing evaluation. A plate is added near the bottom of the unit to separate the two types of particles during particle addition. This ensures the measurement of mixing starting from a completely segregated state. This plate is removed after particle addition, and the central shaft along with the blades is rotated at a specified blade speed to simulate blender motion. Particles get mixed within the periodic section, following which, the blender dynamics are stopped after a mixing time of 20 seconds. The mixing time is determined based on the rate of mixing when it reaches a steady value, as observed from the mixing evaluations highlighted in this article. The PSM blender simulation demonstrated in this article, is performed using EDEM[®] 2019 on a desktop computer with Intel Xeon[®] E5-2650 v4 2.2 GHz (2 processors) and 128 GBs of RAM with the computational time approximating to 8 hours while using 4 CPU cores. DEM simulation, defined above, is used for mixing evaluation and comparison of above-mentioned mixing indices.

4.3. Indices for mixing evaluation

4.3.1. Review of mixing indices

This section focuses on a detailed review of above-mentioned mixing indices considered in the study. These include Lacey's index (LI), Variance amongst bimodal bin counts (VBBC), Variance reduction ratio (VRR), Relative standard deviation (RSD), Mixing segregation index (MSI), Intensity of segregation (ESI), Mixing entropy (ME), Segregation index (SI), Degree of mixedness (DM), Co-ordination number (CN), Average height (AH), Siiria method, Particle-scale mixing index (PSI), Nearest neighbor method (NNM) and Nearest distance method (NDM). Apart from these, there exist other mixing indices like Kramer's index and WLB index, which are derived from the ones mentioned above. To avoid comparison of correlated indices, the derived indices are not considered in this review. Following section enlists each of the above-mentioned indices with respective details on methodology and numerical equations used for mixing evaluation.

4.3.1.1. Lacey's index (LI)

LI is one of the earliest mixing indices, developed by Lacey in 1954 [22]. Since then, there have been many derivatives of this index, which include Kramer's index [23] based on same statistical variances as LI and L.W.B mixing index [45] which was first developed for a cylindrical mixer. LI for evaluation of mixing is based on evaluation of variance of number fraction of particles present in each grid (x_i) as compared to the average value (x_m), shown below with N corresponding to the total number of grids. Number fraction is the fraction of total number of particles of one particle type divided by that of all particle types present in the system.

$$LI = \frac{S^2 - S_o^2}{S_o^2 - S_r^2} \quad (1)$$

Here, S corresponds to the variance of the system to be analyzed, S_o corresponds to the variance of completely segregated state and S_r corresponds to the variance of perfectly mixed state. These are evaluated as shown below.

$$S^2 = \frac{1}{N-1} * \sum_{i=1}^N (x_i - x_m)^2 \quad (2)$$

$$S_r^2 = \frac{x_m * (1 - x_m)}{N} \quad (3)$$

$$S_o^2 = x_m * (1 - x_m) \quad (4)$$

For the periodic section of blender, some variations can be observed in the mixing index profile as a function of time due to the continuous motion of particles from one grid to another due to blade rotations. Being grid dependent, it is important to choose appropriate grid sizing for accurate evaluation of this index.

4.3.1.2. Variance reduction ratio (VRR)

VRR [10] has been widely used for mixing evaluations [16,46,47]. VRR is based on variance of number fraction of each type of particles present in each grid. VRR is the ratio of inlet variance and outlet variance. Thus, this index can be implemented for comparison of the mixed system with the original state and evaluate how much of inlet variance can be removed by mixing within the unit. Thus, VRR ratio is evaluated as shown below in Eq. (5).

$$VRR = \frac{\sigma_{in}^2}{\sigma_{out}^2} \quad (5)$$

$$\sigma_{in}^2 = \sigma_{out}^2 = \left(\frac{1}{N}\right) * \sum_{i=1}^N (X_i - X_{av})^2 \quad (6)$$

Here, X_i corresponds to the number fraction of each type of particles present in grid i , N corresponds to the total number of grids and X_{av} corresponds to the average number fraction of particles present in all grids. Thus, VRR index is grid and color dependent.

4.3.1.3. Variance amongst bimodal bin counts (VBBC)

Recently developed, VBBC [29] is based on calculation of variance of difference between number of particles of each type present in each grid, summed over all grids. It is evaluated as shown below.

$$\sigma_{weibi}^2 = \left(\frac{1}{N}\right) \sum_{i=1}^N \left(\left(C_{a,i}(t) - \frac{C_a(t)}{C_b(t)} * C_{b,i}(t) \right) \right)^2 \quad (7)$$

$$I_{weibi}(t) = \frac{\sigma_{weibi}^2(t)}{\max(\sigma_{weibi}^2(t))} \quad (8)$$

Where, N corresponds to the number of grids, $C_{a,i}$ corresponds to the number of 'A' type particles in grid i . $C_{b,i}$ corresponds to the number of 'B' type particles in grid i . C_a, C_b corresponds to the average of number of 'A' and 'B' type particles. Normalized VBBC ($I_{weibi}(t)$) thus obtained, varies from 1 (unmixed state) to 0 (mixed state). This index can only be used for bimodal particle systems.

4.3.1.4. Relative standard deviation (RSD)

RSD index has been widely used in literature for mixing evaluations [48-50], calculated as shown below in Eq. (9). This index being similar to variance reduction ratio evaluations, calculates standard deviation of concentration of particles present in each grid, divided by average concentration in all grids. RSD, similar to VRR, is also grid dependent.

$$RSD = \frac{\sigma}{C_{av}} \quad (9)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (C_i - C_{av})^2}{N - 1}} \quad (10)$$

Where, C_i corresponds to the concentration of particles of a specific type within grid i , whereas, C_{av} corresponds to the average concentration in all grids (N).

4.3.1.5. *Mixing segregation index (MSI)*

MSI index [51] focuses on fraction variance of number of particles of type ‘A’ with respect to the total number of particles in the system as shown below.

$$MSI = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - X_{av})^2} \quad (11)$$

$$X_i = \frac{\phi_i}{\phi_{av}}, \quad \phi_i = \frac{a_i}{a_i + b_i} \quad (12)$$

Here, N corresponds to the number of grids, a_i corresponds to the number of particles of type ‘A’ present in grid i , b_i corresponds to the number of particles of type ‘B’ present in grid ‘ i ’. This index is mainly used to identify segregation rates by plotting the segregation index as a function of number of rotations. Thus, it is grid and particle size dependent.

4.3.1.6. *Mixed Entropy (ME)*

ME index [27] uses the concept of entropy as a measure to evaluate mixing. Higher the entropy of the system, higher is the value of measured index and thus, higher are the levels of mixing present in the system. This index depends on the number fraction of each type of particles present in each grid. Thus, the entropy of each grid is defined as $e(k)$ and evaluated as shown below.

$$e(k) = x_1(k) \ln(x_1(k)) + x_2(k) \ln(x_2(k)) \quad (13)$$

Where, x_i corresponds to the number fraction of particles of type i present in grid k . Following this, ME index can be evaluated as shown in Eq (14).

$$ME = \frac{1}{N} \sum_{i=1}^N e(k, t) * n(k, t) \quad (14)$$

Here, $e(k, t)$ is the entropy and $n(k, t)$ is the number of particles present in grid k at time t and N corresponds to the total number of grids. Thus, mixed entropy index is grid and color dependent.

After normalizing, this index varies from 0 to 1 where, 0 corresponds to the unmixed state and 1 corresponds to the mixed state.

4.3.1.7. Segregation Index (SI)

SI [26] suggests use of concept of segregation to evaluate mixing. For the index evaluation, the number of contacts with different types of particles (C_{AB}) are compared with the total contacts made by one type of particles. Here, total number of contacts made by particles of 'A' type equals $C_{AA}+C_{AB}$. This information is used to evaluate mixing as shown in Eq. (15) below.

$$SI = \frac{C_{AA}}{C_{AA} + C_{AB}} + \frac{C_{BB}}{C_{BB} + C_{AB}} \quad (15)$$

SI when implemented to evaluate mixing, can range from a value of 2 indicating completely segregated state to 0 indicating mixed state. Index value of 1 indicates randomly mixed state.

4.3.1.8. Degree of mixedness (DM)

DM is a mixing index based on number fraction of particles [28] present in each grid (i), indicated by ' x_i '. The index is calculated as shown below.

$$DM = 1 - \left(\frac{\sum_{i=1}^N \frac{x_i - x_c}{N}}{x_c(1 - x_c)} \right)^{0.5} \quad (16)$$

Where, x_c corresponds to the average number fraction of all particles present in all grids. N equals the total number of grids present in the system. As this index depends on the number of particles present in each grid, the index profile exhibits some level of fluctuations due to constant movement of particles from one grid to another.

4.3.1.9. Co-ordination number (CN)

CN [31] is one of the simplest used indices for mixing. It involves measurement of number of contacts as a function of time. This mixing index is evaluated as ratio of contacts between two types of particles - A and B divided by the total contacts made by all particles separately. This index is closely related to the segregation index [26] and is evaluated as shown below.

$$CN = \frac{C_{AB}}{C_{AA} + C_{BB}} \quad (17)$$

This index varies from 0 indicating unmixed state, to a value of 1 indicating mixed state. Since this index depends on the contacts between particles which can change abruptly based on particle location and system dynamics, the index profile can exhibit some fluctuations while plateauing to the mixed state.

4.3.1.10. Intensity of segregation (ESI)

ESI [30] evaluates mixing based on location of tracer particles within the system. Tracer particles are added in a specific section along the mixing axis and the system is divided into multiple grids. As the system gets mixed, the number of tracer particles present in each grid is measured and the index evaluation is performed as a function of time. This index is evaluated as shown in Eq. (18).

$$ESI = \frac{\sigma}{\sqrt{\frac{N_{tracer}}{N_{total}} * \left(1 - \frac{N_{tracer}}{N_{total}}\right)}} \quad (18)$$

Here, N_{tracer} corresponds to the total number of tracer particles present in the system whereas, N_{total} corresponds to the total number of particles. σ corresponds the standard deviation of number fraction of tracer particles present in all grids. This index has been shown to be applicable for mixing evaluations in axial and radial directions based on placement of tracer particles.

4.3.1.11. *Nearest neighbor method (NNM)*

NNM [52], as the name suggests, evaluates mixing based on measurement of number of neighboring particles. For application of this mixing index, the system of particles is divided initially into two sections before mixing and different colors are assigned to particles in each section. As mixing progresses, the number of neighboring particles with different color is measured as a function of time. For each particle in consideration, 12 nearest neighbors are checked based on their colors. If all the neighbors have the same color, then the system is unmixed whereas, if half of the neighbors have different color, then the system is assumed to be well-mixed and is evaluated as shown below.

$$NNM = \frac{1}{N} \sum_{i=1}^N \frac{2 * n_{diff}}{n_{nb}} \quad (19)$$

Here, n_{diff} corresponds to the number of neighbors having different color than the particle in consideration. n_{nb} equal the number of neighbors which in this case, is considered as 12 and N corresponds to the total number of particles in the system.

4.3.1.12. *Nearest distance method (NDM)*

NDM [34] is based on evaluation of distance between nearest neighbors as a function of time. For a specific particle, the initial neighboring particles at time $t = 0$ are identified and their center-to-center distance is monitored as a function of time. Nearest neighboring particles imply the particles in contact with the selected particle. Due to the continuous motion of powder bed within the system, the center-to-center distance of neighbor particles can vary significantly along time. Thus, to evaluate NDM, the center-to-center distance measured earlier for all particle pairs is normalized by the center-to-center distance for randomly assigned particle pairs in the system as shown below.

$$NDM = \frac{\sum_{N_p} d_{ij} - d_p}{\sum_{N_r} d_{ik} - d_p} \quad (20)$$

Here, d_{ij} corresponds to the center-to-center distance between neighbor particle pairs – i & j . d_{ik} corresponds to the center-to-center distance between randomly selected particle pairs – i & k . d_p corresponds to the particle diameter which is the initial value of the center-to-center distance for neighboring particles in contact. N_p is the number of nearest neighbor pairs and N_r is the number of particle pairs selected randomly.

4.3.1.13. Particle scale index (PS)

PS index [25] is another mixing index based on variance of particle fraction within the powder system. The uniqueness of this index is that it takes co-ordination number also into consideration. A similar approach is applied for this index evaluation as used in nearest neighbor method where, the powder system is divided into two sections – colored differently. During powder mixing, the co-ordination number of each particle ($C_{n,i}$) is measured along with measurement of number of differently colored particles in contact ($C_{n,b}$). Thus, particle number fraction of each particle (p_i) is evaluated as shown below.

$$p_i = \frac{C_{nb}}{C_{n,i} + 1} \quad (21)$$

Similar to LI, PS index is based on the variance of calculated particle fraction, shown in Eq (22).

$$PS = \frac{S_o^2 - S_t^2}{S_o^2 - S_r^2} \quad (22)$$

$$S_t^2 = \frac{1}{N} * \sum_{i=1}^N (p_i - p_{av})^2 \quad (23)$$

Here, p_{av} corresponds to the average particle fraction for all particles and N equals the total number of particles in the system. PS index considers unmixed and completely mixed states through S_o and S_r as shown in Eq. (24-25).

$$S_o^2 = p * (1 - p) \quad (24)$$

$$S_r^2 = p * \frac{1 - p}{C_n + 1} \quad (25)$$

Here, p corresponds to the number fraction of particles present in each section with total number of particles. In this case, it equals to 0.5. C_n is the average co-ordination number for the particle system. Thus, particle scale mixing index is a combination of LI and CN. It is grid independent but cannot be used for powders with different particle sizes as it can lead to bias in measurement of co-ordination number.

4.3.1.14. Siiria method index (SM)

SM index [32] is used for mixing evaluations based on relative change of distance between different particle pairs. For evaluation of SM, all possible particle pairs are identified with measurement of the relative positions between particles of each pair. The mixing index is calculated as shown below with hyperparameters – R and g.

$$M_{ij} = R^{g|x_{ij}|} * (1 - R^{g|x_{ij}-x_{ijk}|}) \quad (26)$$

Here, x_{ij} corresponds to the initial distance between particle pair - i & j whereas, x_{ijk} is the relative distance between the particle pair at time k . Thus, M_{ij} is calculated at each time step and SM equals the averaged value of M_{ij} for all particle pairs as shown in Eq (27) with N as the total number of particles in the system.

$$SM = \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} \frac{M_{ij}}{N^2} \quad (27)$$

Parameters R & g are the dimensionless hyper-parameter and scaling factor used in Siiria index evaluation. Value of R is selected between 0 – 1 based on the relative effect of x_{ij} on magnitude of M_{ij} and a recommended value of 0.87 [32] is used. g corresponds to the scaling factor with dimensions of L^{-1} (length). Value of g is selected such that $(g*L) < 10$ as a rule of thumb, where L corresponds to the largest length between two particles. [32]. This parameter removes the dependency of length scale from the index evaluation. SM thus, is grid and color independent, as it depends on particle pairs within the powder system. Since it depends on calculation of relative position of particle pairs for all particles, this index can get computationally expensive (O^2) depending on the total number of particles considered in the system.

4.3.1.15. Average height method (AH)

As the name suggests, AH index [31] is used to evaluate mixing by calculating the average height of particles present in the system. After particles are added to the system, the particles belonging to the top half section are colored red and the bottom half are colored blue. This color assignment is performed before mixing takes place within the system. During mixing, average height of colored particles is tracked as a function of time. AH is calculated using Eq. (28), as ratio of average height of one type of particles and average height of all particles.

$$AH = \frac{1}{N_{red}} * \left(\frac{\sum_{i-red} Z_i}{\left(\frac{1}{N_{all}}\right) * \sum_{i-all} Z_i} \right) \quad (28)$$

Here, Z_i corresponds to the height of i^{th} particle. N_{red} corresponds to the total number of red particles in the system whereas, N_{all} corresponds to the total number of particles. This index can be evaluated separately for both red and blue colored particles. Index value can fluctuate as particles move throughout the system, eventually plateauing near 1. The index value indicates that the

average height for red particles in Eq. (28) is the same as the average height of all particles, indicating perfectly mixed state. The characteristic feature of this mixing index is that it can be used to measure mixing only in one direction. Thus, to evaluate mixing in all directions, it needs to be evaluated separately for all directions. Above description of AH evaluation corresponds to mixing evaluation in vertical (Z) direction.

4.3.2. Classification of mixing indices

Based on the method of evaluation used in the above mixing indices, they can be broadly classified into three categories as shown in Figure 4.3.

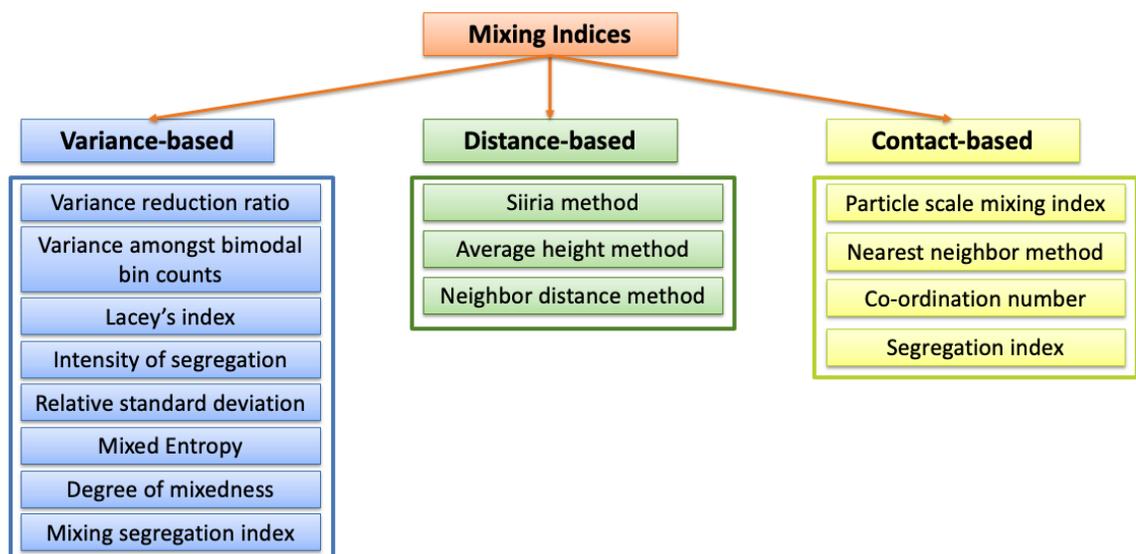


Figure 4.3: Classification of mixing indices

The three categories are variance-based, contact-based and distance-based indices. Variance-based indices include all the indices (VRR, VBBC, RSD, LI, MSI, ESI, DM and ME) which evaluate variance of distribution of tracer particles for mixing. Measurement of distribution of tracer particles within the system is grid dependent. Grid sizing thus, needs to be chosen correctly depending on the required scale of scrutiny in powder system, such that it balances both micro-

mixing and macro-mixing effects within each grid. Coarser grid selection can lead to micro-mixing effects being neglected, whereas finer grid can falsely be evaluated as unmixed systems, especially given the lower number of particles present in each grid. Thus, correct selection of gridding also depends on the average number of particles present in each grid and the end application of the powder blend. Detailed information on effects of grid selection and corresponding selectivity on evaluation of mixing indices can be found in published references [23,27,53-55].

Contact-based mixing indices require measurement of particle contacts within the powder system for mixing evaluation. The contacts within the same particle type and different particle types are identified separately from the DEM framework. This information concerning total number of contacts and corresponding particle IDs in contact is extracted for calculation of these indices. This category includes indices like SI, PS, NNM and CN. Contact-based indices are applicable for systems where the powder bulk remains intact during mixing. Though these indices are easy and quick to evaluate, there are some limitations for their application. For systems like fluidizing bed where the aeration leads to separation of powder bulk into different sections, it is difficult to evaluate contacts and can thus, lead to incorrect measurement of mixing using contact-based indices. Similarly, for agglomerating systems, contact-based indices can fail to capture the true system dynamics during powder mixing.

Third category focuses on evaluation of distance between particles for measurement of mixing. The location of the particles is extracted as a function of time to evaluate mixing within the system. This category includes AH, SM and NDM. Some of these indices based on distance evaluation can get computationally expensive, especially like SM which involves evaluation of distance between initial contact pairs and randomly assigned particle pairs. AH, on the other hand, is quick to evaluate as it involves simple assessments like particle position in only one direction. There are some indices which have been developed on the intersection of these categories like

neighbor distance and particle scale method. NDM first requires information about neighboring particles i.e. particles in contact and then evaluates the distance between them as a function of time. PS also combines variance-based LI with contact-based CN for mixing evaluation.

This classification of mixing indices is useful to identify the underlying technique used in the index calculation for evaluation of mixing. Additional variants [4] of above-mentioned indices can also be categorized in a similar way, paving a clearer path for systematic analysis of mixing indices present in literature.

4.3.3. Detailed comparison of mixing indices

Following the previous classification, this sub-section focuses on a thorough comparison of these indices based on specific features used for mixing evaluation. These features include computational complexity, grid dependency, color dependency, direction, size, and parameter dependency [33].

Computational complexity is used as a feature for comparison based on the overall time required for evaluation of the mixing index. Indices incurring higher computational times, in the order of hours, are indicated by ‘Y’, implying computationally complex indices, whereas indices which provide quick assessments, in the order of a few minutes, are indicated by ‘N’, implying indices which are not computationally complex. This serves as an important feature especially when the end goal is to achieve real-time mixing evaluation in CPM lines. Grid dependency is associated with indices which focus on distribution of tracer particles within the system. Distribution is measured within different grids and grid sizing plays an important role in correct identification of true mixing levels of powder blend. The comparison shown in Table 1, takes this grid dependency into consideration where, ‘Y’ is assigned if the grid dependency is observed. Consequently, ‘N’ is assigned for indices which are not grid dependent. Color dependency implies

that the mixing index takes properties of specific ingredient into consideration, where different ingredients are highlighted by different colors. When colored differently, the ingredient particles can be identified easily and used for mixing evaluation. Some indices like ME, VBBC, SM, NDM, CN, SI, and PS consider all ingredients present in the system, whereas some like VRR, LI, DM, RSD, ESI, MSI, AH and NNM only use one belonging to a specific particle type to evaluate mixing. For comparison purposes, color dependent indices are indicated by 'Y', whereas 'N' is assigned for color independent indices in Figure 4.4. The direction feature is used to classify the indices which evaluate mixing in all directions simultaneously, from the indices which are applied in specific directions. Generally, variance-based, and contact-based indices are direction in-dependent i.e., these indices evaluate mixing in all directions. Such indices are indicated by 'N', implying them to be 'direction in-dependent'. For distance-based indices like AH, the mixing is measured in a specific direction. Similarly, indices like NDM and SM depend on distance between particles, which is generally measured in specific direction. Thus, these indices are direction dependent and indicated by 'Y'. The direction dependent indices based on evaluation of distance can be modified to become direction independent by evaluating distance in 3-D space instead of a specified direction. However, this modification is not considered for the purpose of this comparison as it would require modification of the index proposed originally. Similarly, size feature corresponds to dependency of mixing evaluation on particle size. If the calculation of mixing index relates to the particle diameter, it cannot be used to evaluate mixing for a system with different particle sizes. Further, particle size also affects particle contacts as they are dependent on the particle diameters. Thus, indices relying on particle contacts, particle diameters and evaluation of distance between particles can show different results for particles of different sizes. These include CN, NNM, NDM, SM, PS and SI and are indicated by 'Y' in Figure 4.4, implying size dependency. Other indices are indicated by 'N' implying no size dependency. Some indices which are size dependent, can be modified to be independent, like SM, wherein, scaling parameters [32] can be used to remove scale dependency. Lastly, parameter dependency highlights dependence of mixing index on external

parameters, like SM (R, g), NNM (n_{nb}). These parameters need to be correctly chosen apriori based on their influence on the mixing index for correct evaluation. Indices which are parameter-dependent are indicated by ‘Y’, otherwise ‘N’.

Features	LI	RSD	VRR	VBBC	ESI	ME	DM	MSI	SM	AH	NDM	NNM	PS	CN	SI
Computational complexity	N	N	N	N	N	N	N	N	Y	N	Y	Y	Y	N	N
Color dependency	Y	Y	Y	N	Y	N	Y	Y	N	Y	N	Y	N	N	N
Grid dependency	Y	Y	Y	Y	Y	Y	Y	Y	N	N	N	N	N	N	N
Direction	N	N	N	N	N	N	N	N	Y	Y	Y	N	N	N	N
Size	N	N	N	N	N	N	N	N	Y	N	Y	Y	Y	Y	Y
Parameter dependency	N	N	N	N	N	N	N	N	Y	N	N	Y	N	N	N

Figure 4.4: Feature-based comparison of mixing indices

The above features highlighted in Figure 4.4 can also be interpreted as advantages and disadvantages of these mixing indices. Higher computational complexity accounts as a disadvantage for mixing evaluation. Grid dependency can allow for a deeper insight within the powder bed as it focuses on inclusion of all grids, thus allowing for an evaluation of micro-mixing within the system. However, appropriate choice of gridding is essential to balance micro and macro mixing. Thus, this accounts towards a disadvantage due to additional computational efforts needed to identify appropriate grid size. The color-dependency feature is termed to an advantage, allowing the indices to evaluate and analyze mixing for each ingredient for powder systems with 3 or more components. Parameter dependency requires additional computations to evaluate the index parameters prior to mixing estimation, thus being a disadvantage. The direction feature adds to the advantages of the indices, allowing for a thorough mixing evaluation in different directions. Lastly, size also adds to the advantages for the mixing indices as this feature allows different particle sizes into consideration for mixing evaluation. Thus, the feature-based comparison of mixing indices

shown in Figure 4.4, provides a comprehensive picture of all the indices with their corresponding functionalities. This would be helpful in ensuring appropriate choice of mixing index based on the required index dependencies. To further illustrate the effects of these dependencies for each mixing index, Table 4.1 highlights different mixing applications where the above indices have been previously applied in literature.

Table 4.1: Mixing indices with corresponding applications

Mixing Index	Applications for mixing evaluation
LI	Paddle blade mixer, batch bladed mixer [56], rotating drum [33,57], fluidized bed [58], V-binder and bin-blender [59]
VBBC	Continuous blender with two particle types [29], milli-static mixers [60]
VRR	Continuous blender with multiple particle types [46,61,62]
RSD	Paddle blade mixer [63], continuous mixer [16], Bohle bin blender [64]
AH	Vertical mixing – fluidized bed [34,65]
NDM	Rotating drum, spouted bed [33]
NNM	Fluidized bed, spouted bed and rotating drum [33,52]
CN	Rotating drum [33], vertical cylindrical mixer [25]
PS	Plowshare mixer [54], ribbon mixer [66], lab-scale screw mixer [67], vertical bladed mixer [68]
ME	Horizontal rotating drum [27], rotary kilns [69]
SM	multiple-spouted bed [70], stirred tank [71]
MSI	Rotating drum [72], fluidized bed [34] and tumbling blenders [51]
ESI	Co-rotating screws extruder [73], twin-shaft kneader [74]

SI	Rotating drum, hoop mixer, turbula mixer[26]
DM	Fluidized bed [75,76], V-blender [77], slant-cone mixer [78]

To obtain a quantitative comparison of the above mixing indices, T_{95} mixing time [34] is evaluated for each index. T_{95} mixing time corresponds to the time required for the system to reach 95% of the completely mixed state. This also condenses the information pertaining to a time dependent mixing curve into a single value, which is more suitable for comparison. It is defined as shown in Eq. (29) where, a dampened exponential function is used to fit the mixing curve.

$$M_{fit} = 1 - A * \exp(-\gamma * t) \quad (29)$$

where the parameters A and γ are obtained by fitting the exponential function to the mixing curve over time (t). For indices like average height which show oscillations, the above equation is modified to include harmonic oscillations with frequency (w) as shown in Eq. (30).

$$M_{fit} = 1 - A * \exp(-\gamma * t) * \cos(w * t) \quad (30)$$

Using the above equations, T_{95} mixing time, corresponding to the time at which M_{fit} equals 0.95, is evaluated for all mixing indices considered in this study.

4.4. Results and discussion – DEM simulation and mixing evaluation

4.4.1. DEM simulation and data extraction

Simulation results obtained for periodic section modeling of blender unit as described in Section 2, are shown below in Figure 4.5.

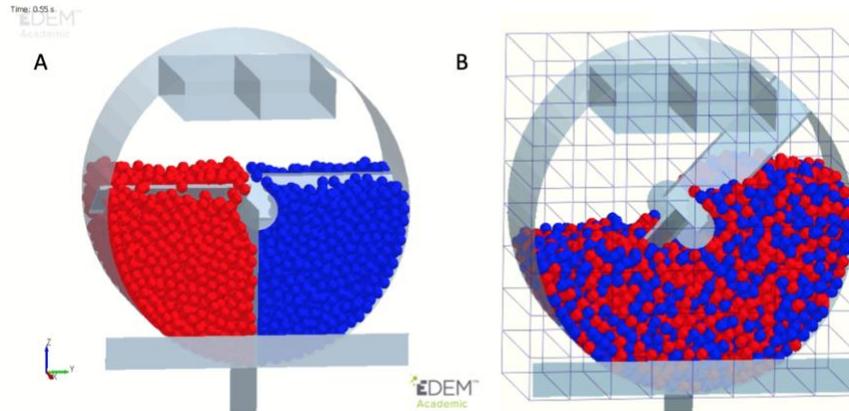


Figure 4.5: DEM simulation of blender using PSM - A) initial state of un-mixed powder system, B) Mixed state of powder system with chosen gridding for mixing evaluation

Figure 4.5A illustrates the unmixed state of powder system within DEM at time $t = 0$ sec whereas Figure 4.5B describes the mixed state of the powder system in DEM at time = 20 seconds. Figure 4.5B also shows the choice of gridding used for mixing index evaluations. Appropriate grid sizing is determined based on the end application of the powder system while determining a balance between micro-mixing and macro-mixing within the system. A recommended gridding size of $1 \times 8 \times 8$ (X, Y, Z axes) based on Boukouvala et al. [20], is intended for this study. It is important to note that periodic section modeling has periodic boundaries, i.e., powder exiting the system along X axis, enters the system again from the opposite side. Thus, the progression of mixing in PSM along time can also be interpreted as progression of powder mixing along the length of the continuous blender (X axis). The mixing levels shown in Figure 4.5B can be used to evaluate mixing at the end of a continuous blender with residence time equal to simulation time of PSM blender model.

For mixing evaluations, particle information including the number of particles of each particle type present in each grid, number of contacts and location of particles is extracted along time, from the DEM simulation. Based on the chosen grid size, there are some grids which contain

particles less than 50% of the average number of particles present in each grid ($0.5 * N_{av} = 0.5 * \frac{6000}{64} \approx 0.5 * 94 = 47$). These grids are considered to be empty or incomplete for data extraction purposes [20] and are not considered for mixing evaluations as they can lead to fluctuations in the index evaluations. These mainly correspond to the grids in the upper half of the periodic section. Grids containing number of particles higher than average number of particles can be counted as representative samples for powder blend. Thus, these grids are considered in the mixing evaluations as shown below.

4.4.2. Evaluation of mixing indices

Based on the data extracted from the DEM simulation as a function of time, mixing indices are calculated. This allows for a time variant analysis of mixing indices. Figure 4.6 illustrates the dynamic profile of all the mixing indices considered in this study for periodic section model of blender.

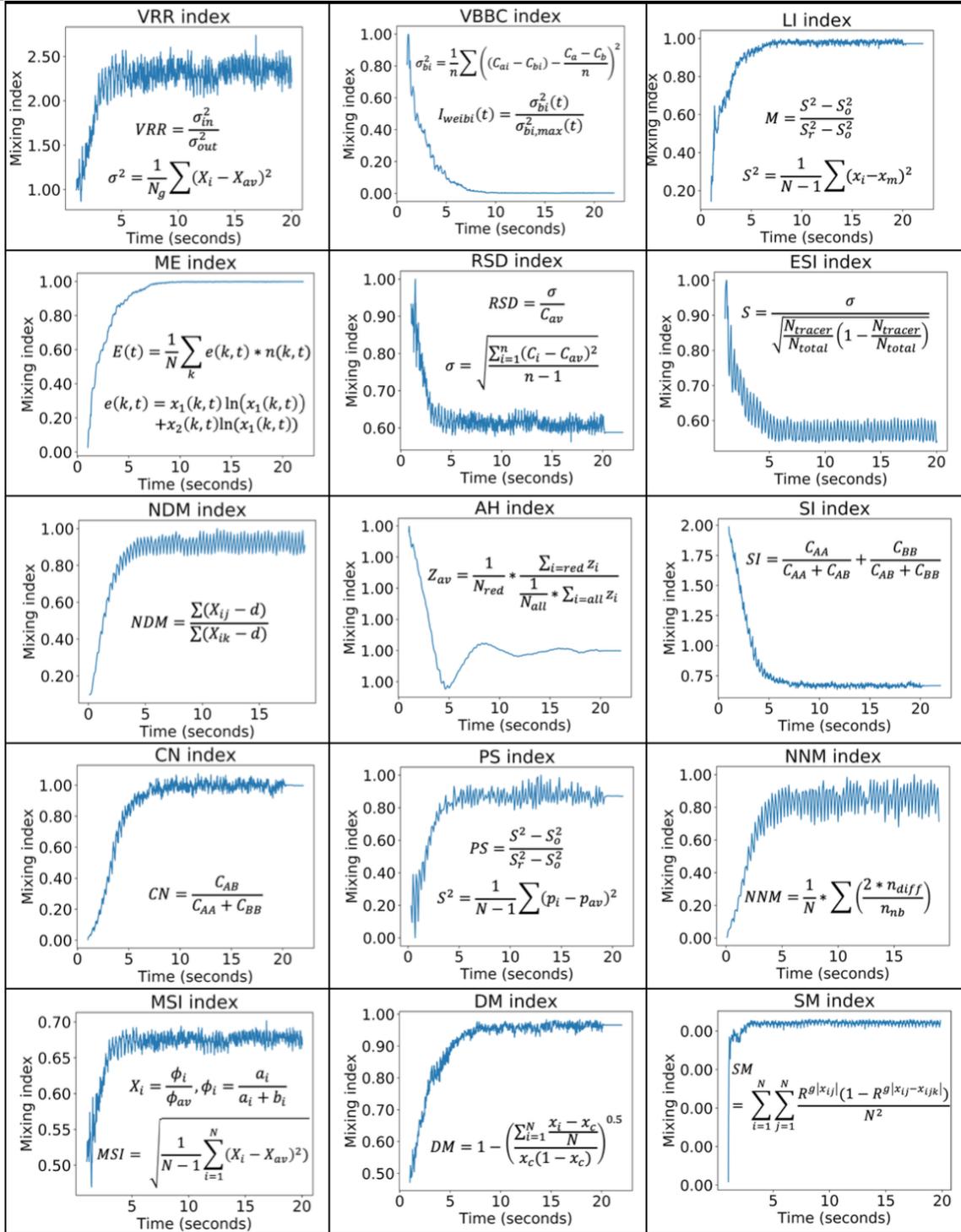


Figure 4.6: Mixing indices for evaluation of mixing in the periodic section with subplots for different indices.

Figure 4.6 highlights index evaluations using red-colored particles for indices which can be evaluated separately for each particle type. These include LI, VRR, RSD, ESI, MSI, DM and AH. Other indices consider both particle types like SM, CN, PS, SI, NDM, NNM, ME and VBBC. All the indices shown in Figure 4.6 are evaluated at a time step of 0.1 seconds. It can be observed from the above index calculations that some indices show higher fluctuations than others. These differences are due to the inherent methods used for mixing analysis and the fluctuations are observed for indices which involve measurement of number of particles in each grid. Due to the movement of blades, particles constantly move from one grid to another, leading to the observed fluctuations. Another important observation is the variety in the range of values obtained for different indices which makes it difficult to compare numerically with each other. To overcome this aspect, the indices can be normalized between 0 to 1. This allows for an easy comparison between different indices.

4.4.3. Comparison of indices based on the T_{95} mixing times

Obtained index profiles are smoothened using a smoothening function for filtering the fluctuations using Savitsky-Golay smoothening filter [79]. These smoothened profiles are then regressed to obtain T_{95} mixing times and corresponding goodness of model fits along with model parameters are recorded as shown in the section 8.2. This is performed for different blade speeds of the blender unit – 100 RPM, 75 RPM and 50 RPM to evaluate the effects of blade rotation speed on the mixing time. The mixing indices which demand higher computation time than one hour are not included in this comparison. These include computationally heavy indices like SM, NDM and PS. Table 4.2 shows the comparison of T_{95} mixing times for different indices at three levels of blade speed. The mixing indices are colored based on the classification proposed in Figure 4.3. Obtained T_{95} mixing times can also be plotted as shown in Figure 4.7 to compare different indices and visually observe the effects of blade speed on the T_{95} mixing time.

Table 4.2: T_{95} mixing times for indices with different blade speeds

Mixing Index	Blade speed (RPM)		
	100 RPM	75 RPM	50 RPM
DM	7.595	10.892	14.349
VBBC	3.902	6.268	8.733
ME	5.004	5.941	8.207
VRR	6.955	8.440	12.257
MSIs	4.436	4.645	6.415
ESI	6.392	8.301	11.729
RSD	6.727	8.110	10.422
LI	3.605	5.617	7.849
CN	8.176	11.943	17.483
SI	6.298	9.203	12.301
NNM	7.180	9.715	13.470
AH	4.648	10.496	20.501

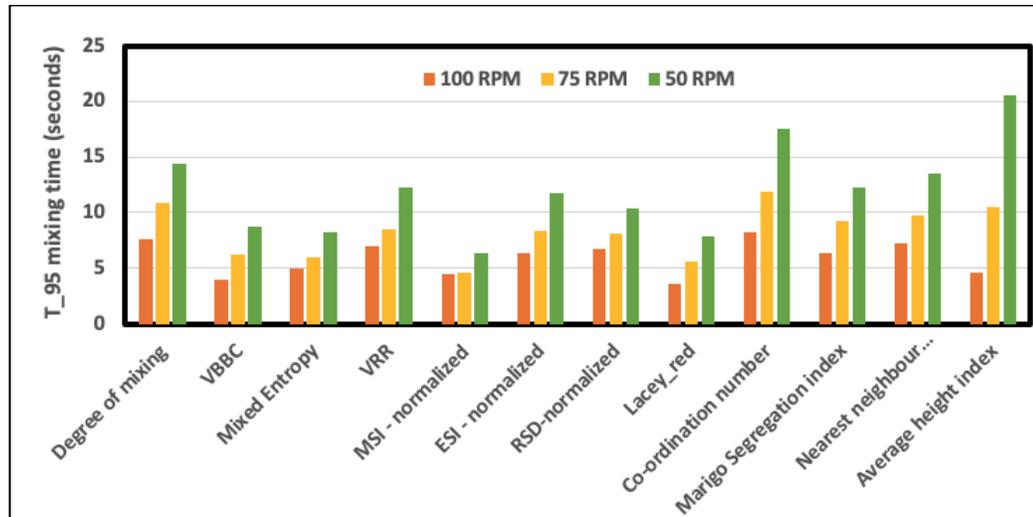


Figure 4.7: T_{95} mixing times for indices at different blade speeds

It can be observed from Figure 4.7 that as blade speed increases, the T_{95} mixing time decreases. This is in agreement with previously recorded observations for blender systems in experimental and computational studies [7,19]. Furthermore, it can be observed that these indices result in different mixing times based on the methods implemented. This extends to the classification of mixing indices listed above – variance, distance and contact-based indices where, they use different methods to measure mixing. Variance based indices (DM, VBBC, ME, VRR, MSI, ESI, RSD, LI) show smaller mixing times as compared to other indices, whereas, contact-based indices including CN, SI and NNM encounter higher values of T_{95} . AH, a distance-based metric also shows higher values, especially for 50 RPM. Since this is a uni-directional index, it can be concluded that the system needs more time to get well-mixed, specifically in Y direction. An alternative would be to use a combination of multiple directions for obtaining accurate estimates for the AH.

4.4.4. Selection of appropriate mixing index

Based on the evaluation of different mixing indices for the powder blending system, this section illustrates the guidelines to be adopted for choice of appropriate mixing index based on the required application. To properly choose a mixing index, following questions need to be addressed prior to the mixing evaluation. The first question is – ‘what powder system is used for the specific study?’. This includes information concerning the particle size, number of particles, types of particles and initial position of these particles within the system. This is important as it allows the user to identify possible risks including segregation, agglomeration and unidirectional mixing associated with the powder system. The second question is – ‘what is the scope of the mixing evaluation?’. This is important as it allows the user to focus on the required scale of scrutiny and accuracy of mixing evaluation necessary for the end application of powder blend. Furthermore, it provides a rough estimate for the appropriate grid sizing balancing the measurement of micro-mixing and macro-mixing within the system. These questions help identify the requirements of mixing evaluation of the powder system and correspondingly select candidate mixing indices. This is illustrated in detail, in Figure 4.8, highlighting the above questions, specific concerns associated with these questions, different options for each concern, corresponding effects of each options on mixing evaluation and a list of indices applicable for each option. Thus, given a new powder system, this diagram can be used to identify which concerns and corresponding options apply to the powder system under study. Following these options, effects on powder mixing can illustrate how the index calculations can get affected based on the selected options. Finally, a list of suitable mixing indices can be selected. Based on the multiple concerns relevant for a powder system, the lists of allowable indices can be compared to select a subset of appropriate indices which together address all concerns. These indices differ from each other based on the specific features highlighted in the review, classification, and comparison of these indices. In-depth scrutiny of these features would further allow the user to identify most appropriate index for the specific system under study.

Questions	Specific concerns	Options	Effects on powder mixing	Indices for powder mixing
Powder system for mixing study	Types of particles	Two types of particles	Index allows two particle types; average concentration is 0.5	Indices considering two particle types like VBBC, ME, SI, CN, and MSI
		Multiple types of particles	Index allows multiple particles types; average concentration to be evaluated	VRR, LI, RSD, ESI, NDM, AH, PS, NNM, DM, SI
	Particle sizes for different particle types	Same sizes for all particles	Particle contacts and distance between particles not affected	All indices – contact, distance and variance-based indices.
		Different sizes of all particles	Different particle size affects contacts and distance between particles	Segregation & variance-based indices - VRR, LI, VBBC, RSD, ESI, MSI, ME, DM.
	Total number of particles ($N_{t,i}$) for each particle type	Same for all particle types	Total number of particles of each type considered to be equal	Indices not accounting for $N_{t,i}$ - VBBC, SI, NDM, CN, NNM, SM
		Different for all particle types	Total number of particles of each type to be considered	Indices accounting $N_{t,i}$ like VRR, LI, ME, RSD, ESI, AH, PS, MSI, DM
	Location of particle addition	Vertical – top to bottom	Direction dependent & distance-based evaluated along vertical direction.	Indices like AH, NDM & SM evaluated vertically, all other indices applicable.
		Horizontal – side to side	Direction dependent & distance based evaluated along horizontal direction	Indices like AH, NDM & SM evaluated horizontally, other indices applicable.
Scope of mixing evaluation	Required scale of scrutiny for powder blend	Balance between macro-mixing and micro-mixing	Choice of appropriate grid sizing based on the end application of powder blend	Grid selection important for variance-based indices; other indices applicable

Figure 4.8: Logical diagram for selection of appropriate mixing index

An additional consideration which needs to be considered while selecting candidate indices includes the overall computational requirement required for mixing evaluation. This is important, especially in digital CPM lines as blend uniformity is measured in real-time where, corrective measures can be devised to rectify the blend non-uniformity in downstream unit operations.

Choice of mixing indices also depend on the dynamics of the powder system under study. For specific applications like fluidized bed and high shear blenders, where powder separates from the bulk due to the system dynamics, the indices need to be chosen such that they incorporate the phenomena of bulk separation. Bulk separation refers to the partial or complete fluidization of powder bed due to the system dynamics. Thus, contact-based indices would not be applicable for such systems due to high variability in contacts observed in the system. Distance-based indices like NNM, NDM and SM, though considers all particles within the system, can get computationally intensive for real-time application. Variance-based indices can be applicable for real-time analysis of such systems, given appropriate selection of gridding. Grid-dependency can also be addressed

by including dynamic gridding within the mixing evaluation. Thus, gridding can be chosen dynamically between coarse gridding (to incorporate process dynamics and bulk separation), and fine gridding (to comply the regulatory requirements of blend uniformity). Variance-based indices incorporating segregation can be utilized for systems with different particle sizes.

4.5. Summary and conclusions

Accurate measurement of blend uniformity is vital for the efficient and regulation-compliant operation of CPM lines. Mixing indices are used for evaluation of blend uniformity at the blender outlet. Since 1950s, there have multiple studies focusing on development of new mixing indices and modifications based on existing indices, for mixing evaluations. Given this overabundance of mixing indices in literature, it becomes challenging to identify the appropriate index for mixing evaluation of specific application. Furthermore, due to the application of different indices, it becomes cumbersome to systematically combine and utilize this vast available literature for understanding and improving blending process through optimal selection of design, operating and process parameters. This article thus aims at addressing the above concern of availability of multiple indices in literature by performing a detailed review, classification, and comparison of different mixing indices. A computational approach is implemented for measurement of blend uniformity, wherein, a DEM simulation of blender is used as a case study for comparison of mixing indices. Following a detailed review, the article aims at addressing the considerations for selection of appropriate mixing index, aiming to capture the true mixing mechanism in the system under study.

Different mixing indices considered in this review can be broadly classified into three categories: Variance-based, Contact-based, and Distance-based. Variance-based indices include LI, VRR, VBBC, RSD, MSI, DM, ME, and ESI. These indices utilize variance of number of particles present in grids for mixing evaluation, hence termed as variance-based indices. For these indices,

appropriate grid sizing is critical to balance measurement of micro-mixing and macro-mixing based on the intended application. These indices are quick to evaluate and can be successfully implemented for real-time mixing measurement. Contact-based indices are based on evaluation of total number of contacts between particles in the powder system. These contacts are measured separately for same and different particle types, which combined, can be applied for mixing evaluation. Contact based indices are also quick to evaluate but can result in erroneous results, especially for fluidizing systems. The bulk powder within such systems can separate, losing contacts between particles intermittently, which in turn affects contact based mixing evaluation. Hence, these indices can only be implemented for those applications where the powder bulk remains intact i.e., particle contacts are not affected by system dynamics. This category of indices comprises of CN, SI, NNM and PS. Lastly, Distance-based index relies on measurement of distance between particles for mixing evaluation. This category includes indices like AH, NDM and SM. Except AH index, other distance-based indices are very computationally intensive as they consider distance between all possible pairs of particles present in the system, making them incompatible for real-time applications. Given the nature of mixing evaluation in distance based and contact-based indices, these are grid-independent but color dependent.

To obtain a qualitative comparison of these indices, a feature-based comparison is conducted, wherein, the above-mentioned indices are compared based on features including computational complexity, grid dependency, color dependency, multi-direction, multi-size and parameter dependency. This feature-based comparison allows user to appropriately select the mixing index based on the requirements of the powder system. To obtain a quantitative comparison of the indices, they are applied for mixing evaluation of a powder blender system. The blending simulation is performed using DEM with PSM methodology wherein, two types of cohesionless, spherical particles with same particle size and properties are intended to be mixed. After completion of PSM simulation for a blade speed of 100 RPM, mixing indices are applied to evaluate function

of mixing levels along time. To obtain quantitative results, T_{95} mixing time is calculated for each index which indicates the time at which mixing level in the powder system reaches 95% of overall mixing. Further, to observe the effects of blade speed on the T_{95} mixing time, the PSM blender simulation is modified for additional blade speeds – 75 RPM and 50 RPM. Overall comparison of indices at different levels of blade speeds show a similar trend of increasing mixing time with blade speed, corroborated by previous studies [7,41].

The above qualitative and quantitative comparison aims towards the addressing the main question of which index should be applied for mixing evaluation. Given various features, advantages and disadvantages associated with each index, the mixing indices need to be chosen based on the intended application of mixing within the powder system. The two questions that need to be taken into consideration before implementation of mixing indices are: information concerning the powder system intended to be mixed and the underlying reason for powder mixing. These together can aid in suitable selection of mixing index. This review article focuses on comparison of these indices for a blender unit. Other case studies including fluidized bed, sprouted beds, and high shear blenders can also be incorporated to justify the choice of mixing index. Given the same properties for both types of particles selected in the study, phenomena like agglomeration, electrostatics and segregation are not observed. However, different particle properties with cohesive powder system can also be incorporated to check for agglomeration issues. Detailed review and comparison of mixing indices can aid in proper utilization of available literature on blender systems and powder mixing for enhancement in process understanding. This can further assist in selection of optimal parameters for process operation by accurate measurement of blend uniformity. Eventually, this would be helpful to identify the suitable index for real-time application of measurement of mixing in digital CPM lines based on the powder properties, where the index incorporates and assesses the true mixing phenomena observed in the powder system.

CHAPTER V

5. Aim 2: Unit operation process modeling using DEM – compartment modeling

Acknowledgement of publication status:

Entire sections of this chapter have been published in a scientific journal written by the author of this thesis with the title: Hybrid multi-zonal compartment modeling for continuous powder blending processes. This article was accepted in the International Journal of Pharmaceutics on 21st April 2021.

5.1. Introduction

CM, since its conceptualization in the last decade[1-3], are being actively researched to enable their adoption in the pharmaceutical industry. Towards this end, innovative research efforts have focused on developing and implementing pioneering initiatives [4-7] to advance drug production and maximize the overall efficiency and robustness of manufacturing line. These initiatives are in accordance with the U.S. FDA’s vision for 21st century [8] aiming to *‘promote a maximally efficient, agile, and flexible pharmaceutical manufacturing sector that reliably produces high quality drugs without extensive regulatory oversight’*. One such initiative is the QbD[9] defined as a systematic approach for advancing the pharmaceutical manufacturing through detailed product and process understanding, based on quality risk management. QbD thus focuses on designing the process such that the product quality is assured, incorporating the detailed process understanding of the overall system. For the successful implementation of this initiative, it is important to develop

the capability to provide detailed process understanding from a mechanistic perspective, along with an accurate assessment of process prediction. There has been substantial emphasis of academic research efforts and industrial collaborations directed towards this effort by developing the required experimental and modeling capabilities over the past decade. These capabilities focus on real-time measurement and evaluation of effects of process parameters on product critical quality attributes (CQAs) of the manufacturing line. FDA Q8(R2) Pharmaceutical development [10] defines CQA as the important properties of the drug substance, excipients, intermediates and drug products that affect the product quality, purity, stability and bioavailability. These properties are measured at the outlet of intermediate unit operations in the manufacturing process, to direct and guide the drug product manufacturing and development. The CQAs identified using quality risk assessment[11], evaluate the extent of impact of change in critical process parameters (CPPs) on the final drug product quality. A detailed description of the important CQAs can be found in literature[12]. To produce oral solid-based drug products, blend uniformity (BU) is an important CQA, measured at the outlet of the continuous powder blender. Measurement of BU ensures that the uniformity of the powder blend [13,14] lies within the limits defined by regulatory guidelines and depends on the extent of powder mixing within the system. This is a crucial CQA, as the accurate measurement of BU provides enough time for the process control strategies to rectify any divergences outside of the regulatory limits[15], before the non-conformant powder blend passes through the manufacturing line, thus preventing the production of off-specification drug products.

Continuous powder blenders have been extensively studied in literature[16-20], including experimental approaches like mixing analysis [14,21-23], residence time distribution studies [19,24-26], holdup analysis [25,27]and modeling approaches like discrete element modeling[28-30], population balance modeling [31] [32], surrogate models[33,34], and RTD models [18,19]. These approaches are mainly directed for obtaining process understanding of powder flow and evaluation of powder mixing within the blender. The assessment of powder mixing performed

along axial direction is measured experimentally through residence time distribution (RTD) and corresponding RTD models are developed accordingly. Furthermore, the influence of various process parameters and material properties on the extent of axial mixing has also been investigated thoroughly in literature[16]. Conversely, the pharmaceutical literature lacks a thorough description on accurate experimental measurement and modeling of powder mixing along the radial direction. The evaluation of radial mixing has been attempted experimentally through use of measurement probes [35] and techniques like positron emission particle tracking techniques (PEPT) [36]. However, these methods have some shortcomings which limit their application. Researchers [35] have reported problems like disruption of powder bed, identification of statistically significant number of samples while using measurement probes for assessment of radial mixing. Though techniques like PEPT can avoid such issues, they can get very expensive due to the requirement of specialized equipment setup for measurement [36]. These difficulties thus complicate the accurate measurement and analysis of radial mixing, which further hinders assessment of blend uniformity for its inclusion in the predictive models for continuous blender units.

To address the above shortcomings, we propose the use of first principle-based simulation techniques like discrete element modeling (DEM) for accurate measurement of powder mixing, followed by the development of a predictive model for the continuous powder blenders. DEM has been widely studied in literature in order to understand particle-level physics during pharmaceutical process operations [37] including continuous blenders[38-40]. This technique is applied to evaluate powder mixing within the blender system, based on particle-scale information extracted from the DEM simulation. However, one of the major drawbacks of using DEM is that it can get computationally intensive due to the evaluation of particle contacts and contact forces for all particles in the system and at all time steps of the simulation[41]. This results in the simulation time of the order of 8-10 hours, and thus cannot be used for quick assessment of powder mixing. To address this challenge and subsequently develop a computationally efficient technique for modeling

of powder mixing within the continuous powder blender, we propose the use of hybrid multi-zonal compartment modeling strategy. Multi-zonal compartment modeling strategy is a technique that transforms the detailed information obtained from high-fidelity simulations like DEM into a network of inter-connected compartments, which can replicate the observed process behavior efficiently. This network of inter-connected compartments is then used to develop a predictive hybrid model for the continuous powder blender, which can then be used in flowsheet modeling of the entire manufacturing line. In the presented work, multi-zonal compartment modeling strategy is developed and demonstrated for a periodic section of the continuous powder blender followed by its extension for the entire blender. Previous publications [23,40] in this area showcase the use of similar methodology of compartment modeling for the blender systems, however the choice of compartmentalization is determined based on the user experience. The dependence of compartmentalization on user-expertise can restrict the applicability of the approach since it does not incorporate process information within the development stage. The novelty of the proposed work stems from the application of mechanistically informed compartmentalization using DEM for the development of a computationally efficient hybrid model using multi-zonal compartment modeling strategy for model predictions.

The chapter is outlined as follows. Initially, the different methodologies utilized in the proposed work are introduced in sections 2, including DEM and multi-zonal compartment modeling strategy. This is followed by their application for the periodic section of the continuous powder blender system as presented in section 3. The results of the proposed approach are demonstrated in section 3, supplemented by the validation. Following the model validation, the applications of the developed model for the periodic section are described, including application of periodic section for evaluation of radial mixing and its extension for modeling the entire continuous blender. This leads to the development of a computationally efficient hybrid model for the continuous powder blender. Lastly, the conclusions and potential avenues for future work are presented in section 4,

including further improvement of proposed methodology and application for other unit operations in the pharmaceutical manufacturing line.

5.2. Materials and methods

5.2.1. Discrete element modeling (DEM)

DEM, first introduced in 1970s[41], has been widely applied for the modeling of particulate systems in the pharmaceutical industry[37] to develop first principle-based understanding of powder dynamics. This method focuses on the numerical modeling of powder mechanics based on Newton's laws of motion and involves evaluation of contact forces acting on each particle based on its interactions with other particles and the walls of the equipment being simulated. The contact forces including normal and tangential forces, are evaluated based on the selected contact model for each particle and at each time step, making this method computationally intensive. The normal and tangential contact forces thus identified, then dictate the trajectory of the particle during the next time step. The equations involved in DEM are described in detail in literature [30,42]. The DEM methodology thus can be applied for different applications including particle tracking, detailed understanding of particle flow, and evaluation of particle mixing within the particulate system.

The detailed steps involved in DEM modeling are summarized in Figure 5.1. The first step involves developing the computer-aided geometry for the desired powder system. Following this, the particles resembling pharmaceutical powder, are added to the system using virtual factories. These particles can be developed based on the user specifications and particle properties, including particle size and shape. Due to the high computational requirement of DEM, generally the particles are scaled up in size around the range of few hundred microns and the particle properties are chosen such that the bulk nature of powder within DEM as observed experimentally. Desired contact

models are chosen, and particles are added to the desired geometry. Next, the particulate system is simulated, and the following events occur recursively for the user defined duration of simulation time. These events include contact detection between particles, evaluation of contact forces – normal and tangential forces, and determination of particle trajectory for the next time step, as shown in Figure 5.1. The time step for the DEM simulation is determined based on the Rayleigh's time step and a value of 10-20% of Rayleigh's time step is recommended [30,42].

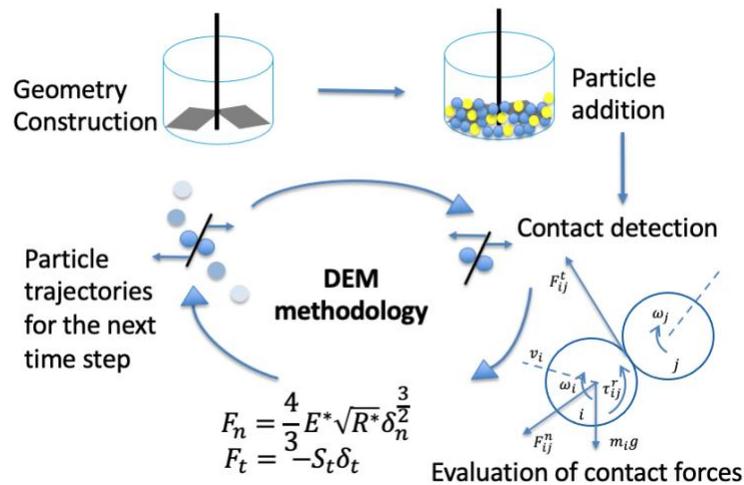


Figure 5.1: Discrete element modeling methodology

In the presented work, we implement the DEM method for simulation of continuous powder blender system. Given the inherently high computational expense associated with DEM modeling, a small section of the continuous powder blender is simulated. This section is chosen specifically to represent the central region of the powder blender, with periodic boundaries on both ends. The periodic boundaries ensure that the particles exiting from one boundary of the periodic section, re-enter the system from the other boundary. Thus, the periodic section simulated for the residence time of the continuous powder blender, can replicate the powder behavior as observed along the axis of the blender. In this work, we develop the periodic section similar to that used in previous publications[33,34,38,43,44] and similar values of DEM parameters are chosen for this work. The details of the DEM simulation along with the chosen model parameters are provided in the section

8.3.1. In the DEM simulation, the walls of the equipment and the blades are considered as hard boundaries, whereas the periodic boundaries are open on both sides to allow periodic movement of particles. The geometric dimensions of the periodic section are 40mm x 80mm x 80 mm along X, Y and Z directions, as shown in Figure 5.2.

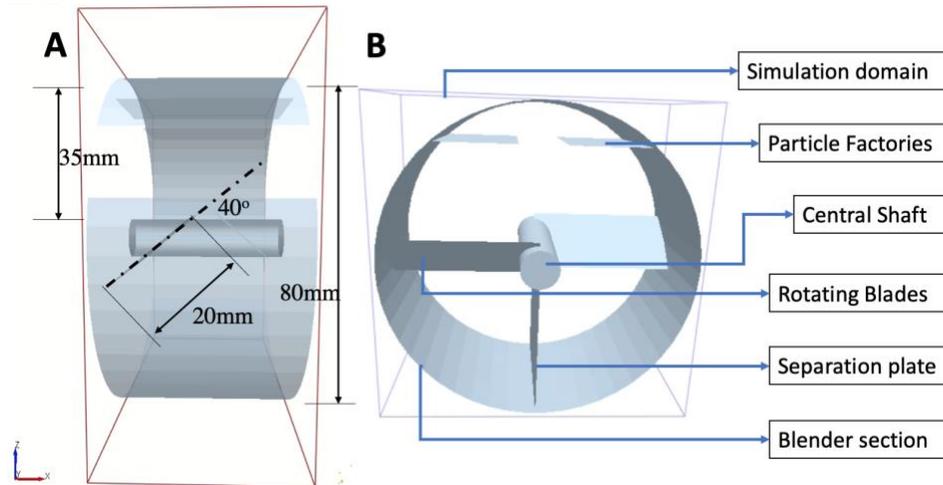


Figure 5.2: Geometric details of the periodic section used in this study (Taken with permission from Bhalode and Ierapetritou, 2020)

Following the development of CAD geometry, cohesionless, spherical particles with particle radius of 1 mm are added to the system. To provide ease of visualization of powder mixing within the periodic section, two types of particles, differing only in color, are added on either side of the periodic section at the start of the simulation. The periodic section at time = 0 seconds with the two types of particles colored as red and blue, is shown in Figure 5.3A. A total of 3000 particles for each particle type are added, resulting in a total of 6000 particles in the DEM simulation [43,44]. During addition of particles, a separation plate is added to the geometry, to ensure the particles are completely unmixed at the start of the simulation. Following the addition of particles, the plate is removed, and the blades are rotated at a blade speed of 100 RPM for 20 seconds, to simulate powder mixing. The mixing time is determined based on mixing index evaluations[43] to ensure the powder mixing reaches a steady state, as shown in Figure 5.3B. The DEM simulation of periodic section,

as demonstrated in this chapter, is performed on a desktop computer with Intel Xeon E5-2650 v4 2.2 GHz (2 processors) and 128 GB RAM using a commercial software EDEM® 2020, resulting in an approximate computation time of 8 hours with 4 CPU cores.

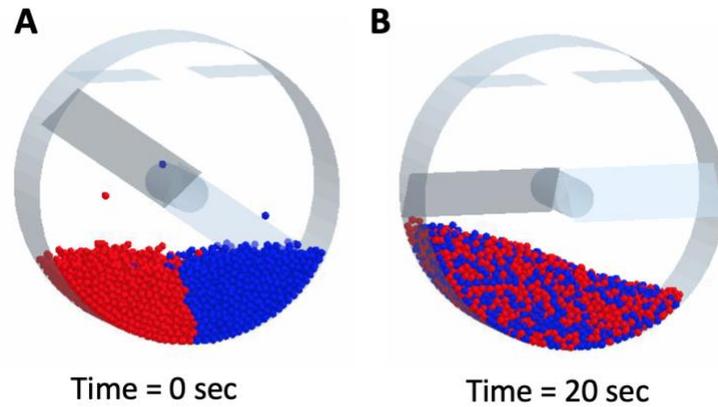


Figure 5.3: Simulation of periodic section using DEM - A) periodic section at $t=0$ sec, B) periodic section at $t=20$ sec

5.2.2. Multi-zonal compartment modeling

Multi-zonal compartment modeling is a hybrid technique to simulate complex systems. This technique comprises of developing a computationally efficient representation of the system based on functional zones called compartments[45]. These compartments are evaluated based on the spatially distribution of important process variables of the system. These when connected to develop the compartment model (CM), the model can then be utilized to represent systemic or global level information of the complex system. Thus, the CM model is constructed such that it provides multi-scale modeling with low computation time and is known by different terminologies in literature[45], including multi-zonal models, network of zones, and multi-block models. Multi-zonal compartment modeling approach divides the system into multiple zones based on the distribution of process variables of interest to construct an inter-connected network of these zones. This modeling approach was first introduced in 1980s[46] and implemented to model the system under study as a network of ideal reactors. Following the introduction of this concept, there have been many publications in literature which apply compartment modeling for applications like

bioreactors[47-49], bubble columns [50], chemical reactors [51,52], crystallizers[53,54], and aerated canal reactors[55,56].

In the proposed work, we aim to implement the CM methodology using the rationalization approach for development of a hybrid model of the periodic section in the form of inter-connected networks of compartments[44,49], which will be extended for modeling the entire continuous powder blender. The proposed methodology for developing multi-zonal compartment model using DEM, includes three steps - initialization, zoning, and compartmentalization as explained in this section.

Step 1. The initialization step focuses on determination of important process variables of interest to be used for evaluation of zones. Zones are the homogenous regions within the system, wherein, process variables are within the homogeneity tolerance limits. The number of process variables can be limited to one variable or multiple based on the process output and the required application of the model. The corresponding user-defined tolerance range for chosen variables are also determined in this step, to discretize the system into separate zones. For application of this step, a high-fidelity simulation like CFD or DEM is simulated and the values for chosen variables are extracted for each element of the simulation. These elements are first grouped together to develop square-shaped grids to ensure uniformity of grid sizing across the system, and the variable information pertaining to these grids is obtained by averaging the variable information for each element present in the respective grid. This is performed to ensure there aren't any irregular shaped grids present in the system[49]. Following this, the user-defined tolerance range of variables is selected based on the degree of variation between maximum and minimum values. This needs to be chosen carefully to balance the number of compartments and the overall computational requirements of the compartment model. Higher the chosen number of compartments, more accurate is the obtained model prediction, but higher is the associated computational cost for the developed compartment model.

Step 2. The zoning step focuses on categorizing all the grids in the system into discrete zones based on the variable parameters chosen in the first step. There exist two methods to perform zoning for high-fidelity simulations – manual zoning and automated zoning. Manual zoning is performed by classifying the system domain into zones based on the user expertise. To limit the dependence of user expertise on zoning strategies, automated zoning is preferred[47,48], which includes classification of grids of the entire domain into spatial zones, based on the variable parameters defined in the initialization step. This step is thus used to develop zone maps of the system.

Step 3. Compartmentalization involves development of compartment maps using the zone maps developed in step 2. The compartment maps are developed as follows. The multi-regional zones in the zone map are separated into multiple mono-regional compartments to develop a network of compartments. This is followed by evaluation of inter-connection flowrates between these compartments, to develop the compartment maps of the original simulation. Initially, the zone maps are checked for presence of any multi-regional zones or spatially disconnected grids belonging to the same zone. If present, each spatially disconnected group of grids is classified into a separate compartment to develop a spatially distributed network of compartments. The distributed network of compartments thus acts as a data-driven surrogate representation of powder flow in the original simulation. Following this, the powder flowrate between the distributed compartments is evaluated based on the original CFD or DEM simulation. The inter-connection flowrates are evaluated separately in all directions to ensure overall systemic flow is well-captured. Lastly, the data-driven inter-connected network of compartments is combined with inter-connection flowrates to provide a computationally efficient hybrid representation of the system under study.

5.2.3. Method for development of compartment model

Following the framework for construction of multi-zonal compartment models described in section 3, this section aims at implementing the proposed framework for the periodic section using DEM. The multi-zonal compartment modeling strategy is implemented to combine the systemic process knowledge pertaining to axial mixing, with local variable information including velocity distribution for radial mixing. In this work, we develop multi-zonal CM based on particle velocities along all directions as the variables of interest, to effectively capture the systemic and local level information of the blender system. After completion of the DEM simulation of the periodic section, the particle-level information is extracted from DEM for the initialization step. This is important for understanding the particle level physics and characterizing the powder flow within the unit. To obtain this information, appropriate grid sizing needs to be chosen for the elements of the DEM simulation. In this work, the chosen grid size is 10 along X direction, and 20 along the Y and Z directions, based on previous publication [43,44]. After developing these grids along X, Y and Z directions, the particle level information for these grids is extracted and the user-defined tolerance limit for zoning is determined based on the range of values of particle velocities. The next step comprises of zoning the periodic section based on these velocities and tolerance limits, using a similar methodology as proposed by Tajssoleiman et al. [49]. Before the zoning step, we break the periodic section down into multiple slices along the axial and radial direction, as shown in Figure 5.4. In this work, 10 slices are chosen for one periodic section along the axial direction. The underlying reasons for the proposed analysis are explained later in section 3. These slices are used to develop zone maps for axial compartmentalization. Each slice is indicated by its corresponding location along the X direction. Each of these slices consist of a filled region with particles present and an empty region where no particles are present. Based on these regions, the zoning strategy is performed wherein, the filled region is divided into zones highlighted by '+1', '0' and '-1', based on the averaged particle velocities of the individual grids being greater than 0, equal to 0 and less than 0. These user-defined tolerance limits are chosen for the compartment model, after balancing

the required model accuracy and associated computational cost. Further, the grids belonging to zone '0' also includes the dead zone, which can later be identified if the averaged particle velocity is equal to 0 for longer durations in the DEM simulation. This distribution of zones for all slices is used to construct zone maps for radial compartmentalization. Following the evaluation of zone maps along axial and radial directions, the compartmentalization step is performed to obtain inter-connected compartment maps. In this work, the compartment maps are developed after a steady state is reached in the periodic section, like that observed experimentally in the manufacturing line, wherein, the powders are added to the blender unit after reaching a steady state of operation. This approach ensures steady flowrates between compartments along time and can be further extended for dynamic compartment model if needed, to include the effects of start-up, shut down and step changes of process variables. Lastly, the inter-connection flowrate values are obtained between the compartments to develop inter-connected network of compartments.

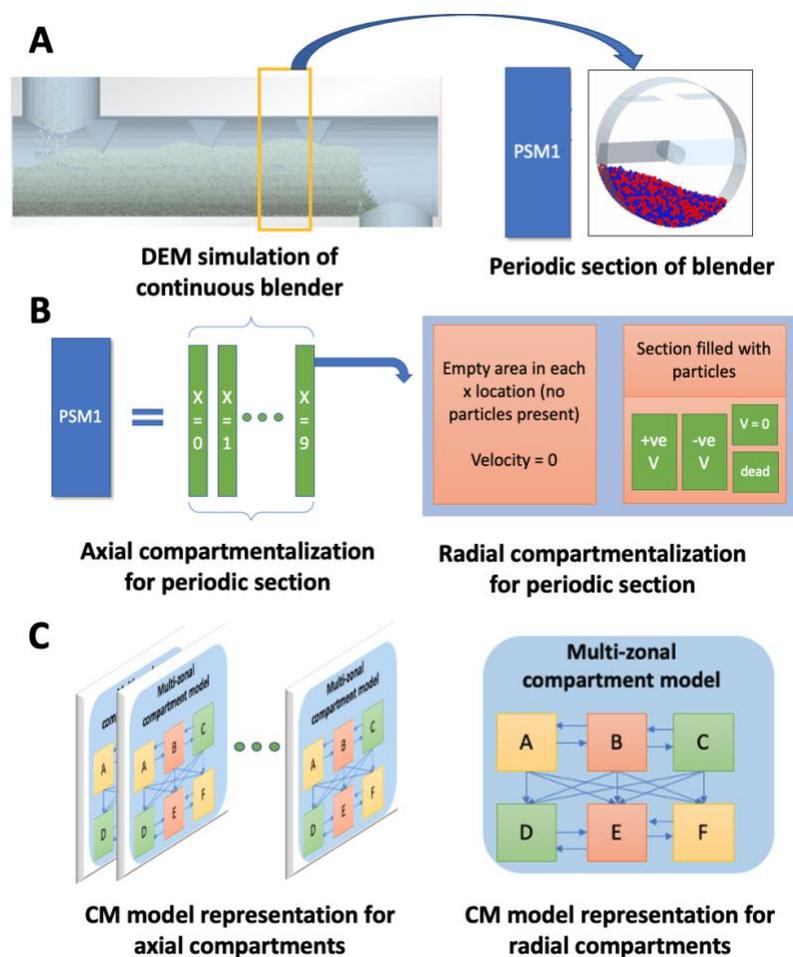


Figure 5.4: Proposed breakdown of periodic section - A) selection of periodic sections from DEM simulation of continuous blender, B) breakdown of periodic section for axial and radial compartments, C) compartment model representation for axial and radial compartments.

The proposed multi-zonal compartment modeling approach is developed in-house using open-source python language, followed by extraction of relevant datasets from the DEM simulation. The developed python code includes the evaluation of compartments and compartment maps for axial and radial directions along with evaluation of inter-connection flowrates. This is shared through GitHub code repository for public use on other applications and is available at <https://github.com/pbhalode/Multi-zonal-Compartment-modeling.git>. The total computational time for developing the inter-connected network of compartment is around 3 mins for the axial compartments and 10 mins for radial compartments, out of which most of the time is required for

loading the data-intensive information of variable distributions from DEM. Once the hybrid model is developed, the simulation is very quick in the order of few seconds and thus, can be easily integrated within process flowsheets for obtaining quick assessment of powder mixing within the blender unit.

5.3. Results and discussions

Based on the proposed breakdown of periodic section highlighted in section 5.2.3, this section aims at evaluating the compartment maps, followed by development of an inter-connected network of compartments for the periodic section along radial and axial directions. Before implementing the proposed strategy of compartmentalization, the empty region in the periodic section is identified for all slices where no particles are present. The evaluation of empty region is performed before the compartmentalization, to optimize the overall time required for evaluation of compartments by focusing only on the region which is filled with particles. This distribution of empty and filled regions is shown in Figure 5.5, for the 1st slice corresponding to $x = 0$. In this figure, the green colored region indicates filled region with particles present and red colored region indicates empty region with no particles. This distribution is obtained for each slice along the axial direction and each slice is further used for evaluating radial compartments.

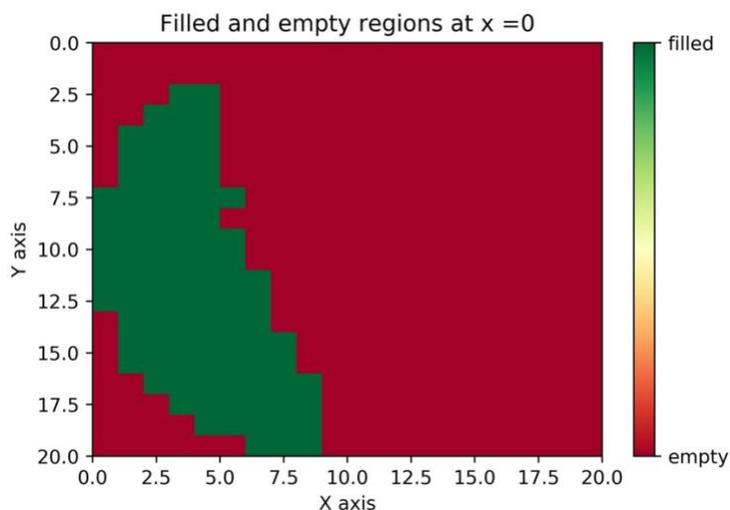


Figure 5.5: Distribution of filled and empty regions in the 1st slice of periodic section, corresponding to $x = 0$

5.3.1. Results for radial compartmentalization

Radial compartmentalization is performed for evaluating radial mixing in the system. For obtaining the radial compartments, zone maps are evaluated based on averaged particle velocities as shown in Figure 5.6A for Y direction and Figure 5.6B for Z direction corresponding to the slice belonging to $x = 0$ in the periodic section. In these zone maps, region corresponding to '1' indicates zones with particle velocities greater than 0, '0' with particle velocities equal to 0, '-1' with particle velocities less than 0 and '-2' indicating the empty zone with no particles. Following the evaluation of zone maps, the total number of spatially separated zones are identified, and compartment maps are developed. The compartment map developed along Y direction is shown in Figure 5.6C, where 7 individual compartments are identified for the slice corresponding to $x = 0$ within the periodic section. Similarly, the compartment map along the Z direction is illustrated in Figure 5.6D, where 6 individual compartments are identified.

After evaluating the compartment maps for Y and Z directions separately, these maps are overlaid together to obtain an overall compartment map for the radial direction. The overlaid map thus

demonstrates the combined distribution of compartments for both directions for quantification of radial mixing and can be used to develop the inter-connected network of compartments for radial direction. The overlaying of compartment maps is performed using the approach shown in literature [44,49], where the two compartment maps are superimposed on top of each other, and each section of the overlaid compartments is identified uniquely to distinguish it from other sections. The overlaid compartment map for radial direction is shown in Figure 5.6E, where the total number of overlaid zones equals 15 for the slice corresponding to $x = 0$. This varies along the periodic section for each slice with the total number of overlaid zones for the entire periodic section equal to 51. Following the evaluation of the overlaid compartment map for radial mixing, the next step focuses on quantification of flowrates between these compartments to obtain an inter-connected network of compartments for radial direction. The inter-connection flowrates thus indicate the extent of powder flow between these compartments. These flowrates $(F_{ij}(k))$ are evaluated as described in literature [44,48] and quantified using interfacial areas (A_{ij}) between different compartments (i, j) and averaged velocities of the grids in the compartments (v_i, v_j) along a specific direction (k) , as shown in Eq 1.

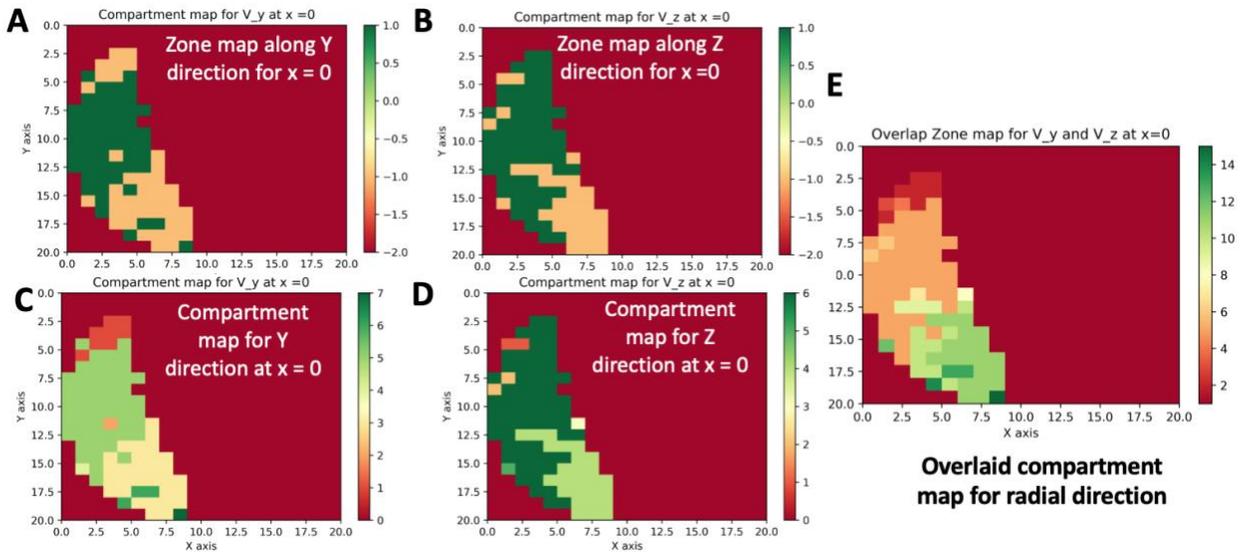


Figure 5.6: Radial compartmentalization for slice corresponding to $x = 0$ within the periodic section - A) identified compartments for $x = 0$ along Y direction, B) identified compartments along Z direction, C) compartment map for Y direction, D) compartment map for Z direction, E) overlaid compartment map for radial direction (including Y and Z directions)

$$F_{ij}(\mathbf{k}) = A_{ij} * (\mathbf{v}_i(\mathbf{k}) - \mathbf{v}_j(\mathbf{k})) \quad (1)$$

These flowrates are evaluated for all combinations of overlaid compartments, based on the interfacial areas for each combination of compartments and averaged velocities for each compartment. After obtaining the flowrates between different compartments, the periodic system is transformed into an inter-connected network of compartments, used for evaluating radial mixing.

5.3.2. Results for axial compartmentalization

A similar methodology is implemented for evaluation of compartments along the axial direction. The proposed zone maps are computed for each slice of the periodic section, wherein, the zones are obtained based on the averaged axial velocities along X direction of individual grids. Identified compartments are shown in Figure 5.7A for slices corresponding to $x = 0$ and Figure 5.7B for $x = 1$ where, '1' indicates zones with particle velocities greater than 0, '0' with particle velocities equal to 0, '-1' with particle velocities less than 0 and '-2' indicating the empty zone with no particles. Since the location of zones is different for each slice as shown in Figure 5.7, a slightly different

approach is used for the axial direction as compared to that used for radial direction. Here, the inter-connected network of zones is developed as shown in Figure 5.7C, with all zones ('+1', '0' and '-1') in one slice being connected to those in the next slice. Such connection is developed for all slices along the length of the periodic section. This connection is determined based on evaluation of flowrates using interfacial areas and axial velocities along the forward and backward directions for each slice, using Eq (1). Further, since the periodic section has periodic boundaries, the periodicity is extended to the slices at the boundary as well, i.e., for slices corresponding to $x = 0$ and $x = 9$. Thus, the inter-connected network of zones is developed for axial direction, to evaluate extent of axial mixing in the periodic section. Here, the flowrates evaluated for axial direction indicate a net forward flow of particles, as observed experimentally due to the forward movement of blades.

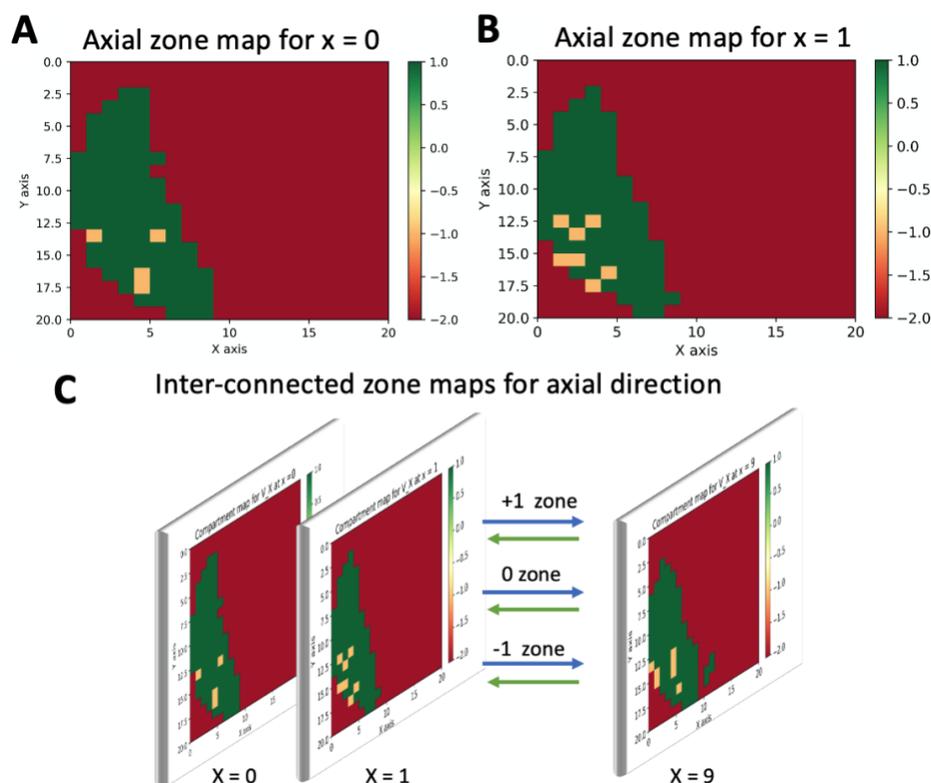


Figure 5.7: Axial compartmentalization for the periodic section - A) zone map for slice corresponding to $x = 0$, B) zone map for slice corresponding to $x = 1$, C) inter-connected zone map for axial direction

5.3.3. Validation of compartmentalization methodology

After implementing the proposed multi-zonal compartmentalization methodology for axial and radial directions, this section aims at validating the developed compartment models. The validation is performed by comparing the systemic information of the periodic section obtained from the developed compartment model with the original DEM simulation. This comparison is performed for quantification of mixing using residence time distribution studies. The respective systemic information is specifically chosen for comparison to ensure that the developed compartment model can preserve the important process knowledge despite the quick model computation in the order of few seconds, as compared to the original DEM simulation.

The validation is performed by comparing the residence time distribution (RTD) profiles of the proposed compartment model with the original DEM simulation. The RTD profile of the periodic section corresponds to the distribution of time that it takes for particles to traverse from one end of the periodic section to the other. For the original DEM simulation, all particles are taken into consideration and the RTD profile is obtained separately for the two particle types considered in the original DEM simulation, as shown in Figure 5.8. The dotted lines (red and blue) indicate one standard deviation from the mean values of both distribution profiles. As the compartment model involves averaged flow between different slices in the periodic section, an averaged value of residence time is obtained using averaged velocities obtained from the compartment model. The residence time thus obtained from the compartment model (RT_CM) is shown in Figure 5.8, which corresponds to the average time it takes for particles to travel the length of the periodic section. It can be seen from Figure 5.8, that the compartment model provides a similar value of averaged residence time as compared to that obtained from the DEM model, thus, validating the selected strategy for compartmentalization.

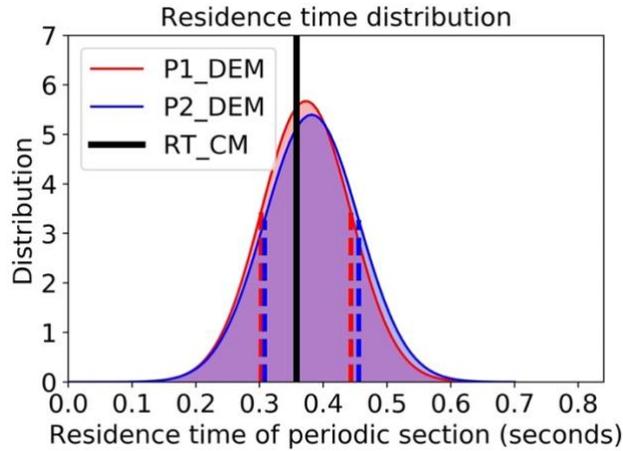


Figure 5.8: Validation of axial compartmentalization - RTD profile for DEM plotted for both particle types (P1_DEM, P2_DEM) and compartment model (RT_CM)

5.3.4. Applications of developed compartment model

After developing and validating the compartment model, it can be used for in-depth process analysis like evaluation of radial mixing. Quantification of radial mixing is important as it affects an important tablet CQA – content uniformity (CU) and is difficult to obtain experimentally [35]. The developed CM model resolves this problem by preserving the relevant information as observed in high-fidelity simulations, while providing quick process computations as observed in process models [57]. The developed compartment model is used to evaluate extent of radial mixing, where radial mixing is quantified using relative standard deviation (*RSD*) mixing index, as shown in Eq. (2).

$$RSD = \left(\frac{1}{C_{av}} * \sqrt{\frac{\sum_{i=1}^N (C_i - C_{av})^2}{N_{compartments} - 1}} \right) \quad (2)$$

Here, the $N_{compartments}$ corresponds to the total number of compartments in the periodic section, C_i represents the concentration of each compartment (i) and C_{av} represents the averaged concentration of all compartments in the periodic section. Using Eq. (2), the *RSD* profile for the compartment model is computed for the entire periodic section and smoothed using Savitsky

Golay filtering[43]. The obtained RSD profile is shown in Figure 5.9, where the RSD profile reduces gradually and stabilizes around a value of 0.57 indicating the stabilization of radial mixing in the periodic section. The RSD profile obtained using the CM model, is visually close to that obtained for the original DEM simulation, as shown in section 8.3.2. The developed CM model can thus provide an insight into the microscopic details of the particle system, as shown by the evaluation of radial mixing for the periodic section, like that obtained by the high-fidelity simulation models like DEM, with significantly lesser computational costs. It is important to note that since there is an inherent discrepancy of the grid sizes between the two models (CM and DEM), a numerical comparison of the microscopic behavior is not applicable, and thus the authors present the RTD profiles of the two models as a validation of the CM model approximation.

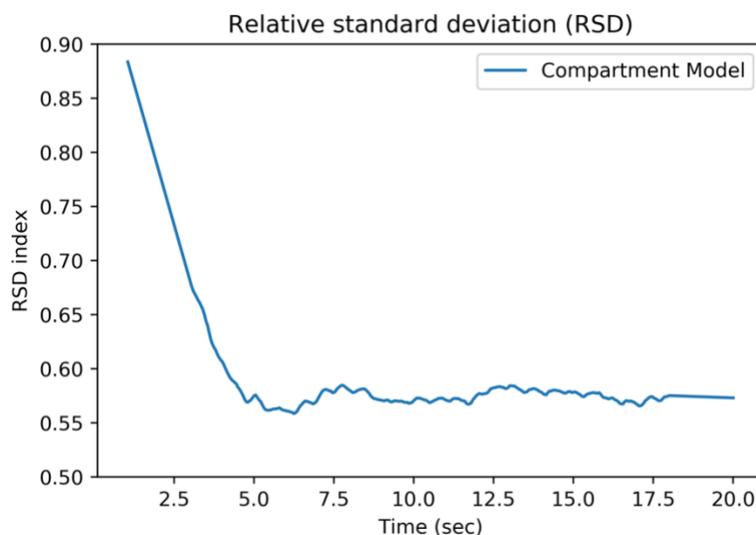


Figure 5.9: Relative standard deviation (RSD) profile for the periodic section as a function of time

Another application of the developed compartment model of the periodic section includes the extension of the developed model for the entire length of the continuous powder blender. Considering the chosen length of the periodic section and the length of the continuous powder blender, 8 periodic sections need to be aligned in series to replicate the process dynamics of the

entire blender [38,44]. Thus, the axial mixing for the entire continuous blender can be obtained by convolution of 8 periodic sections. This convolution ($RTD_{conv}(t)$) is performed using convolution operator, as shown in Eq. (8), where E_1 and E_2 are the two RTD profiles of the periodic sections to be convoluted. This can be represented using the convolution integral as shown in Eq. (9). After implementing the Eq. (9), the convoluted RTD profiles are shown in Figure 5.10 for 1, 2, 4 and 8 periodic sections, along with the respective residence times of the compartment model shown by dotted lines.

$$RTD_{conv}(t) = E_1(t) * E_2(t) \quad (8)$$

$$= \int_{\tau=0}^{\tau=t} E_1(t - \tau) \cdot E_2(\tau) d\tau \quad (9)$$

The convoluted RTD profile for 8 periodic sections thus, represent similar degrees of axial mixing as observed in the entire continuous powder blender.

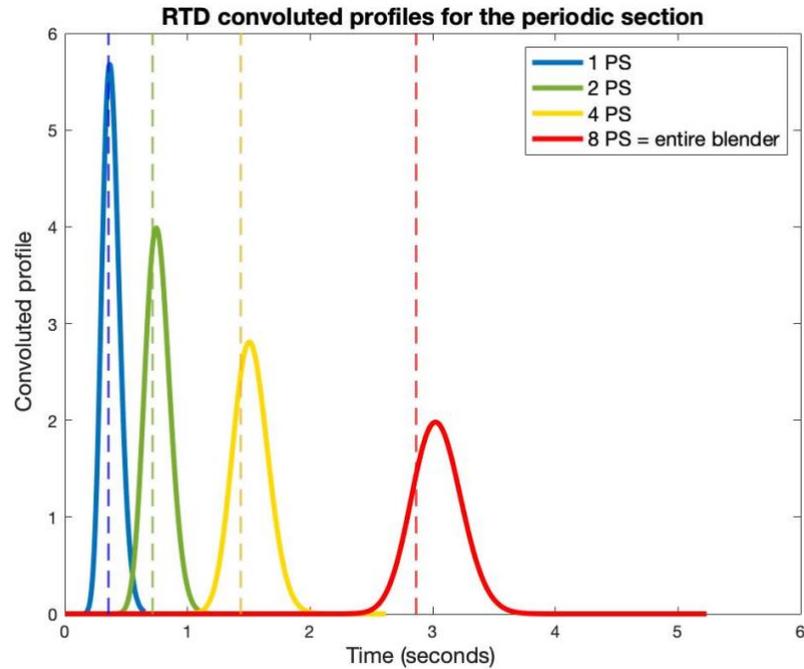


Figure 5.10: Axial mixing for entire continuous blender obtained by convolution of 8 periodic sections

Similarly, the degree of radial mixing observed in the periodic section can also be extended to that for the entire continuous blender based on the evaluation of RSD mixing index as previously suggested by Gao et al. [29]. Using this approach, RSD index is evaluated for the continuous powder blender along the length of the blender (x), as shown in Eq. (10). The model parameters RSD_{ss} , RSD_o and k_b are obtained by fitting the RSD index profile of the periodic section. The averaged velocity along with axial direction (v_x) can be quickly obtained from the developed compartment model and corresponds to 0.1117 m/sec. Thus, the RSD profile for the entire blender ($RSD_c(x)$) can be calculated as a function of axial location, shown in Figure 5.11.

$$RSD_c^2(x) = RSD_{ss}^2 + (RSD_o^2 - RSD_{ss}^2) * \exp(-k_b * v_x * x) \quad (10)$$

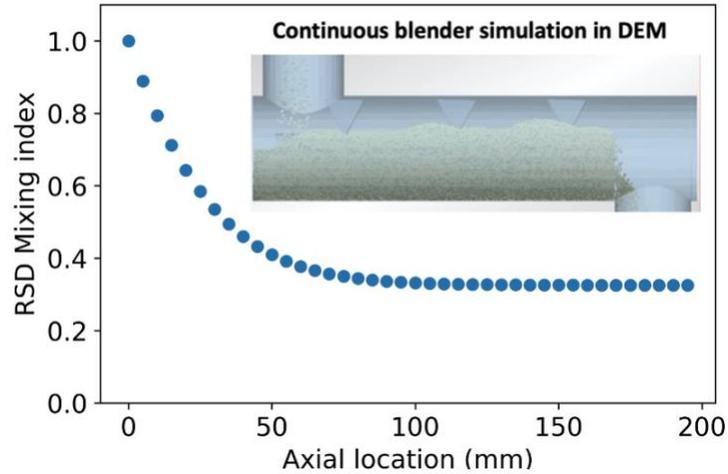


Figure 5.11: Extended RSD profile for the continuous powder blender

5.4. Conclusions and future work

The presented work highlights the development of hybrid model using multi-zonal compartment modeling strategy. The multi-zonal compartment modeling strategy aims at integrating the systemic or global level knowledge with local particle scale information, in a computationally efficient hybrid model. In the presented work, the hybrid model is developed for the continuous powder blender using DEM, to characterize overall powder mixing along axial and radial directions. The novelty of the proposed work stems from the application of DEM models to construct physics-informed, computationally-efficient hybrid models for quantification of overall powder mixing within continuous powder blender.

In this work, a DEM model is constructed for periodic section of the continuous blender. Following this, detailed mechanistic information of the particulate system i.e., particle velocities are extracted from the DEM model and utilized to identify different compartments or zones in the system. These compartments are then used to develop compartment maps, based on the observed process dynamics. Compartment maps are used to develop a simplified representation of the powder system in the form of inter-connected network of compartments. Following this, the inter-connection

flowrates between different compartments are quantified based on the particle velocities of each compartment and interfacial area between different compartments. This inter-connected network along with the flowrates act as a mechanistically informed hybrid model which can be used to replicate the observed process dynamics in a computationally efficient manner. Such hybrid models are developed along axial and radial directions separately to quantify the degree of axial and radial mixing. These models are further validated along axial direction using residence time distribution profiles and along radial direction using mixing indices like relative standard deviation index. Lastly, the presented approach is extended for the entire continuous blender. This approach thus allows to evaluate axial mixing in the powder blender along with quantification of radial mixing, which despite being an important CQA, has been challenging to quantify. In terms of the overall computational cost, the presented methodology provides significant time savings for developing and simulating the multi-zonal hybrid model, as compared to the original DEM simulation. Given such significant time savings coupled with model prediction accuracy, the developed hybrid model can be integrated within flowsheet framework for modeling of integrated manufacturing line.

Some potential avenues for advancing the proposed approach are presented here. The proposed approach assumes a specific selection criterion for evaluation of zones and compartments based on the particle velocities, to balance the required model accuracy and associated computational expense. The model accuracy can further be improved by selecting a finer compartmentalization criterion based on the end requirement of model application. Multiple variables of interest can also be included in the selection criteria based on the end goal of the model development. This strategy can be implemented for other complex systems incorporating cohesion and polydispersity of particles, which are prevalent in the continuous manufacturing line. The overall computational costs incurred in simulation of DEM models, construction, and simulation of developed hybrid multi-zonal compartment models can be improved through use of parallel computing strategies.

CHAPTER VI

6. Aim 3: Data pre-treatment and analysis of manufacturing lines

Acknowledgement of publication status:

Entire sections of this chapter are being prepared to be submitted for publication in a scientific journal written by the author of this thesis. This preparation is being completed under the citation – Bhalode, P.¹, Razavi, S. M.¹, Roman-Ospino, A.¹, Muzzio, F.¹, Ierapetritou M.^{2,*}, “Statistical data pre-treatment and noise removal for residence time distribution studies in pharmaceutical manufacturing”.

Affiliations:

¹ Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), Department of Chemical and Biochemical Engineering, Rutgers – The State University of New Jersey, Piscataway, NJ, 08854, USA

² Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, 19716, USA

* Corresponding author – email – mgi@udel.edu

6.1. Statistical data pre-treatment and denoising for residence time distribution profiles

6.1.1. Introduction

Residence time distribution (RTD) is a commonly used technique to evaluate the time that it takes for any material to traverse through the system under study [101, 197, 198]. Given that

RTD provides an effective technique to observe and understand the bulk fluid patterns within the system, it has been a fundamental tool in characterizing and developing predictive models of unit operations across various industries [101, 197-202]. Since its conception in 1935 [203], RTD studies have been widely used across various industries such as chemical, petrochemical, food, polymer, dyes, pharmaceutical industries for various engineering applications, including stirred tanks, unit operations, distillation, absorption and adsorption columns, microfluidic reactors, and wastewater treatment reactors [102, 152, 197, 198, 200, 204-209].

Within pharmaceutical manufacturing, RTD has also been widely used [10, 74, 102] for various applications ranging from characterization of unit operations and process flowsheets [47, 210-213] to material traceability, real-time quality assurance, process verification, and system health monitoring [46, 47, 214]. These applications are in line with U. S. Food and Drug Administration (FDA)'s vision for the 21st century [215] to enable manufacturing with minimal regulatory oversight, introducing the concept of Quality-by-Design (QbD) [26, 216-218]. QbD essentially encourages developing manufacturing lines where the product quality attributes are ensured in a timely manner and with less intensive product testing. Regulatory agencies, such as FDA [29, 30, 32-34], ICH [31, 35, 36], and ISPE [219-221] have published various guidelines for the industry, providing recommendations to adopt QbD within pharmaceutical manufacturing [218]. A critical component for integrating QbD within manufacturing lines involves developing an accurate and detailed process understanding of the manufacturing line, taking into consideration the effects of critical process parameters (CPPs) and critical material attributes (CMAs) on critical quality attributes (CQAs) of the drug product [218]. To develop this level of detailed understanding, various experimental and modeling techniques have been implemented [10, 12, 59]. RTD is one such technique that has been widely applied for developing process understanding, given the advantage of quick model assessment compared to other techniques [74, 102, 212].

RTD provides a probability distribution of residence time that a material resides within the system [197, 199, 200]. The distribution is obtained by tracking a traceable material, also known as a tracer. Generally, the tracer is selected to be one of the formulation ingredients (typically, API for pharmaceutical applications) or an external tracer material [102, 151]. For pharmaceutical applications, it is advised to choose a tracer material from an extensive material property data library that can replicate the flow properties of the API by using multivariate analysis techniques [222-225]. Recently, Razavi et al. [225] demonstrated that by carefully selecting a tracer that resembles the same physical properties of the API, the RTD profiles of the tracer and API are equivalent under different bulk flow properties. Along with this, it is also important to ensure that the tracer material is easily detectable using PAT techniques (such as commonly used vibrational spectroscopic techniques including Raman and Near Infrared (NIR)) and has a signal that is distinguishable from the bulk material [151, 226] to ensure the obtained RTD can accurately represent the system. Escotet et al. [226] demonstrate the importance of selecting a tracer with differences in spectral data compared to the bulk material by comparing the corresponding concentration profiles determined by NIR for two different tracers (Acetaminophen and Maize Starch) for the same bulk material (Avicel 102). It is observed that since the spectra of Maize starch closely resemble the spectra of Avicel 102, it becomes challenging to identify the tracer concentration accurately and results in a very noisy profile, shown in Figure 6.1A. On the other hand, the spectra of Acetaminophen being distinguishable from the bulk material can provide accurate RTD profiles with a significantly high signal-to-noise ratio, as shown in Figure 6.1B. Thus, it is crucial to select a tracer that is inert to the system, detectable, does not stick to the walls or agglomerate within the unit, and does not alter the flow properties of the system under study [151, 197, 199, 200].

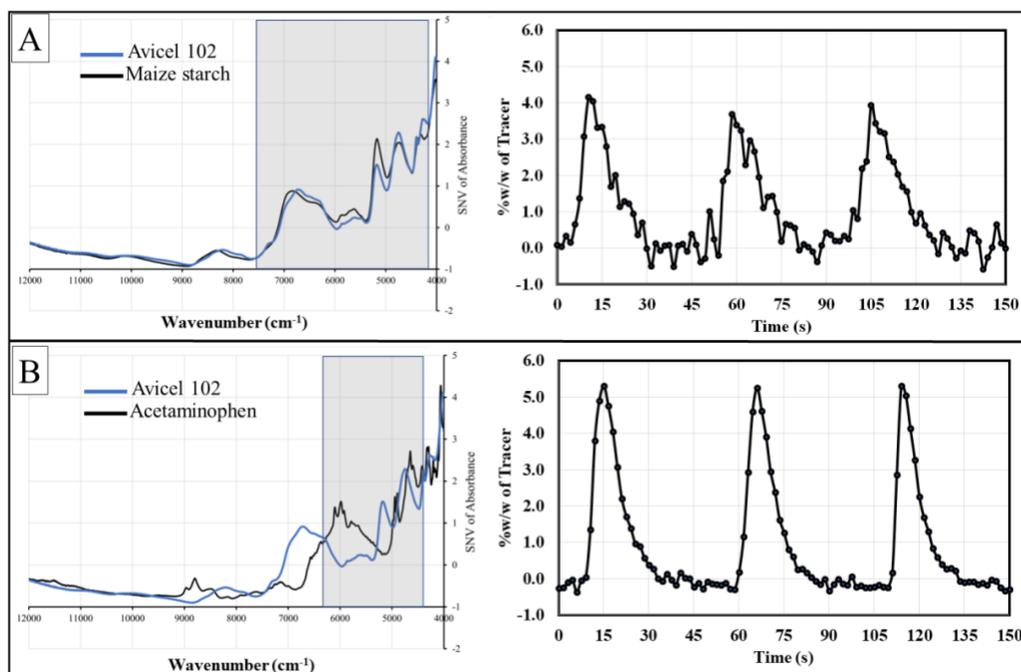


Figure 6.1: Comparison of different tracers for the same bulk material using the NIR spectra and the obtained tracer concentration profiles (Images taken with permission from Escotet et al. [48])

Following the tracer selection, the tracer material is added at the inlet of the system and is tracked at the outlet as it exits the system. Commonly used input profiles for tracer addition include step, pulse, and variable step input [197, 199, 200]. Step input generally involves a step-change in the concentration of one of the formulation ingredients (selected as tracer) at the inlet of the system at a time t . Pulse input provides an impulse of tracer at the inlet whereas, variable step input provides a tracer profile consisting of varying step-up and step-down profiles [151, 214]. These input profiles are discussed in depth along with their advantages and disadvantages in literature [102, 151]. After the tracer is injected near the inlet, the tracer concentration profile is measured at the outlet of the system using process analytical techniques (PAT) [102, 227]. Amongst the PAT techniques, vibrational spectroscopy is generally employed for tracer detection in the pharmaceutical industry, as these techniques do not require sample preparation [102], and are thus suitable for real-time measurement of tracer concentration [29].

To evaluate the tracer concentration profile, it is important to develop appropriate data acquisition and calibration strategy that can capture the changes in tracer concentration using spectroscopic sensors at desired frequency and signal intensity to accurately capture the tracer concentration profile [13, 210]. The calibration model is recommended to be developed for varying levels of tracer concentration to ensure correct measurement of tracer concentration [40]. The concentration range selected for the calibration studies depends on the processing conditions and the limit of detection for the selected tracer material [228]. Thus, multiple calibration runs are performed using a selected design of experiments accounting for various sources of variability, including process, environmental, and instrumental variability [229, 230]. Normalization methods such as standard normal variate (SNV) and multiplicative scatter correction (MSC) and their combinations with 1st and 2nd order derivatives are implemented as part of data pre-treatment strategies to develop the calibration model [102]. These strategies are tested using statistical methods, including root mean square error of cross-validation for varying number of selected latent variables, to ensure the spectral information of the tracer to be the main contributor to the observed variability [229, 231]. During the measurement of tracer concentration, it is ensured that the desired level of measurement frequency of the tracer concentration is maintained. Higher measurement frequency allows capturing the tracer concentration more precisely including all the variabilities of the tracer concentrations, whereas a low measurement frequency results in smoother concentration profiles but can miss out on capturing important RTD information of the system [227, 232]. To avoid loss of information, a higher level of measurement frequency is desired, but this can lead to noisy RTD profiles which may not provide a good signal to noise ratio. A general tendency to circumvent this problem is to increase the tracer concentration in the system to ensure the tracer is accurately captured. However, this can be detrimental to the RTD study if the tracer flow properties differ from the bulk material and can potentially alter the original system. On the other hand, the tracer amount should be enough to ensure that it is detected accurately at the outlet of the system. Thus, the appropriate tracer amount needs to be defined based on the selected tracer, specific bulk

material, and the system under study [102, 151, 225]. After finalizing the required amount of tracer material, powder blends consisting of different tracer concentrations and the bulk material are prepared, and calibration runs are performed. Partial Least Squares (PLS) models are then developed to relate the spectroscopic data to the tracer concentration values, and calibration models are established for RTD studies.

Despite such detailed analysis in tracer selection and developing accurate calibration models to measure tracer concentration profiles for RTD studies, significant level of variability is observed in the measured tracer concentration profiles, as highlighted in literature [152, 225, 226]. The observed levels of variability, leading to noisy tracer concentration profile, can potentially hinder accurate analysis of RTD and subsequently impact process decisions based on RTD studies, leading to drug product recalls. Hence, it is imperative to handle the levels of variability and noise observed in RTD profiles.

Data pre-treatment strategies, including noise characterization and denoising, are thus crucial for RTD studies to quantify and selectively remove the noise from the measured tracer concentration profiles, thus providing a methodology to obtain a true representative RTD of the system under study [102]. Though necessary for accurate characterization of RTD, the literature lacks a detailed analysis and description of the data pre-treatment strategies, making it challenging to replicate or utilize the RTD profiles for important applications such as material traceability and real-time quality assurance. The proposed article aims to address this concern of lack of detailed discussion regarding RTD noise characterization and denoising strategies. In the presented work, the authors focus on discussing the characteristics of noise present in tracer concentration profile along with the application of various denoising strategies for RTD studies. Furthermore, performance metrics are devised to compare the denoising methods quantitatively. This article aims to serve as a guide for researchers in the field to understand the implications of different denoising

strategies on the RTD profiles and subsequently select the most appropriate method for respective applications based on the performance metrics. The chapter is organized as follows. Section 1 provides an introduction of the problem to be addressed and discusses the novelty of the article. Section 2 describes the experimental campaign for obtaining the tracer concentration datasets used in this work. Section 3 focuses on understanding noise and denoising from a signal processing perspective and discusses the denoising strategies investigated in the presented work. This is followed by introducing performance metrics to compare the denoising strategies. Section 4 presents the results of the denoising strategies on the experimental datasets, along with the evaluation of the performance metrics for comparison. Following this, the impact of denoising on important RTD applications such as material traceability and product quality assurance are discussed. Lastly, section 5 outlines the conclusions for the presented work on applying noise handling strategies for RTD studies.

6.1.2. Experimental studies

The experimental strategy used in this work is summarized in this section, where a pulse input method is used for the RTD studies. The details of the experimental work are described in depth in a recent article from the same research group [225], and a summary is outlined below. A single K-Tron KT20 (Coperion K-Tron Pitman Inc., Sewell, NJ) feeder fitted with a coarse auger screw configuration and a C-type motor is used to maintain a steady flow of base powder into the continuous mixing system. A Gericke GCM 250 mixer (Gericke USA, Somerset, NJ) is used as the blending system with alternate blade configuration operating at a rotational speed of 100 RPM. A weir (semicircular disc) is placed at a 60° angle at the exit of the mixer to control powder hold-up. A long vibratory feeder is placed at the exit of the blender to dispense the powder and provide a homogenous sample presentation for the PAT. A Bruker Optics (Billerica, MA) Matrix Fourier Transform Near-Infrared (FT-NIR) spectrometer is installed on a platform above the vibratory

feeder, as shown in Figure 6.2. The throughput of the system is maintained at 12 kg/hr. In this work, only one excipient powder (Avicel PH301) is used as the base powder.

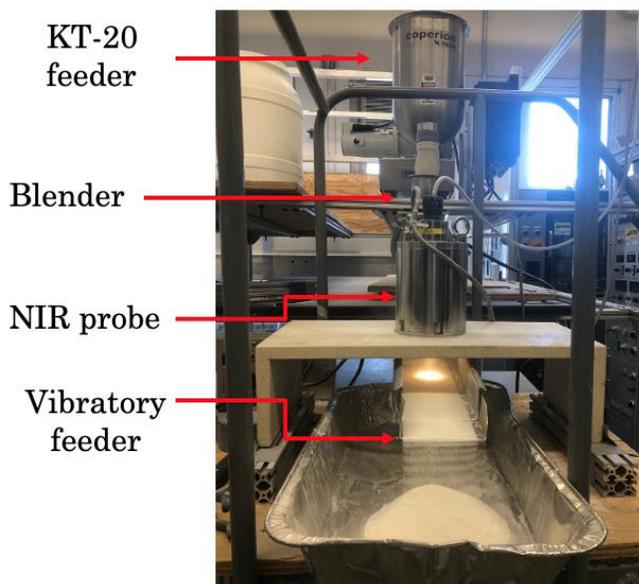


Figure 6.2: Experimental setup including feeder, blender, NIR probe, and vibratory feeder

Following the experimental setup, NIR calibration model is developed, wherein characteristic spectral bands from bulk material and tracer are identified to determine the appropriate spectral range for quantification and calibration models are constructed by testing seven data pre-treatment methods [102, 228], and the best method is determined based on root mean square of cross-validation (RMSEV) by comparing it with explained variation. Calibration spectra is acquired to incorporate sufficient variation within the tracer concentration as it exits the system and thus, calibration experiments are conducted using the same setup and acquisition parameters as the RTD runs. Calibration blends of 0.0 – 9.0 % w/w in steps of 3.0 % w/w are used [228] covering a significant range of tracer concentration. The experimental campaign includes three repeats of the RTD experiment with different tracer amounts (10g, 7.5g, 6g, and 5g) to evaluate its effects on the tracer concentration profiles, leading to a total of 12 experimental runs. For robust

PAT method, the calibration model is built and used for executing the RTD experiments on the same day. The experiments with 6g tracer amount were conducted later resulting in concentration profiles following a different trend in terms of peak intensity, as it is susceptible to environmental and minor setup changes. Nevertheless, all the $E(t)$ profiles overlapped on top of each other and the mean residence time for all the cases showed no statistical significance. Although the PAT models need to be recalibrated often, practitioners may not do it routinely as it is tedious and sometimes costly. This can be another source of variability that is considered in our analysis. The measured tracer concentration profiles for different tracer amounts are shown in Figure 6.3 along with the repeated runs.

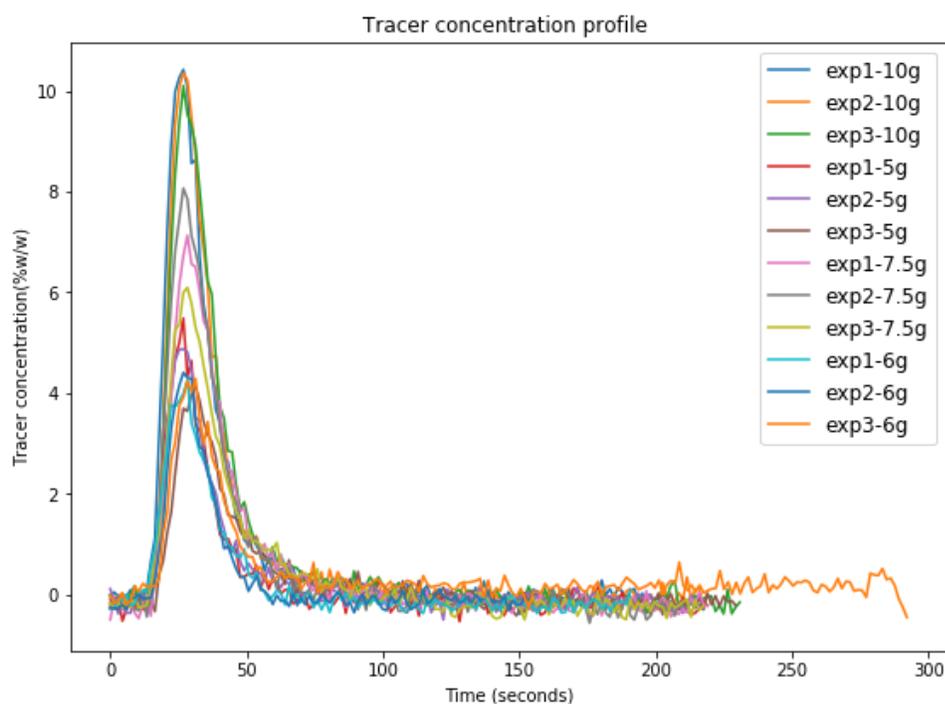


Figure 6.3: Tracer concentration profiles for different tracer amounts with repeated runs

6.1.3.Methods

6.1.3.1. Signal processing – signal and noise

A time-series dataset has two components – signal (desired or important part) and noise (undesired or unimportant part) [233-235]. A crucial problem in the signal processing literature is identifying the true underlying signal from the noise component to correctly analyze the general trend of the dataset and evaluate the degree of confidence of the measurements [233, 236]. Evaluating the extent of noise is relatively easier for datasets that are generally plateaued around a fixed mean value or for situations where the true underlying form of the signal is exactly known. However, characterizing the noise becomes challenging if the underlying signal is of a specific shape or form, such as the tracer concentration profiles [233, 236]. Along with noise characterization, it is also important to consider the limit of detection, which corresponds to the lowest measurement of the signal that the measurement device can reliably capture despite the level of noise present in the system [228]. Roman-Ospino et al. [228] presented an interesting study illustrating probability distributions associated with measurements of low tracer concentrations in a feed frame, thus indicating the difficulty in determining the low concentrations accurately. An alternative way to evaluate the true signal is using repeated measurements of the RTD signal [152, 226], which are used to calculate an averaged profile, known as ensemble averaging [233]. However, obtaining repeated measurements of the RTD experimentally, can become quite laborious and time-intensive, and thus this approach is not always recommended or feasible.

Thus, noise handling and denoising methods are devised to address the fundamental problem of distinguishing the true underlying signal from the noise. These methods focus on smoothing the dataset by removing sharp, abrupt fluctuations in the time-series data to provide a smoother signal which can better capture the underlying trend [233, 235]. Smoothing acts as a low pass filter, which allows signals with low frequencies to pass through and attenuates the noise components associated with high frequencies in the dataset. Before discussing the denoising methods, it is important to

reflect on situations where denoising or noise handling and removal can be important for RTD studies, as outlined below [233],

1. it is important to observe the long-term trend instead of short term
2. the signal is too noisy to correctly identify the important features of the RTD profile, including the height of the peak, the width of the profile and area under the curves
3. the noisy RTD signal is to be used for post-processing such as identification of peaks

For situations where the RTD dataset contains some outlier points (predominantly seen in sensor measurements), it is important to address the outlier detection and removal before implementing noise handling strategies [233], as outlier points, if any, can significantly skew the trend of the dataset during denoising as well as impact model fitting leading to erroneous outcomes [233]. Thus, outlier detection and removal are pre-requisite for denoising and signal processing. Some filtering techniques like median filters [237] can be used to handle outliers in the dataset, using linear interpolation between the points before and after the outlier point, and can thus restore the original trend of the dataset.

6.1.3.2. Noise characterization for RTD studies

During RTD experiments, various sources of noises are captured along with the underlying desired signal. These sources include and are not limited to equipment, environment, experimentation, and human noise [151, 152, 233]. Equipment noise corresponds to the noise due to vibrations of moving parts in the equipment if not balanced or handled properly. Environmental factors include numerous sources, including neighboring activities and vibrations from those activities. The experimentation itself is an important source for unseen errors, comprising material handling and flow, correct positioning and location of sensors, and vibrations that arise during the experiment. Human noise, such as human movements around the equipment, human error during experimentation and tracer introduction, calibration measurement, and material handling, is also a

key component contributing to the overall noise captured in the datasets. For tracer concentration profiles described in section 2, the noise levels captured in the profiles are shown in Figure 6.4. These profiles are first converted to $E(t)$ profiles, followed by taking the absolute difference between a least-square curve fit using a tanks-in-series model and the $E(t)$ profile. The observed types of noises can be identified based on the frequency distribution of the noise and are generally referred to by colors [233, 236, 238, 239]. For a uniform noise distribution, the noise is referred to as ‘white’ or uniform noise. Pink noise has an inverse proportionality between noise and frequency (often referred to as $1/f$ noise) and is relatively difficult to remove compared to white noise. Other types of noises include red noise or Brownian noise, which has the noise proportional to the square root of the frequency, and the blue noise, which has higher power at higher frequencies. It can be seen from Figure 6.4 that there exist different types of noise in the RTD profiles - proportional noise observed near the peak regions of the profiles, inversely proportional or pink noise, and white or constant noise, which is present along the entire duration of the profile. Based on these types of noise, specific denoising methods can be devised that provide higher noise removal [233, 236].

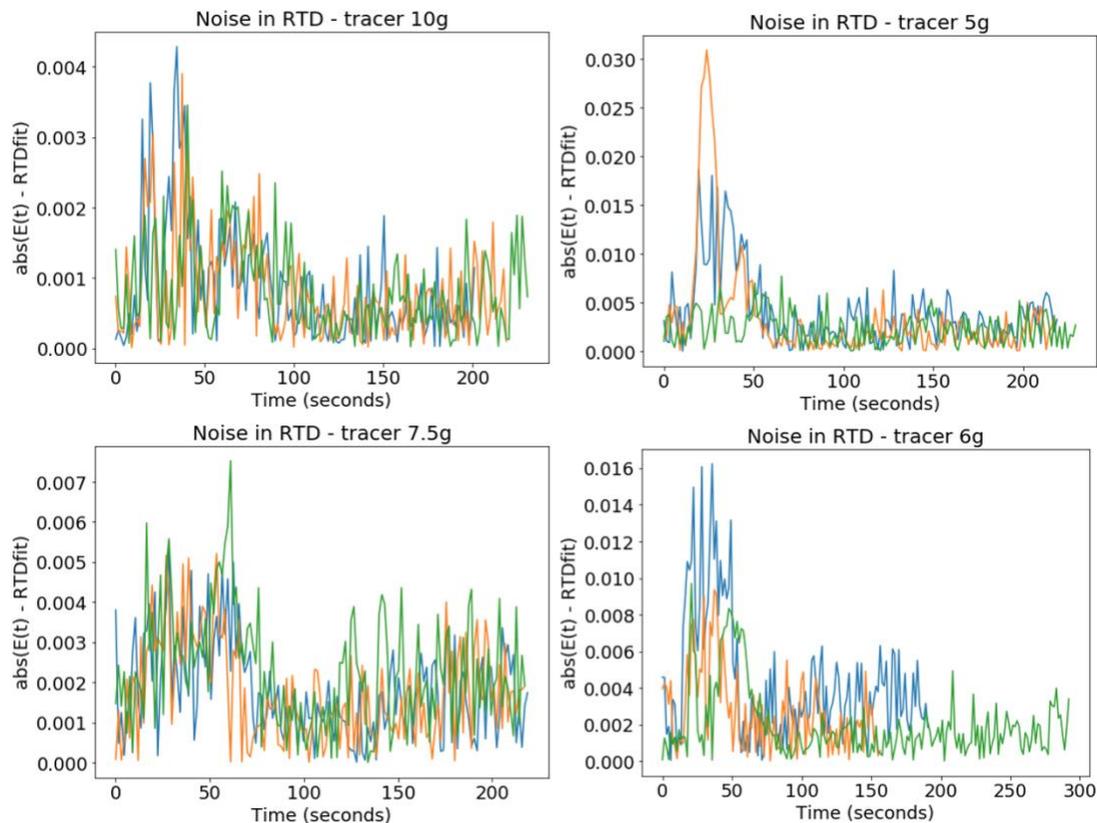


Figure 6.4: Noise present in tracer concentration profiles with different tracer amounts shown in each sub-plots, where colors (orange, green, and blue) are used to show the noise in repeated concentration profiles

6.1.3.3. Denoising methods

There exist various denoising methods in the literature that provide noise reduction in the dataset. It is, however, important to keep in mind that the degree of smoothing can have a significant impact on the important RTD features, such as the height and width of the profile. The higher the degree of smoothing, the greater is the noise removal, but greater is the possibility of signal distortion due to smoothing [233]. Thus, the degree of smoothing should be optimized considering the type of signal, the height of the peak, the width of the profile, and the total number of data points in the dataset.

In this work, we compare time averaging (sliding average and Savitsky Golay filtering) and frequency averaging techniques (Fast Fourier transform and wavelet transform) to evaluate appropriate denoising methods for RTD studies, based on the degree of smoothing and retention of important RTD features.

6.1.3.3.1. Sliding average (SA)

Sliding average (SA), also known as moving or rolling average, is one of the simplest smoothing algorithms [233, 238], involving unweighted averaging of specific number of data points. Commonly used with time series, this method involves replacing noisy data points by averaged values of adjacent M points in the dataset, where M corresponds to the averaging window, thus capturing the overall trend of the dataset [233]. SA is commonly used in signal processing literature and is calculated as shown in Eq. (1).

$$SA_t = \frac{(x_t + x_{t-1} + x_{t-2} + \dots + x_{M-(t-1)})}{M} \quad (1)$$

Using Eqn. (1), SA at time t (SA_t) is calculated as the unweighted average of M data points obtained prior to time t . The averaging window M is chosen based on the desired degree of smoothing vs. the retention of accuracy in the smoothed dataset. The higher the averaging window, the smoother the data, but the lower the accuracy of the retained dataset will be. For the work presented in this article, a preliminary analysis is conducted to evaluate the desired value of M , wherein three values (3, 5, and 10) are tested, and a medium value of 5 is chosen, as it preserves the relevant information while denoising the dataset. A comparison of different values of M for the tracer concentration profile is shown in Figure 6.5.

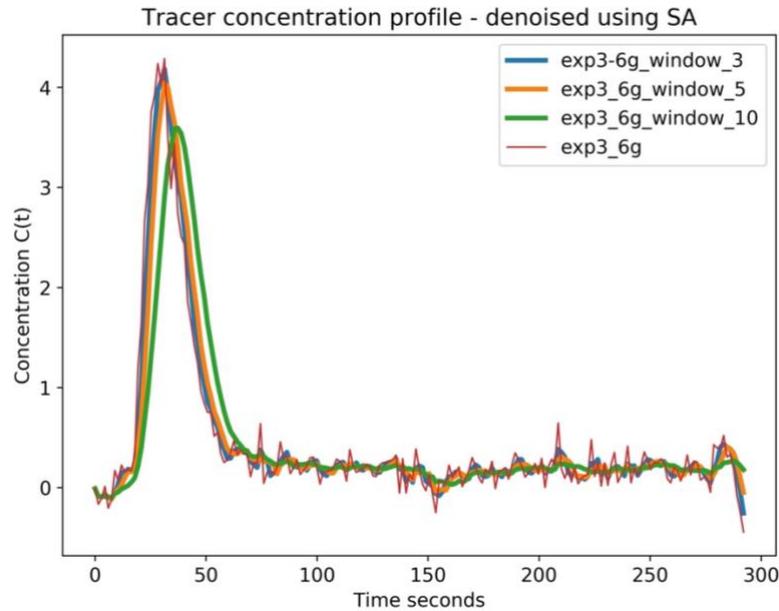


Figure 6.5: Effects of selected smoothing window size on tracer concentration profiles for sliding average (SA) method

SA method can be implemented in python using pandas module [240], as listed below - `(pandas.Series(C1).rolling(M).mean())` where, $C1$ corresponds to the noisy dataset to be smoothed. Alternatively, MATLAB includes the module `movmean`, including parameters as $C1$ and the smoothing window (M) [241]. Other variations of SA include cumulative moving average (CMA), weighted moving average (WMA), and exponential moving average (EMA) methods [233-236]. CMA considers all time points up to the current time t to obtain a smoothed dataset, focusing on capturing the overall trend. WMA assigns specific weights to the time points for obtaining the denoised dataset. Lastly, EMA assigns weights to the data points such that the recent points are assigned more importance than the older data points. Thus, EMA reacts to the changes and fluctuations in the data faster than SA, which seems to be more efficient for fluctuating datasets like RTDs; however, it may not provide desired levels of smoothing for the datasets.

6.1.3.3.2. Savitsky Golay (SV) filtering

Developed and popularized in a seminal paper by Abraham Savitsky and Marcel Golay in 1964 [242], SV is a filtering method wherein the data points are smoothed by fitting subsequent data points with a low-degree moving polynomial using linear least squares [233, 236, 243, 244], which is then used to replace the noisy dataset. Thus, each value of the noisy data is replaced by the value of the polynomial model fit to neighboring $2n + 1$ points with n being greater than or equal to the order of the polynomial. SV is a commonly used filtering method in chemometrics, signal processing, and spectroscopy literature [233, 236, 238, 239, 242, 245]. SV is applied as shown in Eq. (2), where y_{j+1} is the observed value and Y_j is the denoised SV filtered value at specific time points. C_i corresponds to the set of m convolution coefficients chosen based on the selected order of the polynomial model.

$$Y_j = \sum_{i=\frac{1-m}{2}}^{\frac{m-1}{2}} C_i y_{j+1} \quad \text{where, } \frac{(m-1)}{2} \leq j \leq n - \frac{(m-1)}{2} \quad (2)$$

The parameters for SV, including the smoothing window size (indicating the number of powders used to fit a polynomial function) and the order of the polynomial, are investigated using preliminary results and the values that provide the best level of smoothing are chosen. Figure 6.6A shows the effects of order of the polynomial on the selected RTD data whereas, Figure 6.6B shows the effects of the smoothing window size selected for the polynomial. It is observed from Figure 6.6A that the order of 5 tends to overfit the original noisy data near the baseline, whereas a value of 1 does not capture the RTD profile correctly, especially the peak and the width of the profile. Thus, an intermediate value of 3 is chosen, which provides a good balance between accuracy and denoising. Similarly, Figure 6.6B shows the effect of smoothing window sizes, where higher values (21 and 31) do not capture the important features of the RTD profile, whereas a lower value of 5

overfits the noisy data, especially near the baseline. Thus, an intermediate value of 11 is chosen as the smoothing window size for the SV method.

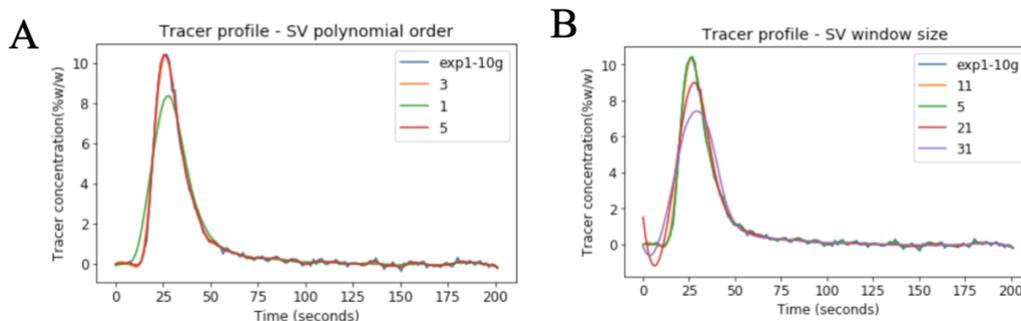


Figure 6.6: Effects of SV parameters on tracer concentration profile - A) effect of polynomial order, B) effect of smoothing window size for polynomial fit

In the presented work, SV is implemented on the noisy dataset using python module `signal.savgol_filter` from the `scipy` parent module as listed here – `signal.savgol_filter(C1, n, m)` [246]. Here, $C1$ corresponds to the noisy data that needs to be filtered or smoothed and smoothing window size (n) and order of polynomial (m) is selected to be 11 and 3, respectively. Alternatively, MATLAB uses the function `sgolay` for performing SV filtering and has similar inputs as listed above for the python module [241].

6.1.3.3.3. Fast Fourier transform (FFT)

Fourier transform involves transforming a signal from time to frequency domain, decomposing the signal into multiple component frequencies contained in the original signal. This mathematical transformation, after its conceptualization in the early 1800s [247] and popularization around the mid-1900s [248], has been widely applied in various fields of science, mathematics, signal processing, and engineering [249-251]. Fast Fourier Transform (FFT) is an algorithm that provides the Discrete Fourier transform (DFT) of a signal, which involves decomposing the signal

into various frequency components, and requires a time of the order of $O(N^2)$ where, N is the data size. FFT reduces the computational requirement to the order of $O(N \log N)$ [233]. This is beneficial for cases where N is large, leading to significant time savings in computing the Fourier transform. FFT allows order reduction as it decomposes the DFT of the signal into multiplication of sparse factors. Figure 6.7 shows the different steps associated with denoising a signal using the FFT method. FFT is commonly used as a low pass filter by first calculating all the frequencies contained in the signal and removing the frequencies associated with the noise component within the original signal to obtain a low pass filter in the frequency domain. Lastly, the signal is transformed back to the time domain by performing Inverse FFT (IFFT), resulting in a denoised signal. FFT is obtained using the `scipy.fftpack` module in python, which includes functions such as, `fftpack.fft` and `fftpack.ifft` for calculating FFT and IFFT of the signal [246]. Similarly, MATLAB uses `fft` and `ifft` functions for obtaining the FFT of the original signal and inverting it back to the time domain [241].

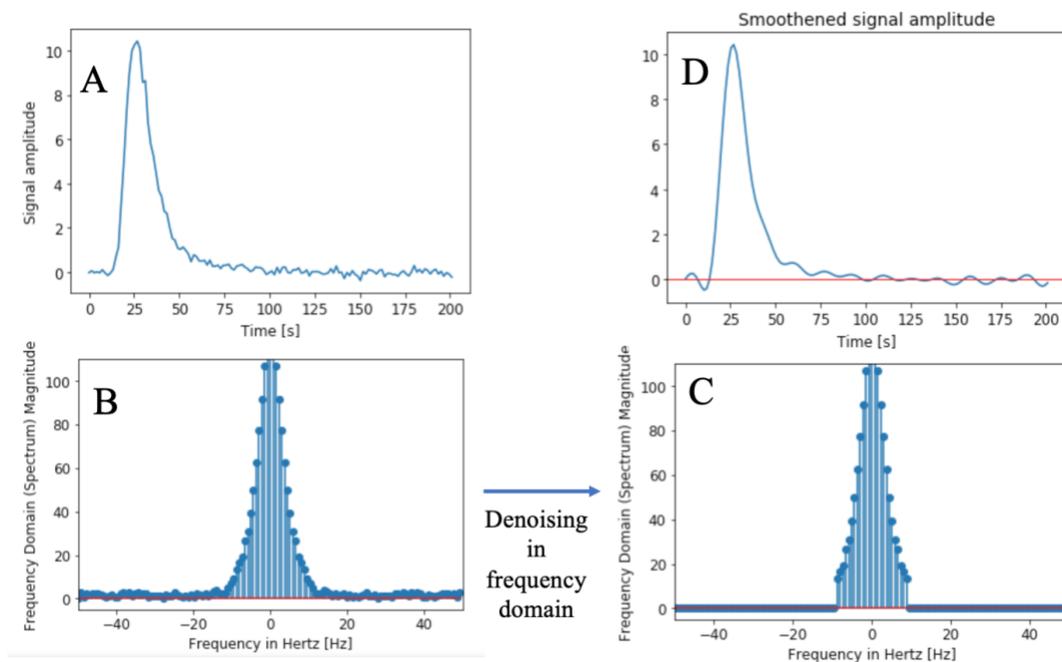


Figure 6.7: Signal denoising using FFT - A) original noisy signal in time domain, B) frequency domain of noisy signal, C) denoising of noisy signal in frequency domain, D) denoising signal in time domain

6.1.3.3.4. *Wavelet transform*

Though widely used as a promising method for denoising, FFT is best applicable for scenarios where the frequencies present in the signal are not time-dependent and are stationary along time. However, most of the signals that are encountered in nature are not stationary and have a time-dependent frequency spectrum associated with the signal. Some examples include ECG signal [252, 253], equipment or sensor data [254], fault detection [255], and stock markets [256]. This is valid for RTD applications as well, as the degree of noise associated with the signal varies over time, and it does not have constant white noise but a combination of white, pink, and proportional noise (as observed in section 2), which are time-dependent. Thus, for dynamic systems where the frequency spectrum varies along time, it is important to record the frequencies present in the signal as a function of time. FFT fails to capture these dynamics as it provides detailed resolution in only the frequency domain but zero resolution in the time domain.

For such systems, wavelet transform (WT) is a powerful method for the analysis of time-series datasets [233, 257]. WT method can provide a high degree of resolution in both time and frequency domains, as it has the dual capability to capture small frequencies (with high resolution in frequency domain and low resolution in time domain), as well as large frequencies (with low resolution in frequency domain and high resolution in time domain). The stretched-out part of the wavelet captures the main shape of the signal whereas, the compressed part of the wavelet captures the sharp changes present in the signal. Amongst the different types of wavelet families that exist with different shapes and compaction levels for denoising of datasets, the most appropriate wavelet family is chosen. These families include discrete wavelets such as Haar, Daubechies, Symlets, Coiflets, Biorthogonal and continuous wavelets such as Mexican Hats, Morlet wavelet, complex Gaussian wavelet [257-261]. Within each wavelet family, there exist sub-categories based on the level of decomposition, refinement, and the number of coefficients.

Discrete wavelet transform (DWT) is used for the studies highlighted in this article, as DWT has discrete nature in frequency domain and is applicable to dynamic systems, i.e., with continuous nature in time domain. DWT method acts as a filter-bank that decomposes the signal into a series of high-pass and low-pass filters [233, 262]. For WT, frequencies are defined in terms of scales, such that high frequencies correspond to small scales. The first step starts with analyzing high frequency behavior. The second step increases the scale by a factor of 2, and the signal is analyzed around half of the maximum frequency. The third stage increases the scale by a factor of 4 and the signal is analyzed around a quarter of the maximum frequency. These steps are continued to reach the desired level of decomposition. Along with these steps, the original signal is transformed into a series of signals with different frequencies. Following this transformation, the denoised signal can be obtained by filtering out the high frequency region of the signal, controlled by a threshold rule such as soft, hard, and garrote rule [263]. Denoising is performed to a minimum of the selected threshold value or $\text{floor}(\log_2 N)$ where, N is the number of samples in the dataset [264, 265]. For the denoising work presented in this article, the DWT is performed using open-source wavelet transform software PyWavelets in python [261]. MATLAB includes a wavelet toolbox including signal denoising function – `wdenoise` is used for obtaining denoised RTD profiles with level set to its default value [264-266]. Different wavelet families and sub-categories are compared with the noisy RTD profile to ensure the most appropriate family and sub-category is selected for RTD

denoising. Figure 6.8 shows a comparison of different wavelet families, where discrete wavelet families are compared to identify the best candidate.

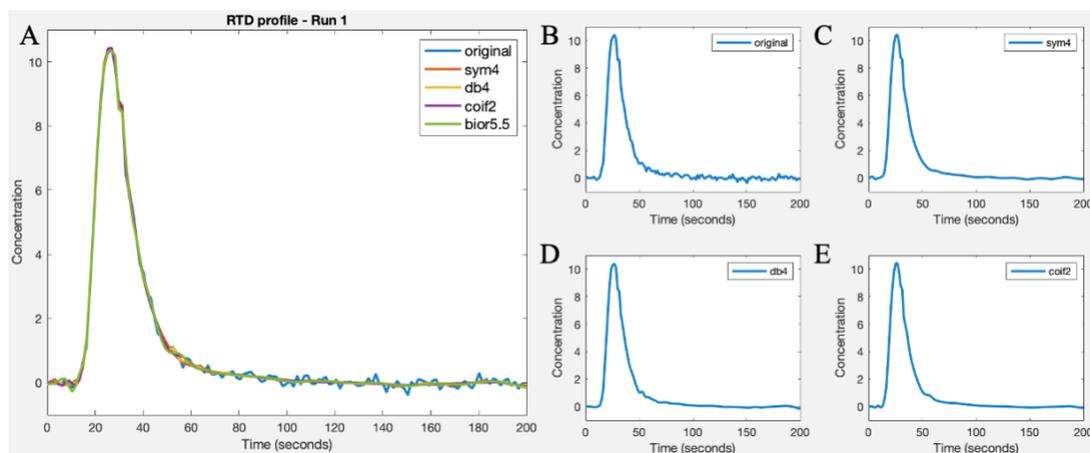


Figure 6.8: Comparison of discrete wavelet families on tracer concentration profiles - A) original and denoised concentration profiles, B) original concentration profile, C) denoised profile using symlets, D) denoised profile using daubechies, E) denoised profiles using coiflets

Based on Figure 6.8A, the wavelet families - symlets, daubechies, and coiflets perform fairly close to each other, as shown in Figure 6.8C-Figure 6.8E, while bior wavelet adds some variability at the start and end of the profile. For the work presented in this article, Symlets wavelet (sym4) is used for denoising tracer concentration profiles using wavelet transform, as it provides smoother profiles in comparison to the daubechies and coiflets.

6.1.3.4. Comparison of denoising methods

It is important to develop a quantitative comparison of the denoised profiles to evaluate the most appropriate denoising strategy. This sub-section aims at developing comparison metrics for denoised tracer concentration profiles. These metrics are developed such that they preserve the relevant information as contained in the original profiles as well as provide adequate levels of denoising. These metrics are outlined as follows -

i. Height of denoised vs. original peak

This metric compares the maximum height of the denoised profile to the maximum height of the original profile [233]. It ensures that denoising does not significantly affect the height of the tracer concentration peak, as the peak height is crucial for RTD-specific applications such as drug product quality and out-of-specification (OOS) analysis.

ii. Width of denoised vs. original peak

This metric compares the width of the denoised tracer concentration profile to the original profile and ensures that the denoising method does not severely impact the width of the tracer concentration profile [233]. As the width varies along the height of the tracer concentration profile, the width is measured at a specific height equal to half of the maximum height (also known as ‘full-width at half-maximum or fwhm’) [267]. The width of the tracer concentration profile is a crucial metric as it is important for applications such as material traceability.

iii. Standard deviation (STD)

The standard deviation of a base signal (constant around a mean value of zero) corresponds to the extent of noise present in the system and, thus, is used as a metric to compare the denoising methods presented in this article. STD is used in two different ways to provide a robust comparison.

a. Signal_STD – This metric compares the standard deviation of the original signal and the denoised signal [233] and thus quantifies the level of denoising.

It is important to keep in mind that this metric assumes the Gaussian profile of tracer concentration in RTD studies itself to be part of the noise and evaluates the standard deviation accordingly.

- b. **Difference_STD** - This metric focuses on comparing STD of the difference between denoised and original signal. The difference between the two signals provides a noisy base signal, and measuring the STD thus estimates the degree of smoothing provided by the denoising method.

iv. **Signal-to-noise ratio (SNR)**

SNR is a widely used metric in engineering literature and defined as the ratio of inherent signal to inherent noise present in the measurement [233-236, 238, 245]. This can be challenging for RTD studies as it is difficult to evaluate the inherent noise present in the profile. Thus, a modified version of SNR is used in this work, defined as a ratio of the mean divided by the standard deviation of the signal [233], and is used in two ways as shown below -

- a. **Difference_SNR** – This metric evaluates the SNR of the signal obtained from the difference between the original and denoised RTD profiles and is evaluated as absolute value of mean of the difference divided by the absolute value of the standard deviation of the difference.
- b. **Improvement_SNR** – This metric is similar to the **difference_SNR** but removes the effect of mean on evaluation of SNR. The mean value of the difference should ideally be 0, but slightly positive or negative values can affect the results of the analysis. Thus, in this case, a mean value of 1 is considered and the metric is evaluated as the reciprocal of the standard deviation of difference between the two signals.

All these metrics, when applied together, ensure that the denoising method preserves the important features of the RTD while providing desired levels of denoising or noise removal. Thus,

these metrics are implemented and analyzed together for a comprehensive comparison of various denoising methods.

6.1.4. Results and discussions

6.1.4.1. Denoising methods on noisy RTD datasets

The denoising methods described in section 3.3 are applied to the experimentally obtained tracer concentration profiles. To provide a robust analysis, the denoising is evaluated for all 12 runs in the experimental plan (as outlined in section 2), where for each experimental setting, 3 replicates are available. The denoised profiles obtained using different denoising methods are presented below in Figure 6.9. Figure 6.9A shows the denoised profiles using SA, and Figure 6.9B shows the denoised tracer concentration profiles using SV with a polynomial of window size of 11 and a polynomial order of 3. Figure 6.9C represents the degree of denoising obtained from FFT, and Figure 6.9D shows the denoising of concentration profiles using WT.

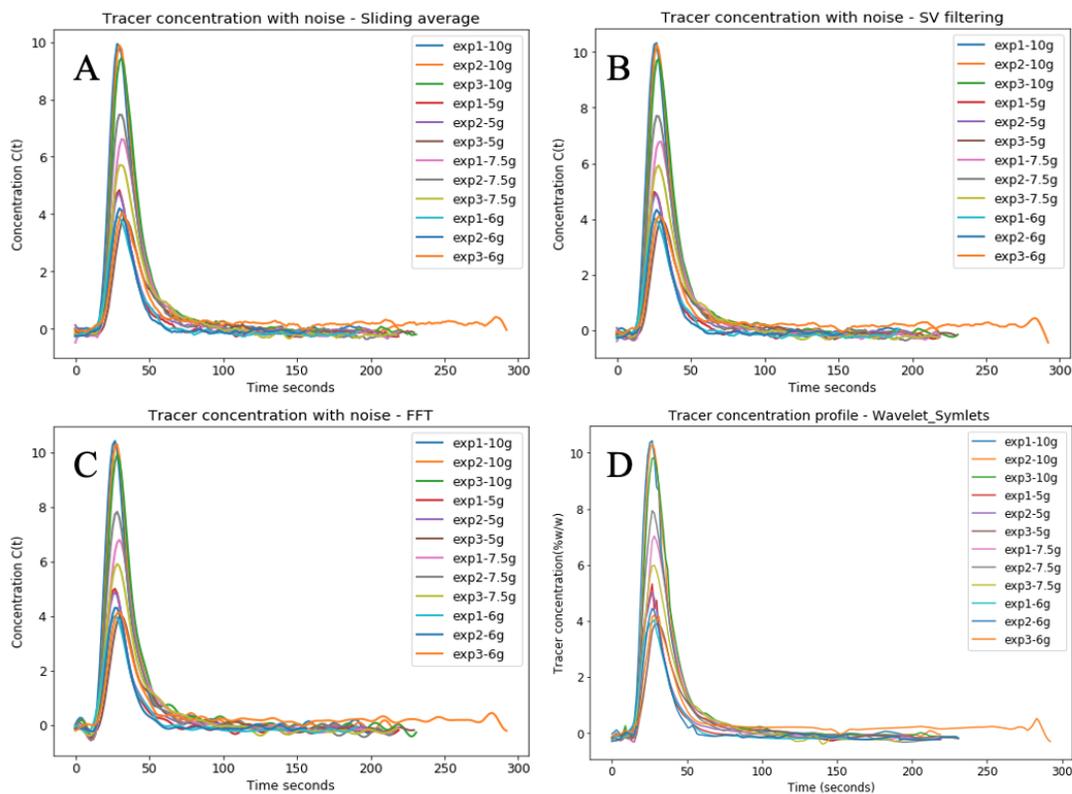


Figure 6.9: Denoised tracer concentration profiles using denoising methods - A) sliding or moving average, B) SV filtering, C) FFT, and D) wavelet transform

6.1.4.2. Comparison of denoising methods based on metric evaluation

Figure 6.10 shows the denoised profiles for the 1st and 2nd run in the experimental campaign showcasing the differences in the denoised profiles using different denoising methods. Here, SA provides a slightly skewed profile as compared to other methods. These profiles are further compared based on comparison metrics calculated for all denoising methods.

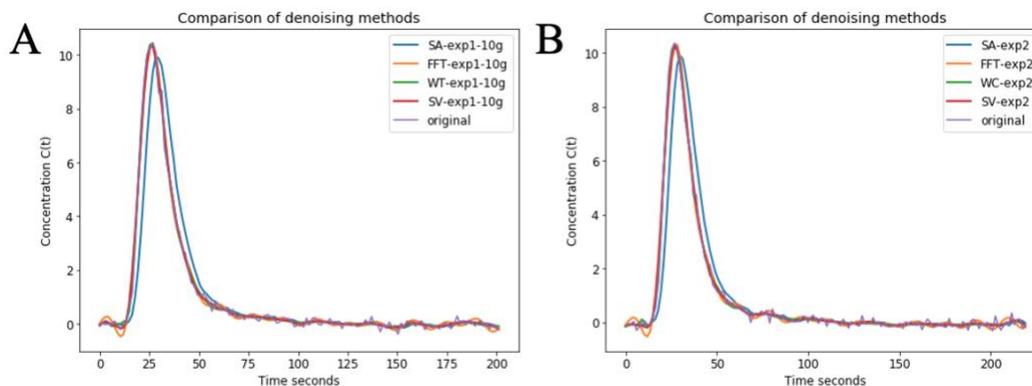


Figure 6.10: Comparison of denoising methods demonstrated for 1st and 2nd experimental run

The comparison metrics - height, width, and standard deviation of the denoised profiles are compared with the original profile, whereas the other metrics, including Difference_STD and Improvement_SNR, are evaluated taking the original profile into account. For the experimental runs performed in this work with 3 repeated experiments for each experimental setting, the denoising metrics are averaged for the repeated experiments. Table 6.1 shows the averaged values for each experimental setting using the different denoising methods. Based on these values, the best denoising method can be investigated.

Table 6.1: Metrics for comparison of different denoising methods on RTD studies

Metrics for comparison	Sliding average (SA)	Wavelet transform (WT)	Savistky Golay (SV)	Fast Fourier Transform (FFT)	Original RTD profile
Height of RTD profile	9.757	10.189	10.111	10.204	10.290
	4.477	4.740	4.605	4.628	4.835
	6.597	6.983	6.804	6.839	7.098
	4.056	4.233	4.140	4.148	4.322
	5.273	5.418	5.380	5.409	5.429

Width of RTD profile	2.475	2.532	2.524	2.534	2.558
	3.848	3.923	3.912	3.921	3.942
	2.319	2.372	2.362	2.367	2.389
Signal_STD	2.239	2.301	2.285	2.297	2.306
	1.051	1.075	1.072	1.076	1.086
	1.634	1.666	1.661	1.665	1.674
	0.985	1.007	1.003	1.005	1.015
Difference_STD	9.089	1.426	1.856	2.443	–
	4.355	1.593	1.608	1.738	–
	5.971	1.621	1.574	1.894	–
	3.825	1.262	1.303	1.465	–
Difference_SNR	1.40E-03	6.90E-03	2.77E-03	3.26E-16	–
	3.09E-03	1.07E-02	8.29E-04	2.03E-16	–
	4.00E-03	8.11E-03	3.59E-03	3.99E-16	–
	9.53E-03	4.37E-03	2.96E-03	5.02E-16	–
Improvement_SNR	1.3338	8.49398	4.96588	6.52627267	
	2.86541	7.6917	7.08962	7.66633434	
	2.06265	7.53036	6.46147	7.75139959	
	3.12194	9.65982	8.13137	9.18556249	

The metrics measuring the height and the width of the tracer concentration profile ensure that the important features of the RTD profile are preserved. It is, however, crucial to remember that the original tracer concentration profile is noisy and metric values for the height and the width of the original profile may be skewed based on the extend of noise present in the system. Thus, for these metrics, the best strategy is to select the denoising method that provides values close to that

obtained for the original profile, albeit it is not required for them to match. Table 6.1 shows that WT provides very close values compared to the original RTD for both metrics – height and width of the RTD profile. The only exception is the height metric for the 1st experiment with 10g tracer, where FFT seems to closely match the values obtained for the original profile.

Signal_STD metric compares the standard deviation of the denoised profiles to that of the original profile. For this metric, the best method should provide the low values close to that of the original profile, as the measured standard deviation considers the Gaussian profile (without noise) itself as noise along with the noise measured during the RTD experiments. Thus, the lower the STD values, the more the RTD signal tends to smoothen out and the RTD features may get lost. Signal_STD metric further shows that the WT method provides the best level of smoothing as compared to other methods.

Difference_STD metric aims at evaluating the standard deviation of the difference of the denoised and the original RTD profile. This difference results in a noisy signal centered around a mean value of 0. Calculating the standard deviation of this signal then provides an understanding of the closeness of denoised and original RTD profiles. Thus, lower value of this metric is better for evaluating denoising methods. However, this value, if equal to 0, indicates that there is no noise reduction between the original and denoised profiles. Table 6.1 shows that WT provides low values of this metric followed by FFT, SV, and SA. For the SNR_Difference metric, which aims at evaluating the SNR of difference between the original and denoised profiles, a lower standard deviation and a correspondingly higher value of the metric would indicate a higher degree of smoothing. From Table 6.1, it is observed that WT provides higher values of the metric in comparison to other methods. SA also provides relatively close values for SNR_Difference. However, the SA method distorts the profile's important features, including width and height. For SNR_Difference, it is also important to keep in mind that the mean values of the difference between

the denoised and the original profiles may affect the outcome. Thus, SNR_Improvement metric is also compared, which considers a mean value of 1 and calculates the SNR of difference between the denoised and original profiles. For this metric, lower standard deviation and correspondingly a higher value of the metric would indicate more denoising. Thus, the WT method outperforms other methods except for one run wherein FFT performs better than WT. These values are highlighted in Table 6.1.

Considering all the above metrics together, it can be concluded that WT performs best, followed by other methods – FFT, SV, and SA. The underlying reason is that WT ensures smoothing across both time and frequency domains. FFT ranks behind WT, as it provides smoothing along only the frequency domain, leading to similar denoising levels only for certain scenarios. It is also important to understand that these values of the comparison metrics are dependent on the meta-parameters for each method (e.g., the averaging window and order of polynomial for SV). Thus, it is imperative to optimize the meta-parameters prior to comparing or utilizing any denoising method.

6.1.4.3. Impact of denoising on applications of RTD

Lastly, following the evaluation and comparison of different denoising methods, this section highlights the impact of denoising of tracer concentration profile on important applications of RTD studies for continuous pharmaceutical manufacturing. These applications include material traceability for tracking material lots and quality assurance for evaluating OOS analysis, which are widely used techniques to ensure robust manufacturing practices [8, 150, 268].

6.1.4.3.1. *Material traceability*

For evaluating the impact of denoising methods on material traceability, a similar approach as outlined by Engisch et al. [210] is implemented, where material lots are identified using lot boundaries defined as 0.5% and 99.5% of the cumulative distribution function (CDF), obtained using the Eq. (3-4). The lot boundaries defined by 0.5% and 99.5% of CDF are used to obtain the boundaries of specific lots for material tracking, as shown in Figure 6.11. Figure 6.11 aims at highlighting the differences in lot boundaries for material tracking of the original profile (shown in Figure 6.11A and Figure 6.11C) and the denoising profile using WT (shown in Figure 6.11B and Figure 6.11D). The lot boundaries for the original profile are 16.375 seconds and 83.384 seconds for 0.5% and 99.5% of CDF of the original tracer concentration profile. These values correspond to 16.379 seconds and 81.895 seconds for 0.5% and 99.5% of CDF for the denoised profile (using WT), leading to an error of 1.7%. Based on these values, it can be concluded that denoising does not largely impact the lot boundaries for material traceability.

$$E(t) = \frac{C(t)}{\int_0^{\infty} C(t) dt} \quad (3)$$

$$F(t) = \int_{t=0}^t E(t) dt \quad (4)$$

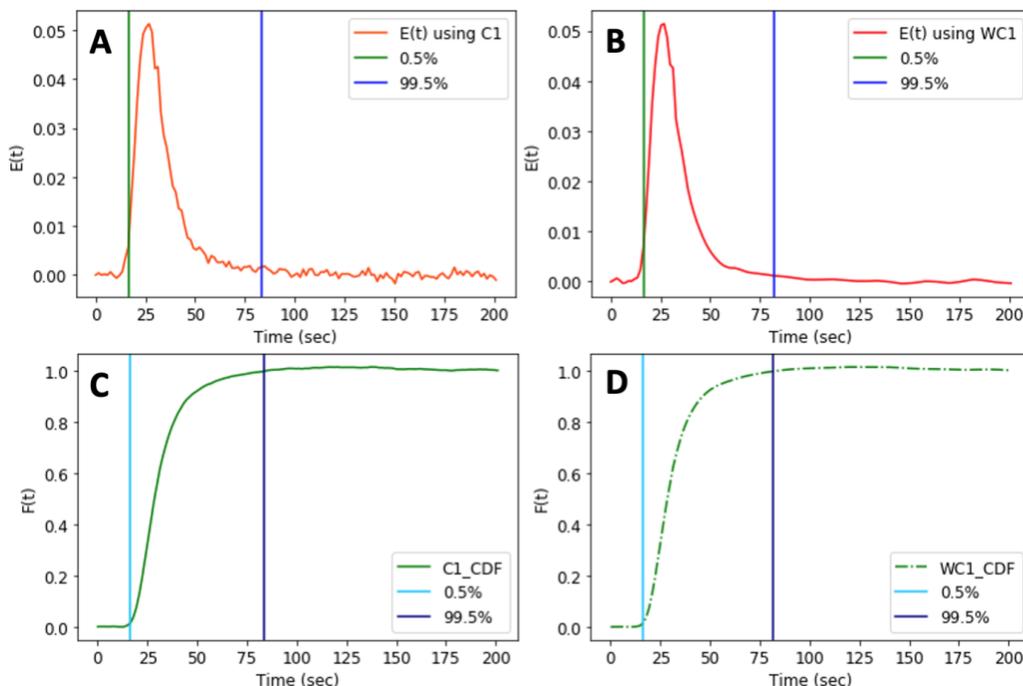


Figure 6.11: Material tracking using original and denoised profiles with lot boundaries defined based on 0.5% and 99.5% of $F(t)$

6.1.4.3.2. Product quality assurance

Similar to material traceability, the impact of signal denoising on product quality assurance is quantified using an OOS analysis. In this approach, the tracer concentration profile is converted to RTD $E(t)$ profiles, which is then used to evaluate OOS product, if any, by convoluting $E(t)$ with different pulse input disturbances with all possible combinations of duration and intensity of the disturbance. Based on the convoluted concentration profile, the product is analyzed for OOS analysis, and the degree of deviation from the desired value is recorded for all these combinations of input profiles. The degree of deviation plotted as a function of the intensity and duration of the input profile is commonly known as the funnel plot. Funnel plots have been widely used to evaluate OOS product and assure product quality [102, 120, 269]. In this work, the funnel plots are developed for the original tracer concentration profile as well as the denoised profile to observe the differences between the two, if any, as shown in Figure 6.12A and Figure 6.12B, where the

denoised profile is obtained using WT. To compare the two funnel plots, they are overlaid on top of each other, as shown in Figure 6.12C, where the two funnel plots align perfectly on top of each other. This indicates that the proposed denoising does not alter the product quality assurance, as observed using the OOS analysis.

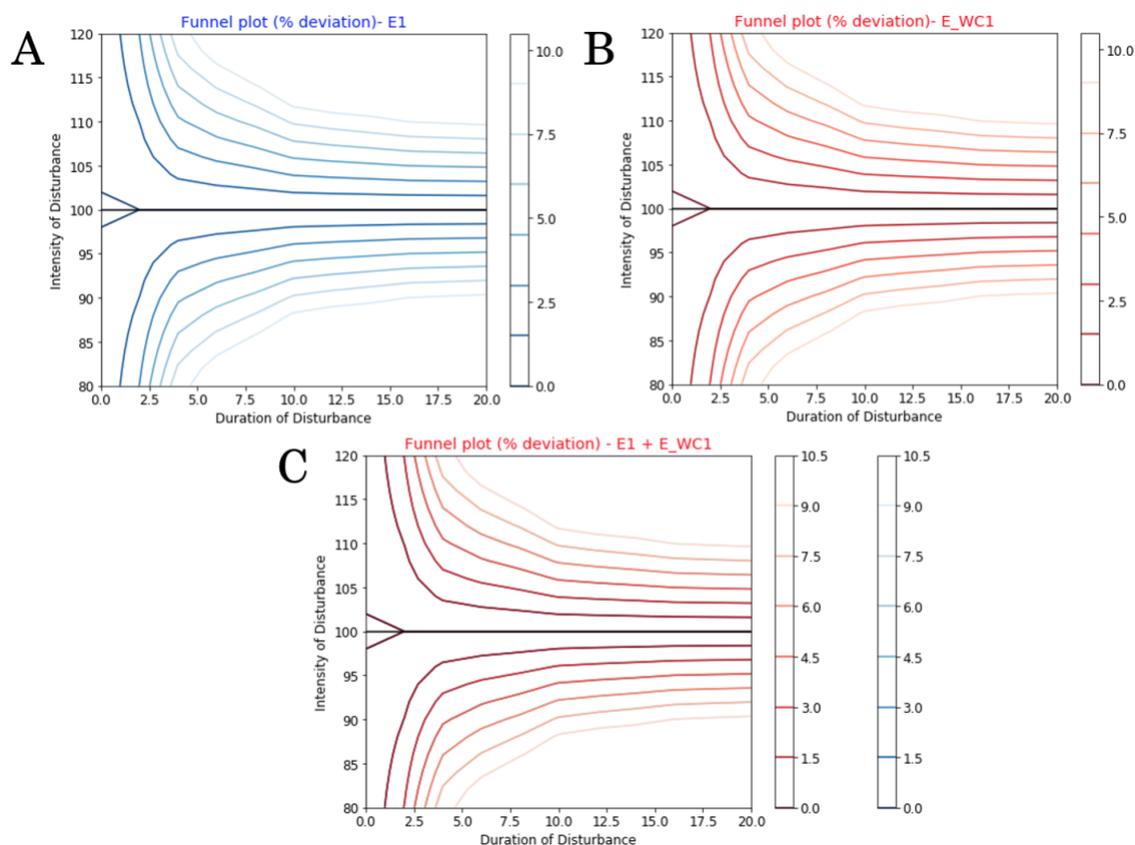


Figure 6.12: Funnel plots developed for A) original concentration profile, B) denoised concentration profile, C) both profiles overlaid together

Thus, it can be concluded that denoising the tracer concentration profile does not impact the RTD applications such as material traceability and product quality assurance and aims at providing greater clarity of the true underlying signal by identifying and smoothening the noise component present in the measured signal.

6.1.5. Conclusions

Despite the importance of accurate signal characterization of RTDs in the pharmaceutical manufacturing industry, the literature lacks description and analysis of appropriate data pre-treatment strategies. This makes it difficult to utilize existing literature for applications such as material traceability and product quality assurance, especially as different data pre-treatment strategies lead to uncertainty regarding the true signal and noise components. To address this problem, the proposed article provides a detailed description of data pre-treatment strategies, including noise characterization and denoising methods for tracer concentration profiles. The article focuses on the characterization of noise present in tracer concentration profiles, discussion of different denoising methods, and quantitative comparison of these strategies using performance metrics. The performance metrics are devised to ensure desired levels of smoothing of the dataset while maintaining the important features of the profile. The results indicate that frequency averaging methods such as wavelet transform and Fourier transform provide higher signal retention than time averaging methods, as observed from the values of the performance metrics. It is important to point out that the specific values of performance metrics for denoising methods are dependent on the respective meta-parameters, and thus, it is recommended to optimize the values of the meta-parameters prior to utilizing or comparing different methods. Further, it is observed that the denoising methods focus on cleaning the dataset to correctly identify the noise and the signal components of the dataset for further analysis and do not scientifically impact the RTD applications such as material traceability and product quality assurance. The presented work on RTD data pre-treatment has been extended to quantify the impact of baseline noise and the need of long RTD tails on product quality assurance through truncation of RTD profiles, aimed towards focusing calibration efforts to capture RTD peak effectively as compared to the tail region.

6.2. Analysis of RTD using truncation

Acknowledgement of publication status:

Entire sections of this chapter are being prepared to be published in a scientific journal written by the author of this thesis. This work is completed in collaboration with Sonia M. Razavi¹, Andrés Roman-Ospino¹, Marianthi G. Ierapetritou², Fernando J. Muzzio^{1*}

Affiliations:

¹ Engineering Research Center for Structured Organic Particulate Systems (C-SOPS), Department of Chemical and Biochemical Engineering, Rutgers – The State University of New Jersey, Piscataway, NJ, 08854, USA

² Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, Delaware, 19716, USA

* Corresponding author – email – muzzio@rutgers.edu

6.2.1. Introduction

The residence time distribution (RTD) is the distribution in time that a material spends or resides in the system under study [197, 199, 200]. It has been widely applied across various technology areas, including continuous manufacturing of petrochemicals, pharmaceutical products, plastics, polymers, food, catalysts, perfumes, oils; distillation columns; absorption and adsorption columns, and chromatography columns; wastewater treatment plants; engineering of flow reactors, fixed bed reactors, microfluidic reactors, continuous flow polymerization, and cascades of stirred tank reactors and other continuous flow chemistry applications [10, 41, 102, 152, 197-200, 212, 270]. Traditionally applied in the continuum domain, RTD aims at understanding the bulk flow patterns and the extent of mixing (and stagnation) for process design and optimization [197]. In the last decade, it has also been extensively employed for applications involving solid granular or powder

systems [102, 212], specifically in the pharmaceutical industry for developing predictive models for powder flow and mixing for various unit operations (such as twin-screw extruders [206, 271, 272], granulators [211, 273-275], fluidized bed dryers [275], continuous blenders [151, 170, 226, 276-279], powder feeders [169, 280], feed frames and tablet press [281-285]) and overall manufacturing lines (including direct compaction [122, 210, 213, 286], dry granulation[47, 211, 278, 287], and wet granulation[287]). These studies have been further extended for important pharmaceutical applications[102], such as material traceability[47, 122, 210, 288, 289], product quality assurance and system health monitoring[214], diversion of out-of-specification (OOS) material[120, 122, 269, 290], aimed towards ensuring final drug product quality and developing robust manufacturing practices.

RTD is normally obtained by introducing a high-contrast traceable material (commonly known as a “tracer”) at the inlet of the system under study along with the bulk fluid and measuring its concentration as a function of time at the outlet [10, 41, 102, 152, 197-200, 212, 270]. For RTD studies, it is essential to ensure that the tracer does not affect the bulk flow of materials, which would prevent capturing the true RTD in the system. Thus, the tracer material is chosen specifically to have similar physico-chemical flow properties as that of the material being traced while being inert and detectable[151, 226]. Once the tracer is selected, the RTD experiments are performed under controlled conditions that minimize the impact of external disturbances on the system, including the disturbance introduced during the addition of tracer at the inlet [152, 291]. The tracer concentration is measured at the outlet of the system, using process analytical technology (PAT) techniques, with a measurement frequency that is at least one order of magnitude faster than the expected duration of the RTD peak. Commonly used PAT techniques in the pharmaceutical industry are based on vibrational spectroscopy, such as NIR and Raman⁴⁵. A PAT strategy involving a relatively fast data acquisition method, based on a chemometric calibration model spanning the entire range of variation of tracer concentration, is developed to correlate the obtained

spectral measurements of the outlet stream to the tracer concentration ($C(t)$). Lastly, the RTD probability distribution function ($E(t)$) is obtained for the system using Eq. (5). Also known as the exit age distribution, $E(t)$ mathematically describes the distribution of time that each portion of the bulk fluid resides in the system before exiting out [101, 197, 199, 200]. Thus, it can be used to understand the flow patterns and develop predictive models to replicate the system dynamics. This is typically done using simple dynamic elements such as ideal stirred tanks and plug-flow pipes. Non-ideal mixing is captured using components such as dead volume, bypass, and recycling.

$$E(t) = \frac{C(t)}{\int_{t=0}^{t=\infty} C(t)dt} \quad (5)$$

The RTD is normally characterized using metrics corresponding to its 1st, 2nd, and 3rd centered moments along time, commonly known as the mean residence time (MRT), the mean-centered variance (MCV), and the skewness of RTD, respectively [101, 197, 199, 200]. These metrics characterize the main features of the RTD profile, such as the peak, width, and skewness, as compared to a Gaussian profile, and the tail region, and are used to quantitatively compare multiple RTDs. Typically, RTD studies (including the experimentation and development of PAT strategy) are conducted such that all the above-mentioned features of the profile are captured as accurately [231, 292, 293]. The chemometric calibration models included in the PAT strategy are best constructed based on an experimental design incorporating process, material, and environmental variability [231]. The obtained spectral measurements at the outlet stream are then pretreated using data normalization techniques to ensure the variability in tracer concentration is the main source of observed variance [231, 294]. The range of concentration values considered in the calibration model is normally defined by the range of potential values observed in the system during the RTD experimentation and depends on the target amount of tracer added to the system. The target amount of tracer is selected to be high enough to be easily detectable and low enough so as not to affect the inherent physico-chemical flow properties of the system [102, 151, 294]. Once the target amount of

tracer is defined, the calibration range is evaluated, and powder blends are produced based on the experimental design spanning the desired calibration range for RTD experiments. Based on these experiments, the calibration models are developed to determine the correlation between spectral measurements and tracer concentration values. It requires extensive experimentation to develop a rigorous calibration model capable of detecting the entire range of observed values of tracer concentration and is further extended for situations where the tracer material is yet to be defined, and different combinations of tracer/bulk powders are considered [229, 231, 294-296].

In previous studies, there has been much emphasis on developing a robust experimental campaign capable of capturing all features of the RTD profile, including the tail and the peak region [226]. However, when developing a chemometric model, the methods distribute the error of prediction across the entire range of measurements. As a result, as the range of concentrations becomes wider, accuracy is somewhat lost across the entire range. This loss of accuracy is most significant in the tail of the distribution because concentrations are smallest, and therefore, relative errors (due to error of measurement) tend to be highest. Efforts to minimize tail errors by, for example, including data sets for a longer time duration, come at the expense of loss of accuracy in the peak region [152]. Conversely, if the calibration model is developed focusing on capturing the peak accurately, it is then likely that accuracy will be lost in the tail region, leading to noise near the RTD baseline (and *vice versa*). Despite the extensive efforts required in obtaining accurate RTD profiles, there exists a possibility that the low values of tracer concentration may not be predicted correctly, especially if the values are below the limit of detection. Román-Ospino et al. [228] present an interesting finding concerning the limit of detection of tracer concentration, where the authors demonstrate the existence of overlapping probability distributions for low values of tracer concentration, which make it challenging to identify these values accurately. The uncertainty introduced in the RTD profiles, due to a typically noisy tail region, eventually impact decisions concerning the “end” of the profile, the RTD metrics that best characterize the profile, and affect

RTD-based applications such as material traceability, and out-of-specification (OOS) analysis [122]. Despite some efforts being directed towards developing appropriate denoising strategies [152, 291] for RTD profiles to address such issues, it may become challenging to distinguish the intrinsic systemic profile from the external noise, especially considering the limit of detection.

The main takeaway point made in this article is that for some RTD applications, the effect of the tail is negligible, and it is possible to minimize effort and maximize accuracy by focusing only on the peak of the RTD. To this end, a novel approach focusing on analyzing the effects of RTD features is investigated from a quality assurance perspective. In the presented work, the experimentally obtained RTD profiles are truncated at different baseline levels to examine the impact of different features of the RTD profile and limit of detection on the final product quality. Ensuring that the quality of drug products is within the desired specification as part of a quality assurance framework is essential in manufacturing pharmaceutical solid drug products [59, 216, 218]. OOS analysis focuses on identifying the products that may not be within the specified limits defined by the regulatory agencies, and if not detected and diverted, can lead to risk to patients, and to product recalls, causing considerable losses to the pharmaceutical companies [150, 268]. Thus, OOS analysis is vital in identifying OOS products beforehand to maintain the quality of the final drug products.

The remainder of this article is organized as follows: Section 2 focuses on the different methods used in the article, including experimental work, OOS identification, and the different case studies for evaluating the important features of RTD using truncation. The results of the case studies are provided in section 3, followed by discussion in section 4, including work focused on identifying optimal levels of truncation in RTD profiles and the effects of truncation of RTD itself by comparison of truncated profiles with the original. Conclusions are outlined in section 5.

6.2.2. Methods

6.2.2.1. Experimental details

The experimental setup used here is shown in Figure 6.13A as described in a previous publication from our group [225]. The details of the setup and experimental procedure are briefly discussed below. In the presented RTD studies, a pulse test is performed, with a high contrast tracer added at the inlet of the system and the tracer concentration tracked at the outlet to obtain the RTD of the system under study. The experimental setup consists of a single K-Tron KT-20 feeder (Coperion K-Tron Pitman Inc., Sewell, NJ) with a coarse auger screw configuration, with a C-type motor to maintain a steady flow of powder at the outlet of the feeder unit. The powder is then fed to the inlet of the continuous blender, where a Gericke GM 250 mixer (Gericke USA, Somerset, NJ) is used with alternate blade configuration operated at 100 RPM, and a weir (semi-circular disc) is placed at 60° at the outlet to control powder hold up inside the unit. The blender outlet is connected to a long vibratory conveyor feeder, providing a homogeneous sample presentation of powder for PAT measurement. A Bruker NIR spectrometer is installed above the conveyor to measure the tracer concentration. The overall setup is arranged as shown in Figure 6.13A, and the system is operated at a flowrate of 12kg/hr with one excipient powder (Avicel PH301) as the bulk powder and Mannogem XL as the tracer. The RTD experiment is repeated 3 times to ensure robustness. To obtain the tracer concentration profile from the NIR measurements, an NIR calibration model is developed with powder blends containing 0.0 % - 9.0 % Mannogem XL in steps of 3.0 % to cover the anticipated range of tracer concentration values. The details of NIR chemometric models are described in Razavi et al. [225]. Here, the tracer concentration profile obtained using 10g tracer amount is used for downstream analysis of important RTD features as shown in Figure 6.13B, and similar analysis can easily be extended for other tracer amounts.

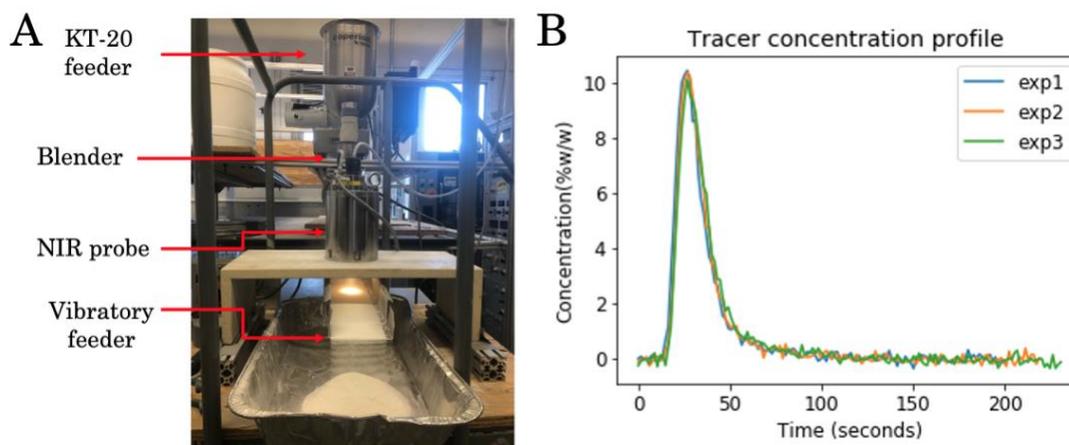


Figure 6.13: RTD studies, A) experimental setup, B) tracer concentration profiles with repeated runs.

6.2.2.2. Predictive identification of OOS product

Formation of OOS product can be predicted using RTD studies, and the steps are listed below [46, 102, 150, 210, 290]. After obtaining the tracer concentration profiles experimentally, these profiles are used for OOS analysis, focused on quantifying the probability of OOS material being produced in the system subject to various tracer input profiles or disturbances in the system. This is quantified by the convolution of input profiles with the RTD ($E(t)$) profiles of the system under study. The first step involves obtaining the $E(t)$ profiles from the tracer concentration profiles ($C(t)$), using Eq. (1). The second step involves convoluting the obtained $E(t)$ profiles with input profiles representing process disturbances or fluctuations. Convolution is performed as shown in Eq. (6), where $E_1(t)$ corresponds to the RTD of the system and $E_2(t)$ represents the input disturbances [210]. For convolution, the RTD dataset can be used directly (i.e., without fitting an RTD model). The convoluted concentration profile provides an estimate of OOS material that deviates from the base value, e.g., having a % deviation of more than 5%, and it can be used to identify the time limits within which the OOS product is formed, during which the material is diverted [210].

$$E_{conv}(t) = E_1(t) \otimes E_2(t) = \int_{-\infty}^{\infty} E_1(\tau) * E_2(t - \tau) d\tau \quad (6)$$

The steps mentioned above are repeated for various combinations of duration and intensity of disturbances applied to the system, which are convoluted with the $E(t)$ profile to obtain % deviations of the signal for all possible combinations of the duration and intensity of disturbance. Funnel plots are developed based on the % deviations of concentration from the base value. First conceptualized in 2017, funnel plots have been widely used in the literature for performing OOS analysis [102, 120, 122, 269]. Funnel plots are mainly categorized into 2 regions – a feasible region where the combinations of duration and intensity of disturbances do not lead to OOS product (i.e., maximum % deviation is less than 5%) and an infeasible region where OOS product is formed (i.e., where the maximum % deviation is greater than 5%) and can potentially lead to product recalls [122]. The steps involved in developing funnel plots are outlined in Figure 6.14. These steps are repeated to obtain funnel plots for different scenarios.

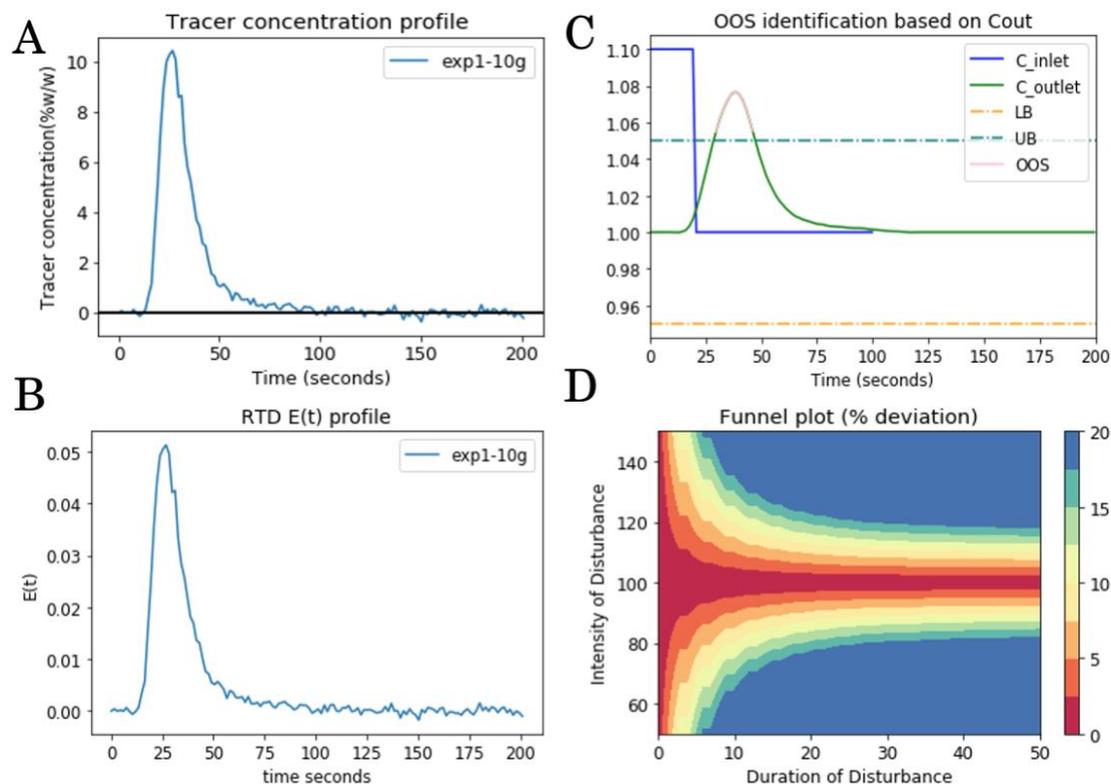


Figure 6.14: Steps involved in OOS analysis - A) obtaining tracer concentration profile, B) evaluating RTD $E(t)$ profile, C) quantifying OOS for input profiles when convoluted with $E(t)$, D) developing funnel plot analysis

6.2.2.3. Evaluation of important features of RTD using truncation

This section focuses on identifying important features of RTD and their respective effects on funnel plot or OOS analysis from a quality assurance perspective. In the presented work, the important features are evaluated using truncation of tracer concentration profiles. Several case studies are showcased in Figure 6.15, where specific parts of the profiles are truncated to understand their impact on the funnel plots. The details of these case studies are outlined below.

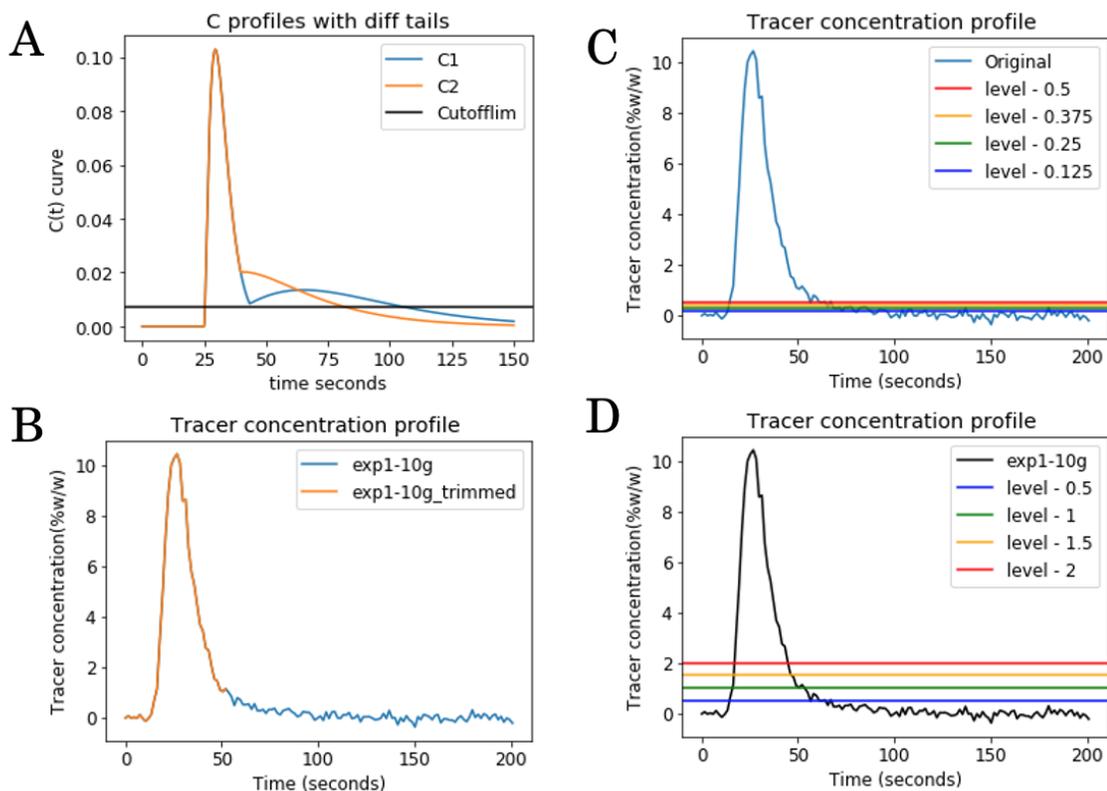


Figure 6.15: Case studies for truncation of RTD profile - A) truncation of tail region, B) truncation along time, truncation along Y axis – focusing on C) low limits of detection, and D) high levels of truncation.

6.2.2.3.1. Truncation of tail regions – do tails matter?

This case study focuses on analyzing the effects on OOS analysis of truncation of the RTD's tail based on a minimum concentration of tracer. This analysis is directly related to the effort required to measure the tail accurately; to accurately quantify lower tracer concentration levels, the experimenter requires more extensive data, with faster and more sensitive (and more expensive) instrumentation. Moreover, retention of a longer tail would necessarily incorporate more error into the overall data set, which would then reduce the accuracy of peak measurements. In this study, two RTD profiles with different tails are compared, and the differences are evaluated based on funnel plots aimed towards analyzing the optimal efforts required to obtain the RTD of a system

experimentally. The tracer concentration profile is simulated to obtain the profiles with different tail regions, obtained by superimposing two Gaussian profiles representing the main RTD peak and the tail region, as shown in Figure 6.15A.

6.2.2.3.2. *Truncation along time axis*

Truncation of the RTD profile along the time axis or vertical truncation can be useful to evaluate the impact of the time extension at the end of the RTD profile on the production of OOS material. This is an important case study as it helps understand and investigate the need for capturing an RTD beyond its peak. As discussed below, this analysis provides an opportunity for significant experimental savings. To investigate this case study, a tracer concentration profile is obtained for a long duration of 200 seconds (well beyond the span of the RTD's peak), and it is compared with a truncated version (truncated at 50 seconds), as shown in Figure 6.15B, to understand how time truncation of the RTD's tail affects the OOS analysis. The corresponding effects are evaluated using funnel plots.

6.2.2.3.3. *Truncation along Y axis*

Truncation of the RTD profile along different values of tracer concentration (shown on the Y axis), i.e., horizontal truncation, is useful in understanding the impact of different regions in the RTD peak on the funnel plot using OOS analysis. This case study is quite useful in evaluating the probable effects of an NIR calibration model that is not calibrated for values below a certain limit (e.g., 2% wt/wt of tracer concentration) on the RTD profile and funnel plots. This case study is explored in two parts. First part focuses on analyzing the impact of low values of tracer concentration on OOS analysis. Capturing low levels of tracer concentration values accurately has been an important issue in the experimental characterization of RTD profiles, especially combined with the efforts to characterize the baseline correctly and distinguish the true signal from noise. In

this study, the tracer concentration profile is truncated at low limits (0.125, 0.25, 0.375, and 0.5 % wt/wt) as shown in Figure 6.15C, and subsequent effects on the funnel plot analysis are analyzed to evaluate the need for accurately obtaining low values of tracer concentration, especially given the inherent concern about their reliability [228]. The second part of this case study focuses on analyzing the impact of higher degree of truncation along Y axis, to identify optimal levels of profile that need to be captured without significant changes in the funnel plot. For this study, the RTD profile is truncated at different levels (0.5, 1, 1.5, and 2% wt/wt) as shown in Figure 6.15D, and the respective effects are evaluated using the funnel plot.

6.2.3.Results

In this section, the case studies described in section 2.3 are evaluated using funnel plots to understand the effects of truncation on OOS analysis. The funnel plots are developed for original and truncated RTDs for ease of comparison, as shown in Figure 6.16. It is important to note that the truncation of tracer concentration profiles is performed such that the total number of measurements are preserved. Here, the total time duration of profile is maintained during truncation to ensure accurate calculation of the area under the curve for $E(t)$. This prevents the need to renormalize the RTD after truncation.

The first case study focusing on truncation of tail regions is performed to mainly answer the question – “do tails matter?”. Towards this end, the differences in the overlaid funnel plots of the two RTD profiles differing in the tail regions is evaluated and shown in Figure 6.16A. Here, the funnel plot highlighted in blue corresponds to C1 (as shown in Figure 6.15A) and that highlighted in orange corresponds to C2 (as shown in Figure 6.15A). The funnel plots closely fit on top of each other in the regions of interest - low duration and low intensity region of the perturbation space. These regions are quite important as the disturbances in these regions are difficult to track given their duration or intensity being too small to be captured easily (i.e., large perturbations are easier

to detect and to track). The fact that the funnel plots overlay on top of each other in these regions is noteworthy, as it provides additional support to the previous finding that OOS analysis is not affected in these regions specifically despite minor changes in the tail region of tracer concentration profile. As mentioned, this finding provides an opportunity for substantial savings in experimental efforts by focusing on the meaningful portion of the RTD profile. For the regions indicated by longer duration and intensity, small fluctuations are observed in Figure 6.16A, and these fluctuations are dependent on the specific tail profiles, as shown in Figure 6.15A.

Figure 6.16B shows the overlaid funnel plots developed for the second case study comparing the original tracer concentration profile and the profile truncated along time (at $t = 50 \text{ sec}$). The funnel plots obtained for the two profiles overlay exactly on top of each other, indicating no visible differences in the funnel plots, despite the tracer concentration profiles being truncated along time at $t = 50 \text{ sec}$. Thus, it can be concluded that the measurement timeline of the selected RTD profile can be effectively reduced at least by a quarter, significantly reducing the time and efforts required in RTD measurement without resulting in a detectable consequence on the final product quality. This is important because RTD studies can be lengthy, consuming large amounts of expensive line time and materials; the savings achieved by time truncation would enable consideration of larger sets of designed experiments, augmenting process knowledge without sacrificing accuracy.

Lastly, Figure 6.16C and Figure 6.16D depict the overlaid funnel plots developed for truncation along the Y axis focusing on identifying low limits of detection and high values of truncation. Figure 6.16C shows the funnel plot developed for tracer concentration profiles being truncated at low limits of detection and it is observed that the funnel plot grows narrower as the truncation limit increases. For example, the funnel plot in blue indicates a truncation level of 0.125 % wt/wt, which has a slightly larger feasible region compared to a truncation level of 0.5 % wt/wt colored in red. Thus, the case study demonstrates an important fact: funnel plots become more conservative as

truncation levels increase, albeit resulting in loss of a small number of acceptable tablets that would be wrongly labeled as OOS but providing an increased degree of confidence on not accepting bad tablets and thus increasing the level of confidence of the QA system. Similar analysis is supported by the results shown in Figure 6.16D, where funnel plots are developed for tracer concentration profiles truncated at higher limits of truncation, as shown in Figure 6.15D. The funnel plot colored in red (for a truncation level of 2% wt/wt) is observed to be more conservative than that shown in blue (for a truncation level of 0.5% wt/wt). Thus, it reinforces the previous findings that conservative estimates of funnel plots are obtained with a higher degree of truncation. This observation is remarkable as it indicates that the experimental savings would in fact result in a more reliable QA system by yielding a conservative funnel plot. In other words, excluding the tail region, where measurements have lower accuracy due to low values of tracer concentration, both reduces the required effort and increases the reliability of the QA system.

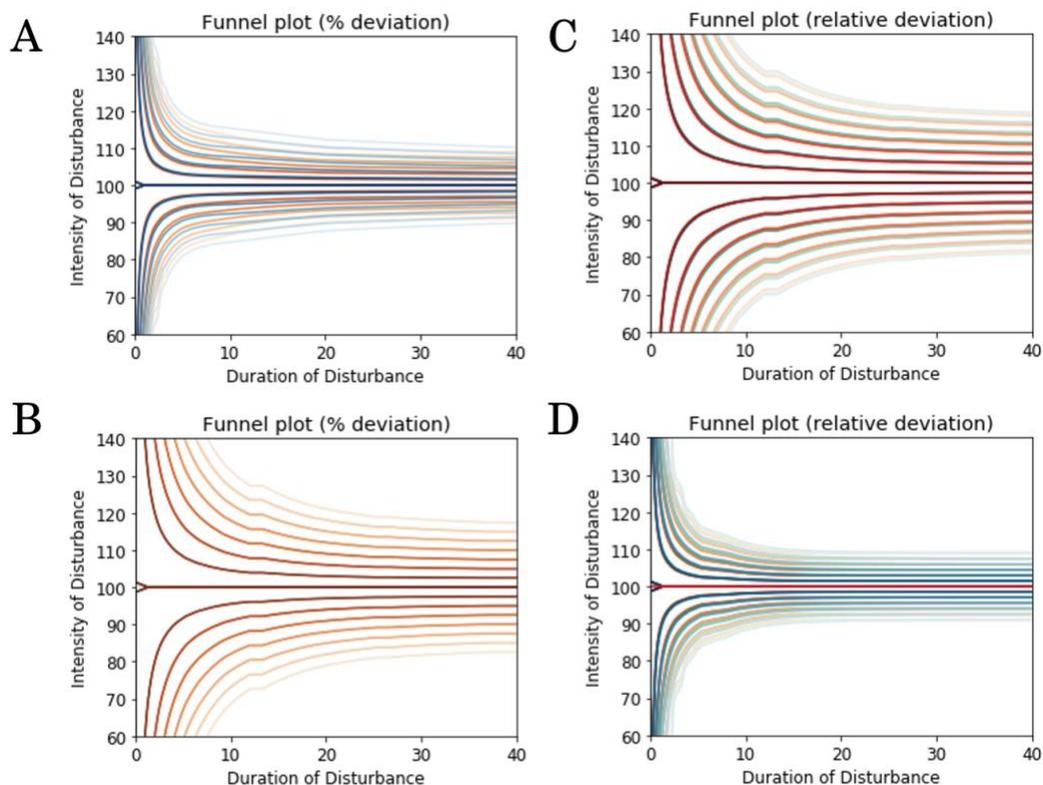


Figure 6.16: Funnel plots for different case studies - A) truncation of tail regions, B) truncation along time, truncation along Y axis – C) low limits of detection and D) high levels of truncation.

6.2.4. Discussions

Following the results of obtaining conservative funnel plots with higher truncation levels, it brings to light the question of computing optimal levels of truncation of RTD. This question is tackled using three different approaches as described below. To quantitatively understand the relationship between the funnel plots and the level of truncation of the tracer concentration profile, and subsequently compute optimal levels of truncation for RTD studies, the first approach involves evaluating the effect of truncation on the degree of conservativeness (reduction) of the feasible region in the funnel plots. Similar to previous studies [210], we define the feasible region in a funnel plot as the region where the combination of intensity and duration of disturbance results in a deviation less than 5% from the base value. To quantify the effect of truncation on the feasible

region, a similar OOS workflow (as implemented in section 2.2) is applied and outlined in Figure 6.17. The first step of this workflow involves truncation of the tracer concentration profile at varying levels indicated by 1.25, 2.5, 3.75, 5, 7.5, and 10% of the maximum height observed in the original profile. The truncated profile is subsequently smoothed using wavelet transform to remove noisy components leading to noise in the funnel plots [233, 291]. Following this step, the funnel plots are developed to identify the effect of truncation on OOS product determination. The funnel plots are obtained for all the truncation levels mentioned above, and the feasible region is quantified in the respective funnel plots. The reduction of feasible region in the funnel plots thus indicate more tablets being classified as OOS with higher levels of truncation. Lastly, this increment in the OOS tablets (assuming a uniform probability distribution of disturbances of varying duration and intensity) is plotted as a function of the truncation levels.

Figure 6.17 demonstrates an interesting finding that the increment in OOS tablets does not increase linearly with the truncation level. Initially, it increases up to a truncation level of 7.5%, and then it plateaus to a % increment of 0.14% as the truncation level is increased to 10%. This is an interesting finding as it indicates that the % loss of tablets (classified as OOS) is not linear with the truncation criteria. This indicates that a higher truncation level would not significantly reduce the tablets within specification limits but instead, would provide substantial savings in terms of experimental time and efforts required to obtain the tracer concentration profile.

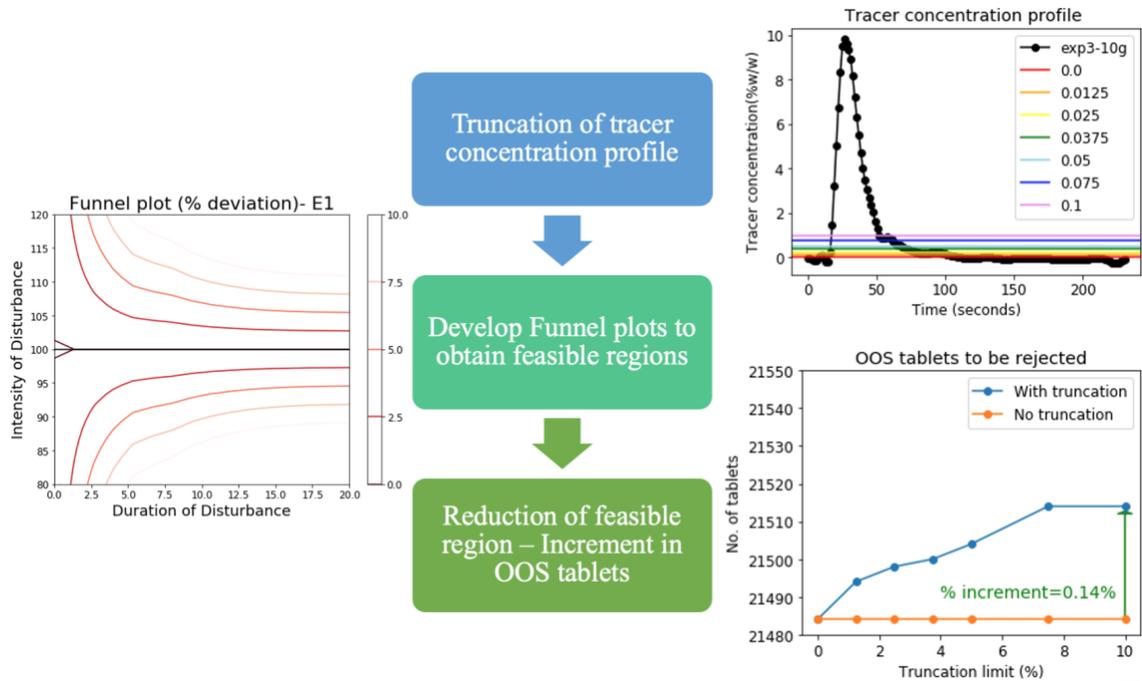


Figure 6.17: Workflow for determining truncation criteria using funnel plots

The second approach to determine an optimal level of truncation focuses on investigating the effect of the truncation on the RTD model parameter estimates. This approach helps evaluate the effect of truncation on other RTD-based applications, such as unit operation and flowsheet modeling, and material traceability. To this end, a comparison of the RTD profiles with different truncation levels is performed using a model-dependent approach [151, 226]. This approach compares the model parameters statistically for the commonly used axial dispersion (AD) model. The AD model replicates the RTD of a system wherein particles are transported by convective and dispersive methods along the length of the unit, normally combined with a plug-flow unit to represent time delay in the system. It is mathematically described using Eq (7), where τ_d is the delay time, τ_{AD} is dispersion time, and Pe is the Peclet number.

$$E(t) = \frac{u(t - \tau_d)Pe^{0.5}}{(4\pi(t - \tau_d)\tau_{AD})^{0.5}} e^{-\left(\frac{Pe\left(1 - \frac{(t - \tau_d)}{\tau_{AD}}\right)}{\frac{4(t - \tau_d)}{\tau_{AD}}}\right)} \quad (7)$$

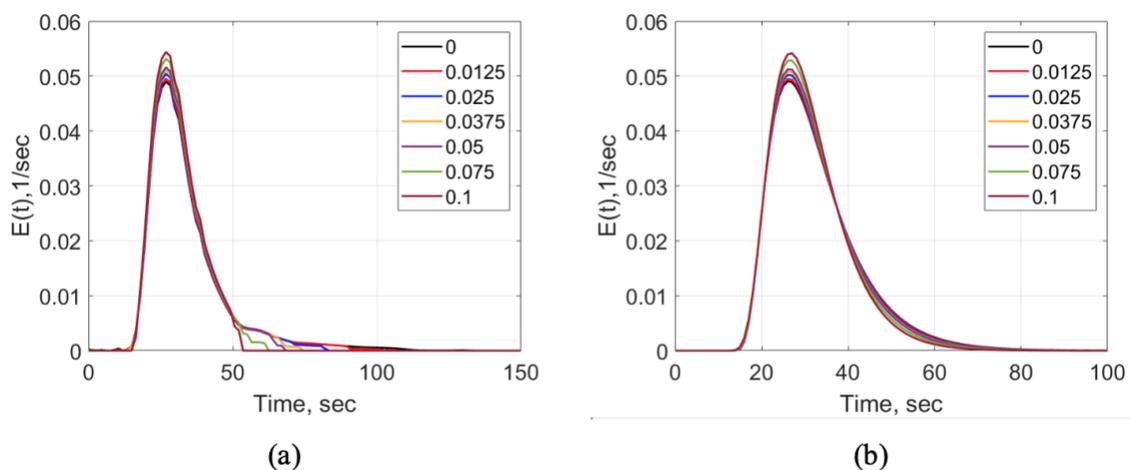


Figure 6.18: $E(t)$ profiles for different levels of truncation (a) averaged for all repeats and (b) regressed using axial dispersion model

The averaged $E(t)$ profiles of the experimental data truncated at 1.25, 2.5, 3.75, 5, 7.5, and 10% of the total maximum height of the tracer concentration profile are obtained using Eq. (1) and plotted in Figure 6.18A. As the truncation levels are increased, the area under the curve decreases, and the intensity of the RTD peak increases. The regressed $E(t)$ profiles using Eq. (3) are presented in Figure 6.18B, and the averaged AD model parameters for the three RTD replicates are provided in Table 6.2. In general, higher truncation levels result in slightly lower τ_d and inversely, larger τ_t , and Pe values. Larger Pe values correspond to less axial back-mixing, which can be attributed to the RTD profiles getting narrower as the truncation levels increase.

Table 6.2: Regressed model parameters for AD model with the student's t -test comparison for different truncation levels

Truncation level	τ_d (sec)			τ_{AD} (sec)			Pe		
	Ave.	St. Dev.	Group	Ave.	St. Dev.	Group	Ave.	St. Dev.	Group
0	11.92	0.57	A	16.37	0.51	A	7.67	0.13	A

0.0125	11.81	0.60	A	16.46	0.46	A	7.90	0.26	A
0.025	11.61	0.67	A B	16.63	0.44	A B	8.35	0.20	A B
0.0375	10.99	0.55	A B	16.59	0.13	A B	9.01	0.07	B
0.05	11.24	0.44	A B	16.96	0.56	A B	9.10	0.26	B
0.075	10.51	0.90	B C	17.65	0.49	B C	10.56	0.71	C
0.1	9.47	0.90	C	18.66	1.13	C	12.47	0.82	D

One-way ANOVA is performed on the regressed model parameters for the seven truncated profiles, including the original profile, which results in a statistically significant difference for all the model parameters. The student's t-test of each pair means comparison grouped the truncation levels based on the statistical significance between the regressed coefficients, as shown in Table 6.2. Truncation levels that share the same letter are found to have similar statistically significant means (p -value < 0.05). Thus, by truncating the concentration profile up to 2.5% of the total maximum height, the differences in regressed model parameters are statistically insignificant, as they share the same letter. On the other hand, the difference between the 0 and 0.1 truncation levels is apparent as they do not share any letters for the three model parameters. These results confirm the sensitivity of the model parameters to the truncation levels and shed light on the possibility of misinterpretation of results. More importantly, how one could interpret the results differently by avoiding any truncation or truncating too much.

Lastly, the third approach of RTD truncation is investigated, taking the inherent system variability captured in RTD measurements. From the understanding that external variability along with the precision of PAT instrument used for RTD measurements affects the RTD profile, especially near the tail of the profile and the baseline, it is important to consider the sensitivity of the PAT instrument for identifying the optimal truncation levels for the tracer concentration profile. Thus, this approach utilizes the sensitivity of the PAT instrument used for measuring RTD profile to

determine optimal truncation levels. The instrument sensitivity is determined based on the extent of variability captured in the measurements when minimal or no tracer is present in the system. The RTD runs captured for a long duration of time (as described in section 2) are used to quantify this extent of variability in the measurements recorded approximately beyond $t = 120$ s where there is minimal or no tracer present in the system. These measurements are separated from the RTD profile and used to calculate the average (μ) and standard deviation (σ) of the measurement variability. For the three RTD run replicates, μ corresponds to -0.03, -0.05, and -0.074 whereas, σ corresponds to 0.125, 0.130, and 0.132. Assuming the variability to have a normal distribution, a truncation level of $\mu \pm 2\sigma$ will capture 95.5% of the overall variability observed in the system, and $\mu \pm 3\sigma$ will capture 99.5% of the total variability. Thus, a conservative truncation limit of $\mu \pm 3\sigma$ is implemented to capture most of the variability observed in the system. The truncated concentration profiles after denoising are shown in Figure 6.19 for all repeats along with the original profiles. The effect of the truncation limit is further quantified using the two approaches mentioned above, i.e., funnel plots for OOS analysis and RTD model parameter estimates. The funnel plots obtained for truncated and original profiles (Figure 6.20) almost overlay on top of each other and thus indicate no significant differences between the two profiles from a quality assurance perspective.

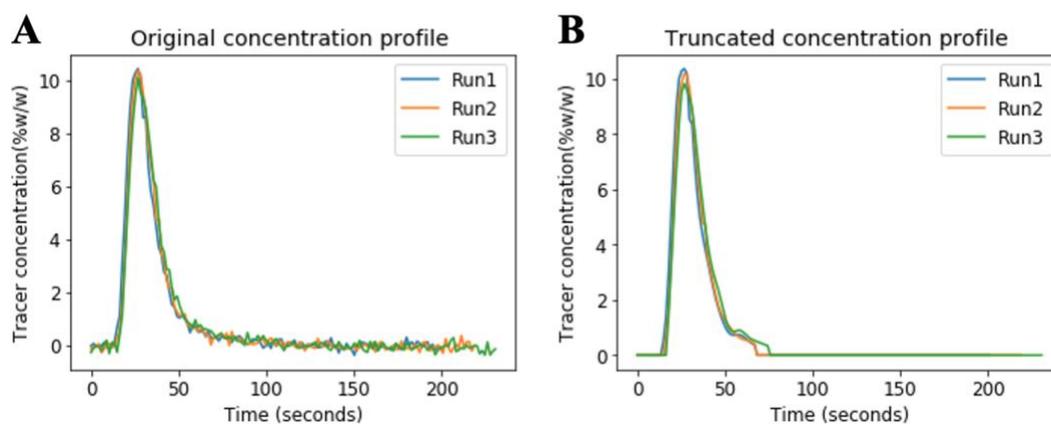


Figure 6.19: Original and truncated tracer concentration profiles

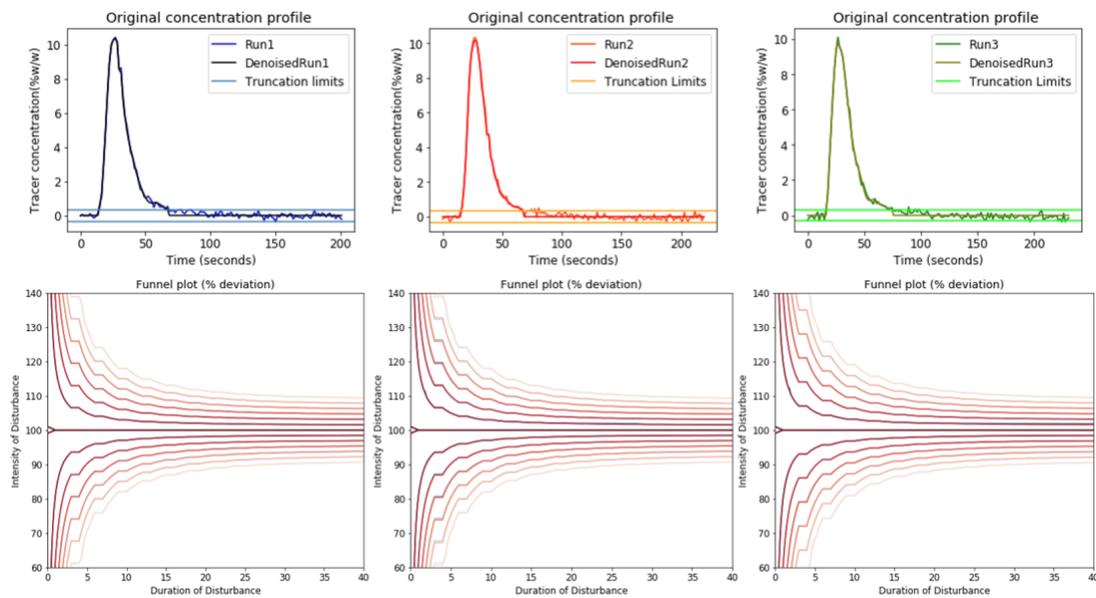


Figure 6.20: Overlaid funnel plots for truncated and original tracer concentration profiles

The truncation strategy proposed in the third approach is further validated by comparing the regressed RTD model parameters for AD model (as described in the second approach). Figure 6.21 demonstrates the experimental and regressed $E(t)$ profiles using the PAT-based truncation method, which shows that the regressions capture the distribution properties truthfully.

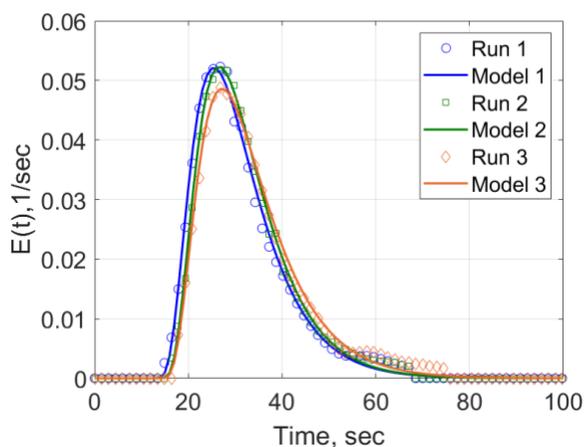


Figure 6.21: The original and regressed $E(t)$ profiles using PAT-based truncation method

Table 6.3 presents a comparison between the regressed AD model parameters of the original (non-truncated) and PAT-based truncated $E(t)$ profiles. Comparing the results in Table 6.2 and Table 6.3 shows that the PAT-based truncation falls around the 3.75% truncation level. This seems to be a practical truncation strategy for three specific reasons:

1. it provides a reasonable truncation level that does not result in very different model parameter estimates
2. it is adjustable to different PAT instruments in use and can be assessed before executing the RTD experiments
3. unlike the second approach, it is independent of the amount of the tracer being pulsed in

Table 6.3 Regressed AD model parameters for the original and PAT-based truncation

Truncation strategy	τ_d (sec)		τ_{AD} (sec)		Pe	
	Ave.	St. Dev.	Ave.	St. Dev.	Ave.	St. Dev.
Original RTD profile	11.92	0.57	16.37	0.51	7.67	0.13
PAT-based truncation	11.40	0.64	16.82	0.35	8.73	0.41

6.2.5. Conclusions

This article aims at investigating the impact of different features of RTD profile on an RTD-based application, OOS analysis, to optimize the time and effort required to obtain an accurate RTD for the system under study. In this work, the RTD tail is truncated along time to observe the respective effects on the determination of OOS product, specifically by developing funnel plots of truncated profiles and comparing them with that developed from the original profiles. Different case studies are developed to investigate the impact of profile truncation on funnel plots. These case studies include truncation along the tail regions, time, and Y axis including low values of tracer

concentration, to identify the impacts of limit of detection. For each of these case studies, the truncated RTD profiles are developed, and respective effects are captured on the funnel plots analyzing the possibilities of production of OOS material. The results for these case studies show that truncation along the tail regions and along time do not have a significant impact on funnel plots, and truncation along the main RTD peak and along lower values of tracer concentration intriguingly provides a more conservative estimate of funnel plots. The conservative nature of the funnel plots obtained by truncation indicates higher confidence of not accepting a bad tablet, at the cost of rejecting a small number of good tablets erroneously labeled as bad. The conservative nature of the funnel plots thus reveals an avenue for significant savings in the amount of time and effort required to capture the tracer concentration accurately, especially at the low concentration levels in the tail region of the RTD. Lastly, efforts to quantify optimal levels of truncation are investigated using a three-way approach. It includes examining the relationship between the degree of conservative nature of the funnel plots with truncation levels, the effect of noise, and truncation on regressed RTD model parameters for the axial dispersion model.

CHAPTER VII

7. Conclusions and future work

7.1. Conclusions

With the collaborative efforts of academia, industry, and regulatory organizations to modernize pharmaceutical manufacturing and move towards realizing FDA's vision for the 21st century for the pharmaceutical industry, there have been immense developments towards Pharma 4.0 using digital twin frameworks. These developments have focused on various aspects of developing digital twins including the virtual and physical components, connection between the two components and integration of process data using cloud computing technologies. The work presented in this thesis mainly focuses on work related to the first two aspects – the virtual and physical components of digital twins and involves efforts directed towards developing multi-scale computationally efficient process models that incorporate detailed process physics to be incorporated within the virtual component along with statistical data handling and analysis of experimentally obtained dataset from the physical component. These efforts thus contribute towards the efficiency of the digital twin as well as establishing connection between the two components. The major contributions of the thesis are outlined below –

- First aim of the thesis focuses on developing high-fidelity simulation models using discrete element modeling (DEM) of unit operations to understand particulate flow from a particle-scale perspective. This is especially beneficial for situations where powder properties can vary with the environment and external conditions and can eventually impact the predictive models. Understanding particulate flow from a particle-scale perspective thus provides an insight into the powder mechanics and flowability while accounting for powder flow based on its properties. DEM calibration is also investigated to develop realistic simulations of

powders within the DEM environment, aimed to capture the true behavior of powder flow. DEM calibration is developed using the bulk calibration approach and aims at identifying a unique solution for the calibration problem by incorporating multiple bulk measurements together to identify calibrated set of DEM parameters.

- The second aim of the thesis mainly focuses on integrating the detailed particulate insight (investigated in aim 1) to develop accurate process models for process flowsheets. DEM models being very time intensive cannot be integrated directly within process flowsheets. Thus, multi-zonal compartmentalization methodology is investigated to develop computationally efficient process models based on the time-intensive DEM simulations of unit operations. The work thus aims at transferring information from a particulate scale to a unit operation scale to develop accurate and efficient process models. The proposed strategy is developed for continuous powder blenders based on DEM simulation of periodic sections of the blender. Blend uniformity, being an important critical quality attribute for regulation compliant manufacturing of pharmaceutical drug products, is quantified from DEM using mixing index – relative standard deviation (RSD). Given the plethora of mixing indices available in literature, a detailed literature review is performed to classify the indices into three main categories, with guidelines for selecting appropriate indices based on the specific application. Using these guidelines, RSD is selected for developing process model of blender using DEM simulation. The unit operation model thus developed can be integrated within process flowsheets for quick simulation and accurate prediction.
- The third aim of the thesis focuses on the physical component of the digital twin – specifically the statistical data handling and pre-treatment of systemic data obtained from the manufacturing lines. The experimentally obtained data normally encounters various fluctuations and disturbances given external and environmental conditions and may potentially lead to erroneous conclusions about the systemic behavior. Thus, this aim focuses on data pre-treatment and handling for residence time distribution (RTD) profiles

obtained from manufacturing line. In this work, different statistical denoising methods are used to denoise RTD profiles and compared based on quantitative metrics to ensure sufficient degree of denoising. The degree of denoising needs to be balanced with retention of important features of RTD and degree of smoothing. The quantitative metrics are devised to balance the degree of denoising of RTD profiles. Denoised RTD profiles can thus effectively remove the effects of unwanted noise and be used towards various RTD-based applications such as characterization of manufacturing lines, material traceability, and OOS analysis. Along with the data pre-treatment, efforts focusing on optimizing the time and efforts required for obtaining RTD experimentally are performed by identifying the important features of RTD and capturing those well. This is performed using truncation of RTD profiles to isolate the effects of different features and their effects are respectively investigated on OOS analysis, ensuring regulation compliant drug production. The investigation including evaluation of tail and baseline regions, led to interesting conclusions indicating that mainly the RTD peak is important for OOS analysis and the tail or baseline regions do not affect the OOS profiles as much, resulting in conservative funnel plots with truncation.

7.2. Future work

Directions for future research are outlined below. These research areas are highlighted for the different scales of process modeling (particle, unit, and flowsheet scale) highlighted in the thesis.

- For particle-scale simulations outlined in aim 1 specifically the DEM calibration, incorporation of electrostatics has not been explored in literature and would be an interesting addition to ensure realistic predictions of powder flow. Electrostatics can lead to agglomeration or sticking of powder to the walls of the unit operation due to charge transfer and this may affect the overall powder flow dynamics. DEM calibration is normally performed

for a specific powder or a set of bulk measurements and need to be repeated for different powders in the material database. Repetition of calibration efforts can be immensely time-consuming and hinder its applicability. An interesting idea would be to develop calibrated material database in DEM parameters using the pharmaceutical material database. This database can thus provide the calibrated parameters of all powders and be used as a reference guide for future DEM simulations. DEM simulations (developed for feeder units) can be extended for complicated unit operations in the pharmaceutical manufacturing lines such as granulators, roller compactors, extruders along with incorporation of fluid. This will help understand the system dynamics from a detailed perspective. Such simulations can be developed for overall manufacturing lines as well, especially to understand the effect of transfer from one unit to another on powder flowability and dynamics.

- Multi-zonal compartment models have shown to translate important particle-scale information to a unit-operation scale and useful for integration within process flowsheets for simulation of unit operations. Such models can be extended to more complicated unit operations such as extruders to develop computationally efficient process models. Such models when integrated together to develop process flowsheets, can also be investigated for applications of manufacturing lines such as material and disturbance traceability.
- For the RTD data handling and pre-treatment outlined in aim 3, an interesting and important future work would include developing a python library for pre-processing of RTD profiles resulting in denoised RTD profiles with comparison of different denoising methods. Outlier detection and detection of RTD profiles can be incorporated within the python library to develop a complete data pre-treatment library for RTD profiles. RTD truncation can further be extended for material traceability applications. An important challenge for material traceability involves identification of start and end of RTDs, especially considering the variabilities observed near the baseline. Developing a variability or uncertainty range based on the PAT system used for detection of tracer concentration,

would be an interesting avenue for future work. This can be developed to obtain quantification of this variability, based on which accurate RTD profiles can be evaluated by discarding the environmental and external noise. Such RTD profiles can accurately capture the systemic information of manufacturing lines. Lastly, these can be incorporated within the digital twin frameworks as well to ensure data pre-treatment and handling for digital twins.

CHAPTER VIII

8. Appendices

8.1. Aim 1 – Developing high-fidelity DEM simulations for pharmaceutical unit operations with DEM calibration

8.1.1. Calibration results for bulk measurements

Different bulk measurements are used in the study for material calibration including shear angle, angle of repose, maximum flowrate at opening and mass in lower container. For these bulk measurements, the important main and interaction effects are identified using stepwise regression with minimization of AIC, BIC criteria and p-values. Following the results obtained for shear angle, analysis of variance and parameter estimates are obtained and shown in Figure 8.1 for main effects. Figure 8.2 shows the ANOVA results with parameter estimates for main and interaction effects corresponding to shear angle measurement. The results for other bulk measurements are shown below where Figure 8.3 corresponds to the important main and interaction effects for mass in lower container. Angle of repose evaluations are shown in Figure 8.4 whereas maximum flowrate evaluations are shown in Figure 8.5.

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	1436.4083	478.803	19.1527
Error	17	424.9874	24.999	Prob > F
C. Total	20	1861.3957		<.0001*

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	1.7904356	2.919069	0.61	0.5478
PP_SF	17.470889	4.713975	3.71	0.0018*
PP_RF	81.251157	25.98335	3.13	0.0061*
PW_SF	27.464222	4.713975	5.83	<.0001*

Figure 8.1: Analysis of variance and parameter estimates of main effects for shear angle

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	5	1673.2603	334.652	26.6818
Error	15	188.1353	12.542	Prob > F
C. Total	20	1861.3957		<.0001*

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	12.07512	3.159112	3.82	0.0017*
PP_SF	17.470889	3.338976	5.23	0.0001*
PP_RF	84.068879	18.41601	4.56	0.0004*
PW_SF	27.464222	3.338976	8.23	<.0001*
(PP_SF-0.35)*(PP_SF-0.35)	-113.9805	36.25152	-3.14	0.0067*
(PW_SF-0.35)*(PW_SF-0.35)	-80.78046	36.25152	-2.23	0.0416*

Figure 8.2: Analysis of variance and parameter estimates of main and interaction effects for shear angle

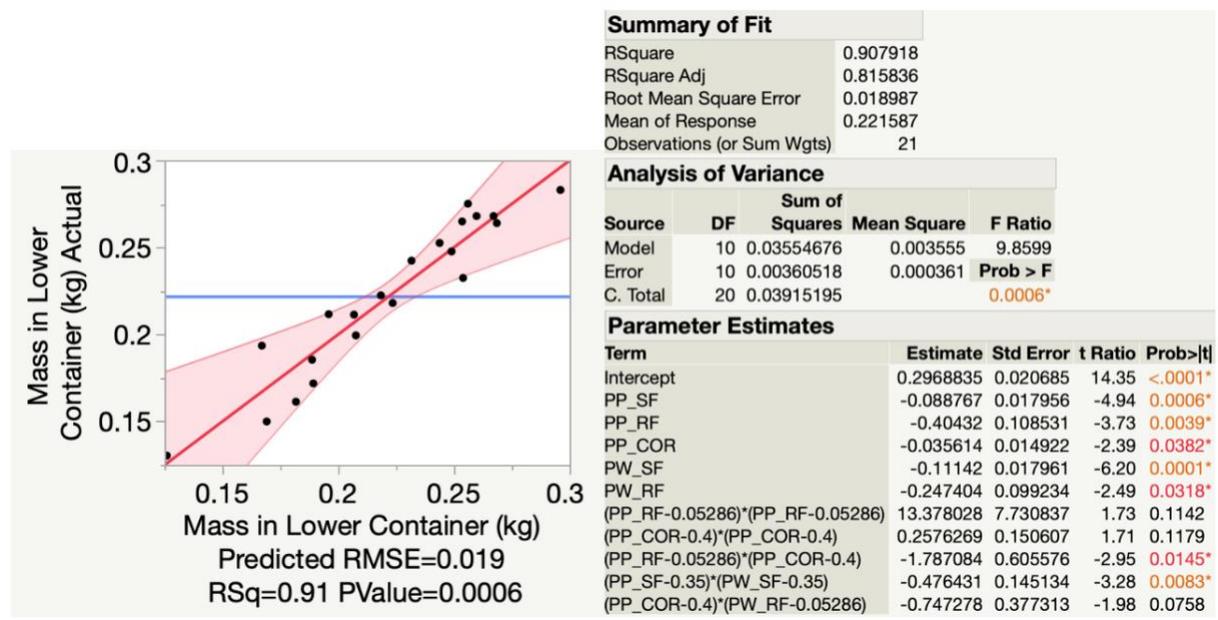


Figure 8.3: Main and interaction effects for mass in lower container of the drawdown test

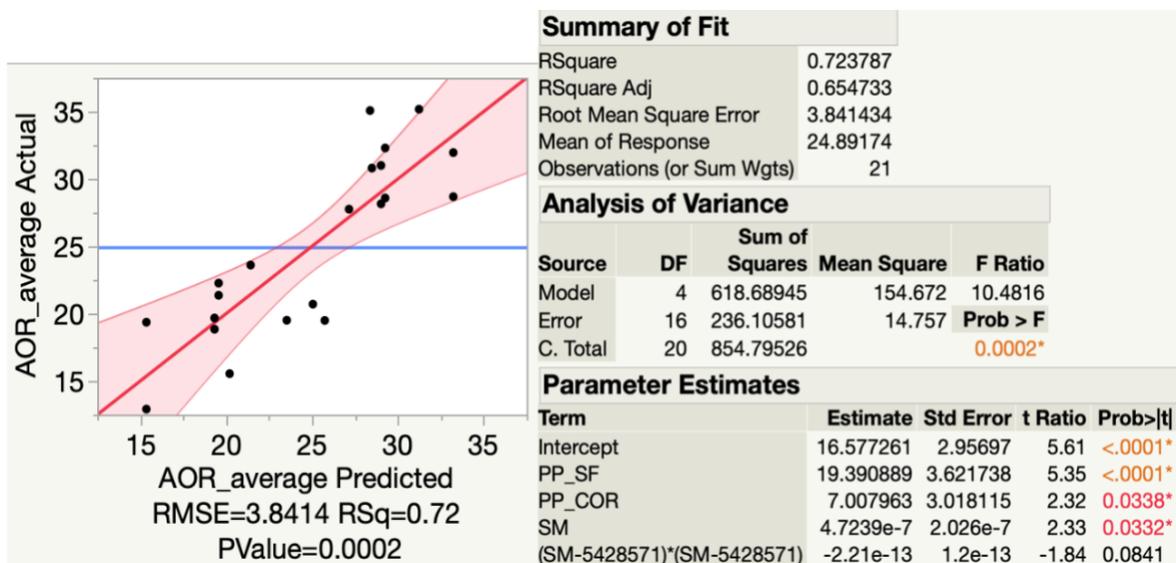


Figure 8.4: Main and interaction effects for bulk measurement of angle of repose

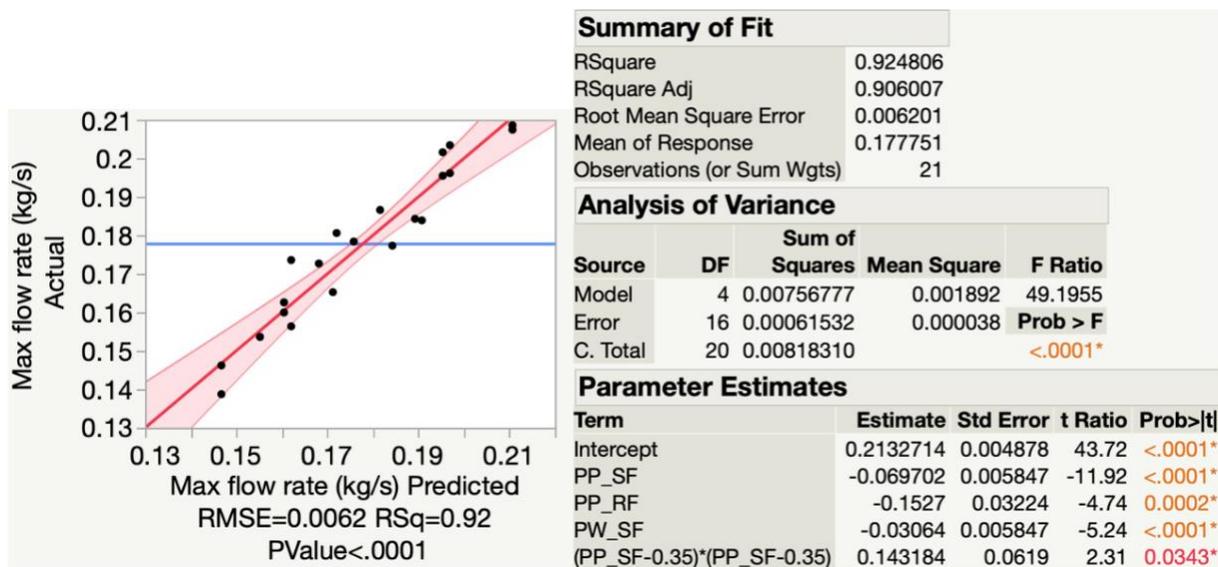


Figure 8.5: Main and interaction effects for bulk measurement of maximum flow rate

8.1.2. Evaluation of feed factor

Table 8.1 highlights feed factor parameter values regressed from feed factor vs feeder weight profile. Goodness of the fit measured using R^2 values are also measured in the table below.

Table 8.1: Regressed feed factor parameters for the selected DOE

Run	FF_{max}	FF_{min}	β	R^2
1	0.03158	0.00177	11.91720	0.89611
2	0.05161	0.00004	4.46254	0.95363
3	0.06896	0.00004	4.78237	0.96879
4	0.05764	0.00023	6.68204	0.97840
5	0.04572	0.00263	6.45111	0.93269
6	0.06569	0.00148	4.25551	0.95585
7	0.02180	0.00727	12.28483	0.76931
8	0.06500	0.00194	4.81382	0.96475
9	0.04334	0.00133	9.85730	0.93485
10	0.03663	0.01062	10.74547	0.84227
11	0.05342	0.00280	5.47477	0.94281
12	0.03386	0.00422	7.82309	0.89391
13	0.02355	0.00518	11.42371	0.78380
14	0.03487	0.01204	11.21864	0.81465
15	0.06751	0.00156	4.51660	0.95554
16	0.03477	0.00164	10.50590	0.89336
17	0.05112	0.00495	5.64182	0.88750
18	0.02466	0.00574	11.93730	0.81609
19	0.03448	0.00340	7.60655	0.90767
20	0.04859	0.00314	6.34797	0.93488
21	0.07235	0.00022	4.79984	0.95953

Figure 8.6-Figure 8.8 provide parity plots for the feed factor parameters (FF_{max} , FF_{min} , and β) along with an effect summary of important parameters for their predictions.

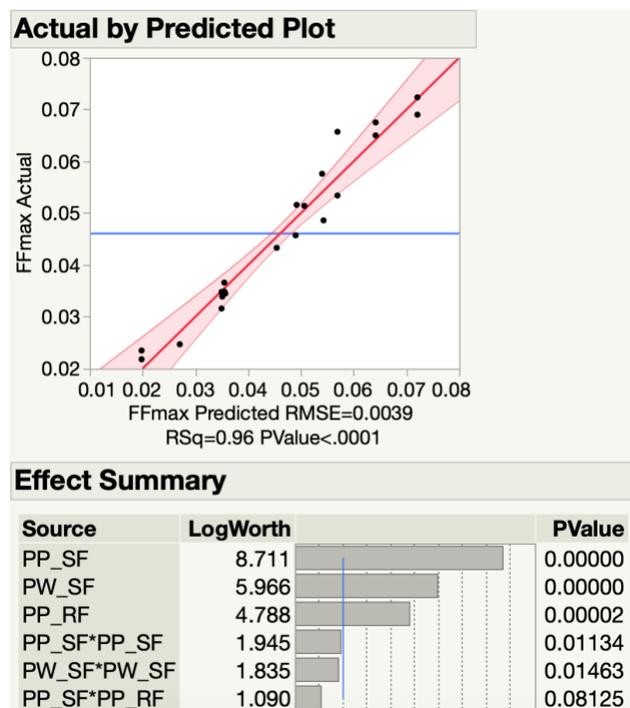


Figure 8.6: Parity plot for FF_{max} with effects summary

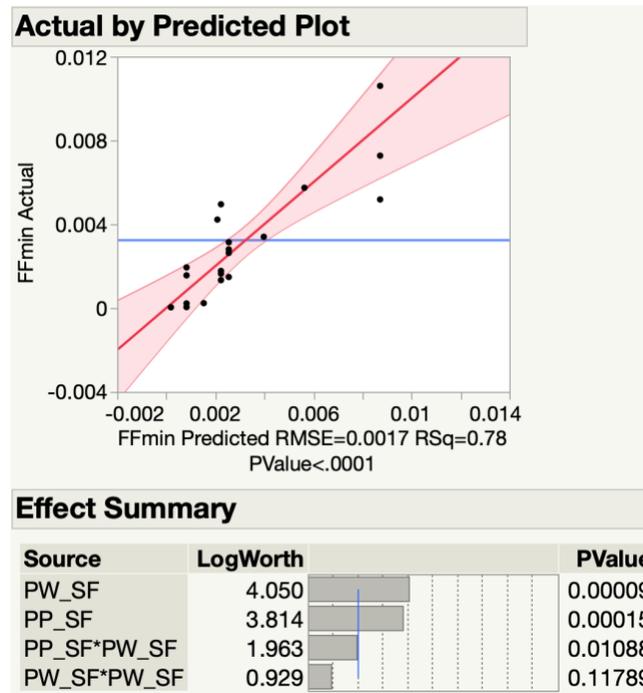


Figure 8.7: Parity plot for FF_{min} with effects summary

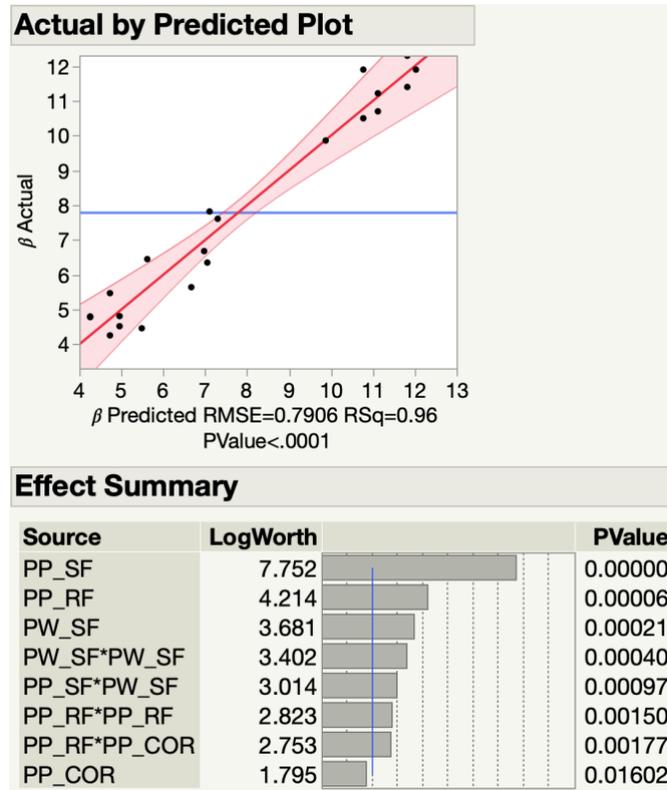


Figure 8.8: Parity plot for β with effects summary

Prediction profiles for all feed factor parameters (FF_{max} , FF_{min} , and β) are shown in Figure 8.9- Figure 8.11, which elaborates the dependence relationship between important DEM material and feed factor parameters.

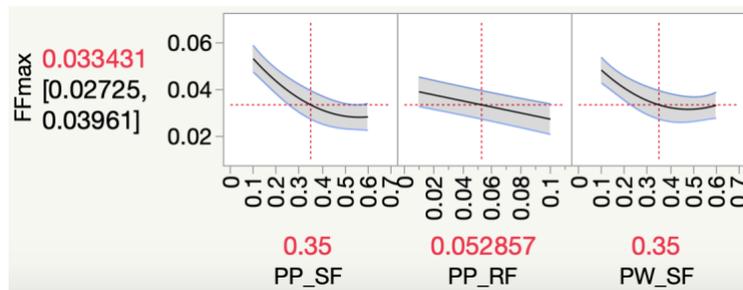


Figure 8.9: Prediction profiles for FF_{max}

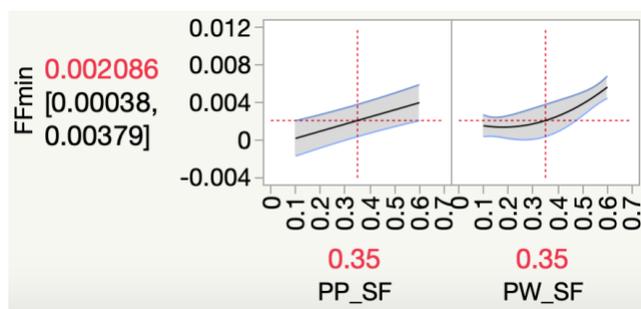


Figure 8.10: Prediction profiles for FF_{min}

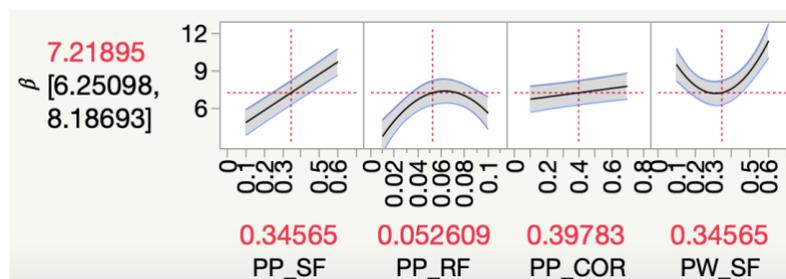


Figure 8.11: Prediction profiles for β

8.2. Aim 2: Unit operation process modeling using DEM – mixing indices

This section includes supplementary results illustrating goodness of model fits for evaluation of T_{95} mixing times for different indices. Table 8.2 highlights the goodness of the model fits for blender operation with blade speed of 100 RPM, Table 8.3 for a blade speed of 75 RPM and Table 8.4 for a blade speed of 50 RPM. Indices – SM and NDM are not evaluated for 75 and 50 RPM due to the high computational requirements associated with these indices.

Table 8.2: Results of model fit for blender operation with a blade speed of 100 RPM

Model Parameters	R ² fit	Mixing Indices
------------------	--------------------	----------------

Afit	Gamma	T_{95}		
0.6968	0.3469	7.5948	0.9345	DM
2.6007	1.0127	3.9020	0.9728	VBBC
6.4894	0.9724	5.0042	0.9897	ME
1.0721	0.4407	6.9552	0.9039	VRR
0.4265	0.4833	4.4355	0.8142	MSI
1.2514	0.5038	6.3915	0.8365	ESI
1.2906	0.4833	6.7268	0.8142	RSD
2.2095	1.0508	3.6053	0.8998	LI
1.8814	0.4437	8.1758	0.9648	CN
1.9524	0.5819	6.2980	0.9802	SI
1.0453	0.4234	7.1799	0.9866	NNM
0.6447	0.2639	9.6891	0.6554	PS
-0.8400	0.6071	4.6475	0.9501	AH
0.9998	2.9268	1.0235	0.9367	SM
0.9763	0.5023	5.9165	0.8012	NDM

Table 8.3: Results of model fit for blender operation with a blade speed of 75 RPM

Model Parameters			R ²	Mixing Indices
Afit	Gamma	T_{95}		
0.6666	0.2378	10.8918	0.9688	DM
1.6018	0.5531	6.2675	0.9875	VBBC
1.6174	0.5852	5.9412	0.9958	ME
0.9379	0.3473	8.4402	0.9185	VRR

0.4717	0.4832	4.6445	0.9652	MSI
1.2089	0.3837	8.3010	0.9808	ESI
1.3132	0.4030	8.1100	0.9182	RSD
1.2171	0.5683	5.6170	0.9622	LI
1.6061	0.2905	11.9429	0.9570	CN
1.5099	0.3703	9.2033	0.9893	SI
1.1115	0.3192	9.7151	0.9841	NNM
-0.0008	0.3966	10.4956	0.9048	AH

Table 8.4: Results of model fit for blender operation with a blade speed of 50 RPM

Model Parameters			R ²	Mixing Indices
Afit	Gamma	T ₉₅		
0.5941	0.1725	14.3492	0.9854	DM
1.3663	0.3788	8.7326	0.9861	VBBC
1.2852	0.3956	8.2073	0.9936	ME
0.7970	0.2259	12.2566	0.9702	VRR
0.3765	0.3147	6.4148	0.9815	MSI
1.0740	0.2615	11.7292	0.9740	ESI
1.3288	0.3147	10.4225	0.9815	RSD
0.9122	0.3700	7.8488	0.9958	LI
1.3962	0.1904	17.4830	0.9707	CN
1.3295	0.2667	12.3011	0.9959	SI
1.0891	0.2287	13.4705	0.9899	NNM
-0.0006	0.2163	20.5012	0.9054	AH

8.3. Aim 2: Unit operation process modeling using DEM – compartment modeling

8.3.1. Details of DEM simulation

The details of the DEM simulation of periodic section are provided in Table 8.5, along with the particle properties of the bulk powder used in the DEM simulation. The walls of the equipment and the blades are made of the same material – steel and DEM parameters are assigned correspondingly as shown in Table 8.5.

Table 8.5: Details of DEM model parameters used for periodic section

DEM Parameters		Values
Powder particles	Particle radius	1 mm
	Particle density	1500 kg/m ³
	Shear modulus	2E6 Pa
	Poisson's ratio	0.25
Particle-particle interactions	Static friction	0.5
	Rolling friction	0.01
	Coefficient of restitution	0.1
Equipment (walls and blades)	Shear modulus	8E10 Pa
	Poisson's ratio	0.29
	Solid density	7800 kg/m ³
Particle-wall interactions	Static friction	0.5
	Rolling friction	0.01
	Coefficient of restitution	0.1
Simulation time step	10% of Rayleigh's time step	

8.3.2. Radial mixing profiles using RSD

Radial mixing profile using relative standard deviation (RSD) mixing index, obtained from the original DEM simulation of periodic section is shown in Figure 8.12. The RSD profile RSD_{DEM} is evaluated using the Eq. (B.1). Here, C_j is the concentration of each grid (j) in the DEM simulation, N_{grids} is the total number of grids and C_{av} is the averaged concentration of all grids.

$$RSD_{DEM} = \left(\frac{1}{C_{av}} * \sqrt{\frac{\sum_{j=1}^N (C_j - C_{av})^2}{N_{grids} - 1}} \right) \quad (B.1)$$

To obtain the RSD profile as shown in Figure 8.12, the original DEM simulation is broken down into 10x6x6 grids along X, Y and Z directions for evaluating the RSD profile. The gridding is selected such that it provides a grid size equivalent to the final pharmaceutical tablet of size 100 mg.

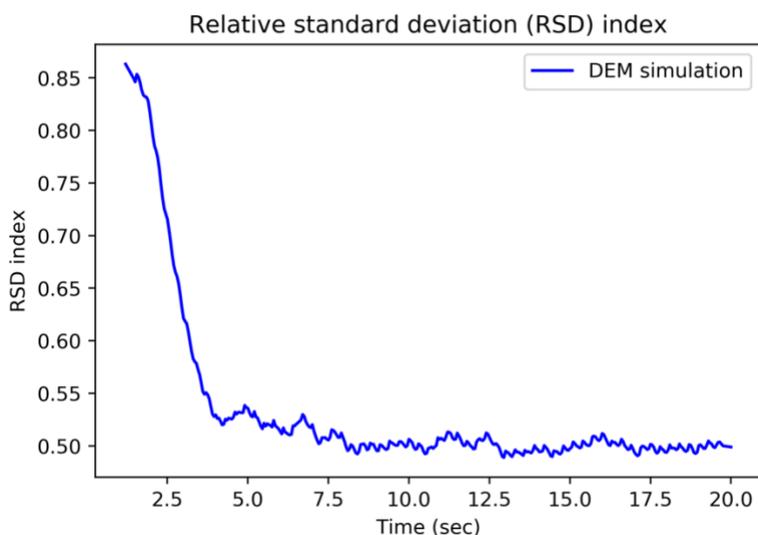


Figure 8.12: RSD profile obtained from DEM simulation of periodic section

It can be seen from Figure 8.12 that a similar trend is observed as compared to that obtained for the developed compartment model, with the profile stabilizing to a value close to 0.5.

CHAPTER IX

9. References

1. Mikulic, M., May 4). Global Pharmaceutical Market Size 2001-2019. Statista. Retrieved February 23, 2022, from <https://www.statista.com/statistics/263102/pharmaceutical-market-worldwide-revenue-since-2001/>
2. Ltd., R.a.M. *Research and Markets - Market Research Reports*. Pharmaceuticals global market report 2021: Covid-19 impact and recovery to 2030 25th February 2022]; Available from: https://www.researchandmarkets.com/reports/5292738/pharmaceuticals-global-market-report-2021-covid?utm_source=GNOM&utm_medium=PressRelease&utm_code=4gx334&utm_campaign=1518358%2B-%2BGlobal%2BPharmaceuticals%2BMarket%2BReport%2B2021%3A%2BMarket%2Bis%2BExpected%2Bto%2Bgrow%2Bfrom%2B%241228.45%2BBillion%2Bin%2B2020%2Bto%2B%241250.24%2BBillion%2Bin%2B2021%2B-%2BLong-term%2Bforecast%2Bto%2B2025%2B%26%2B2030&utm_exec=chdo54prd.
3. Mikulic, M., August 25). Total prescription drug revenue global projection 2020-2026. Statista. Retrieved February 23, 2022, from <https://www.statista.com/statistics/309387/global-total-prescription-drug-revenue-projection/>.
4. *Novel Drug Approvals for 2020*. [cited 2022 February 23]; Available from: <https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/novel-drug-approvals-2020>.
5. Jhamb, K., *Continuous Manufacturing in Pharmaceuticals: Implications for the Generics Market*. Drug Dev. Deliv., 2019. **19**: p. 1-5.
6. Research and development in the pharmaceutical industry. Congressional Budget Office. (n.d.). Retrieved February 23, from <https://www.cbo.gov/publication/57126>
7. M. Ierapetritou, F.J.M., G. Reklaitis, *Perspectives on the continuous manufacturing of powder-based pharmaceutical processes*. AIChE Journal, 2016. **62**: p. 1846-1862.
8. Lee, S.L., et al., *Modernizing Pharmaceutical Manufacturing: from Batch to Continuous Production*. Journal of Pharmaceutical Innovation, 2015. **10**(3): p. 191-199.
9. Matsunami, K., et al., *Decision Support Method for the Choice between Batch and Continuous Technologies in Solid Drug Product Manufacturing*. Industrial & Engineering Chemistry Research, 2018. **57**(30): p. 9798-9809.

10. Teżyk, M., et al., *Recent progress in continuous and semi-continuous processing of solid oral dosage forms: a review*. Drug Development and Industrial Pharmacy, 2015.
11. Plumb, K., *Continuous Processing in the Pharmaceutical Industry: Changing the Mind Set*. Chemical Engineering Research and Design, 2005. **83**: p. 730-738.
12. Wahlich, J., *Review: Continuous Manufacturing of Small Molecule Solid Oral Dosage Forms*. Pharmaceutics, 2021. **13**(8).
13. Vanhoorne, V. and C. Vervaet, *Recent progress in continuous manufacturing of oral solid dosage forms*. International Journal of Pharmaceutics, 2020. **579**: p. 119194.
14. Gernaey, K.V., Cervera-Padrell, A. E. & Woodley, J. M. , *Development of continuous pharmaceutical production processes supported by process systems engineering methods and tools*. Future Medicinal Chemistry, 2012. **4**(11): p. 1371-1374.
15. Gernaey, K.V., Gani, R., *A model-based systems approach to pharmaceutical product-process design and analysis*. . Chemical Engineering Science 2010. **65**: p. 5757–5769.
16. Gernaey, K.V., A.E. Cervera-Padrell, and J.M. Woodley, *Development of continuous pharmaceutical production processes supported by process systems engineering methods and tools*. dx.doi.org, 2012. **4**(11): p. 1371-1374.
17. Gernaey, K.V., A.E. Cervera-Padrell, and J.M. Woodley, *A perspective on PSE in pharmaceutical process development and innovation*. Computers and Chemical Engineering, 2012. **42**: p. 15-29.
18. Gernaey, K.V. and R. Gani, *A model-based systems approach to pharmaceutical product-process design and analysis*. Chemical Engineering Science, 2010. **65**(21): p. 5757-5769.
19. Schenkendorf, R., et al., *Model-Based Tools for Pharmaceutical Manufacturing Processes*. Processes, 2020. **8**(1): p. 49.
20. Helal, N.A., et al., *Integrated continuous manufacturing in pharmaceutical industry: current evolutionary steps toward revolutionary future*. Pharmaceutical Patent Analyst, 2019. **8**(4).
21. Karttunen, A.-P., et al., *Comparison between integrated continuous direct compression line and batch processing – The effect of raw material properties*. European Journal of Pharmaceutical Sciences, 2019. **133**: p. 40-53.
22. Domokos, A., et al., *End-to-end continuous manufacturing of conventional compressed tablets: From flow synthesis to tableting through integrated crystallization and filtration*. International Journal of Pharmaceutics, 2020. **581**: p. 119297.
23. Içten, E., et al., *A Virtual Plant for Integrated Continuous Manufacturing of a Carfilzomib Drug Substance Intermediate, Part 1: CDI-Promoted Amide Bond Formation*. Organic Process Research & Development, 2020.
24. Mascia, S., Patrick L. Heider, Haitao Zhang, Richard Lakerveld, Brahim Benyahia, Paul I. Barton, Richard D. Braatz et al., *End-to-end continuous manufacturing of*

- pharmaceuticals: integrated synthesis, purification, and final dosage formation*. Angewandte Chemie International Edition, 2013. **52**(47): p. 12359-12363.
25. JM., J., *Juran on quality by design: the new steps for planning quality into goods and services*. 1992, New York: The Free Press.
 26. Janet Woodcock, p.q.i.t.s.c.-a.i.s.a., AAPS workshop: pharmaceutical quality assessment - a science and risk-based CMC approach in the 21st Century, North Bethesda, MD, 5-October-2005.
 27. Pharmaceutical CGMPs for the 21st century: a risk-based approach; a science and risk-based approach to product quality regulation incorporating an integrated quality systems approach, U.S.D.o.
 28. Badman, C., Cooney, C.L., Florence, A., Konstantinov, K., Krumme, M., Mascia, S., Nasr, M. and Trout, B.L., *Why we need continuous pharmaceutical manufacturing and how to make it happen*. Journal of Pharmaceutical Sciences, 2019. **108**(11): p. 3521-2523.
 29. *PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance | FDA*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pat-framework-innovative-pharmaceutical-development-manufacturing-and-quality-assurance>. Accessed April 18, 2021.
 30. U. S. Food and Drug Administration. *Guidance for Industry: Q8 (2) Pharmaceutical Development*. 2009.
 31. International Conference on Harmonisation. 2009. ICH: *Pharmaceutical Development Q8(R2)*. [accessed 2020 Nov 20]. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q8_R1/Step4/Q8_R2_Guideline.pdf.
 32. U. S. Food and Drug Administration. *Guidance for Industry: Q9 Quality Risk Management*. 2006.
 33. U. S. Food and Drug Administration. *Guidance for Industry: Q10 pharmaceutical quality system*. 2009.
 34. U. S. Food and Drug Administration. *Guidance for Industry: Q11 development and manufacture of drug substance*. 2012.
 35. ICH. ICH Quality IWG: *Points to Consider for ICH Q8/Q9/Q10 Guidelines*; ICH: Geneva, Switzerland, 2012; 17p. Available online: https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use/q9/q10-guidelines_en.pdf (accessed on 2 July 2021).
 36. ICH. *Final Concept Paper ICH Q13: Continuous Manufacturing of Drug Substances and Drug Products*; ICH: Geneva, Switzerland, 2018; 2p. Available online: https://database.ich.org/sites/default/files/Q13_EWG_Concept_Paper.pdf (accessed on 2 July 2021).
 37. Burcham, C.L., A.J. Florence, and M.D. Johnson, *Continuous Manufacturing in Pharmaceutical Process Development and Manufacturing*. Annual Review of Chemical and Biomolecular Engineering, 2018.

38. Byrn, S., et al., *Achieving Continuous Manufacturing for Final Dosage Formation: Challenges and How to Meet Them May 20–21 2014 Continuous Manufacturing Symposium*. Journal of Pharmaceutical Sciences, 2015. **104**(3): p. 792-802.
39. Myerson, A.S., et al., *Control Systems Engineering in Continuous Pharmaceutical Manufacturing May 20–21, 2014 Continuous Manufacturing Symposium*. JOURNAL OF PHARMACEUTICAL SCIENCES, 2015. **104**(3): p. 832-839.
40. Nasr, M., et al., *Regulatory Perspectives on Continuous Pharmaceutical Manufacturing: Moving From Theory to Practice: September 26-27, 2016, International Symposium on the Continuous Manufacturing of Pharmaceuticals*. JOURNAL OF PHARMACEUTICAL SCIENCES, 2017. **106**(11): p. 3199-3206.
41. Rogers, A. and M. Ierapetritou, *Challenges and opportunities in modeling pharmaceutical manufacturing processes*. Computers & Chemical Engineering, 2015. **81**: p. 32-39.
42. Schaber, S.D., et al., *Economic Analysis of Integrated Continuous and Batch Pharmaceutical Manufacturing: A Case Study*. Industrial & Engineering Chemistry Research, 2011. **50**(17): p. 10083-10092.
43. Woodcock, J. *Modernizing pharmaceutical manufacturing—continuous manufacturing as a key enabler*. in *International Symposium on Continuous Manufacturing of Pharmaceuticals*. 2014. Cambridge, MA.
44. Escotet-Espinoza, M.S., *Phenomenological and Residence Time Distribution models for unit operations in a continuous pharmaceutical manufacturing process.*, in *Department of Chemical and Biochemical Engineering*. 2018, Rutgers University: Piscataway, New Jersey.
45. Abbasi, M.R., et al., *Dynamic process modeling and hybrid intelligent control of ethylene copolymerization in gas phase catalytic fluidized bed reactors*. Journal of Chemical Technology and Biotechnology, 2019. **94**(8): p. 2433-2451.
46. Kruisz, J., et al., *RTD modeling of a continuous dry granulation process for process control and materials diversion*. International Journal of Pharmaceutics, 2017. **528**(1-2): p. 334-344.
47. Martinetz, M.C., et al., *RTD-based material tracking in a fully-continuous dry granulation tableting line*. International Journal of Pharmaceutics, 2018. **547**(1-2): p. 469-479.
48. Barrasso, D., et al., *A multi-scale, mechanistic model of a wet granulation process using a novel bi-directional PBM–DEM coupling algorithm*. Chemical Engineering Science, 2015. **123**: p. 500-513.
49. Barrasso, D. and R. Ramachandran, *Multi-scale modeling of granulation processes: Bi-directional coupling of PBM with DEM via collision frequencies*. Chemical Engineering Research and Design, 2015. **93**: p. 304-317.
50. Kumar, A., *Experimental and model-based analysis of twin-screw wet granulation in pharmaceutical processes*. 2015, Ghent University. Faculty of Bioscience Engineering.
51. Matsunami, K., et al., *Analysis of the Effects of Process Parameters on Start-Up Operation in Continuous Wet Granulation*. Processes, 2021. **9**(9).

52. Vercruyssen, J., et al., *Continuous twin screw granulation: influence of process variables on granule and tablet quality*. Eur J Pharm Biopharm, 2012. **82**(1): p. 205-11.
53. Chaudhury, A., et al., *Population Balance Model Development, Validation, and Prediction of CQAs of a High-Shear Wet Granulation Process: Towards QbD in Drug Product Pharmaceutical Manufacturing*. Journal of Pharmaceutical Innovation, 2014. **9**(1): p. 53-64.
54. Jain, M.S., Lohare, G. B., Bari, M. M., Chavan, R. B., Barhate, S. D., & Shah, C. B., *Spray drying in pharmaceutical industry: A review*. Research Journal of pharmaceutical dosage forms and technology, 2012. **4**(2): p. 74-79.
55. Nizar Al-Zoubi, S.G., Ahmad Aljaberi, and Ioannis Nikolakakis, *Spray Drying for Direct Compression of Pharmaceuticals*. Processes, 2021. **9**(2): p. 267.
56. Eitzlmayr, A. and J. Khinast, *Co-rotating twin-screw extruders: Detailed analysis of conveying elements based on smoothed particle hydrodynamics. Part 1: Hydrodynamics*. Chemical Engineering Science, 2015. **134**: p. 861-879.
57. Eitzlmayr, A.K., J. and *Co-rotating twin-screw extruders_ Detailed analysis of conveying elements based on smoothed particle hydrodynamics. Part 2_ Mixing*. Chemical Engineering Science, 2015. **134**: p. 880-886.
58. Matić, J., et al., *Developing HME-Based Drug Products Using Emerging Science: a Fast-Track Roadmap from Concept to Clinical Batch*. AAPS PharmSciTech, 2020. **21**(5): p. 1-18.
59. Ierapetritou, M., F.J. Muzzio, and G. Reklaitis, *Perspectives on the continuous manufacturing of powder-based pharmaceutical processes*. AIChE Journal, 2016. **62**(6): p. 1846-1862.
60. *The future is digital: In silico tissue engineering*. Current Opinion in Biomedical Engineering, 2018. **6**: p. 92-98.
61. Chen, Y., et al., *Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review*. Processes, 2020. **8**(9): p. 1088.
62. Kumar, S.H., et al., *Adaptations of Pharma 4.0 from Industry 4.0*. search.ebscohost.com.
63. Liao, Y., et al., *Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal*. International Journal of Production Research, 2017.
64. Steinwandter, V., D. Brochert, and C. Herwig, *Data science tools and applications on the way to Pharma 4.0*. Drug Discovery Today, 2019. **24**(9): p. 1795-1805.
65. Chen, Y., Yang, O., Sampat, C., Bhalode, P., Ramachandran, R., Ierapetritou, M., *Digital Twins in Pharmaceutical and Biopharmaceutical Manufacturing: A Literature Review*. Processes, 2020. **8**(1088).
66. Cao, H., et al., *A Systematic Framework for Data Management and Integration in a Continuous Pharmaceutical Manufacturing Processing Line*. Processes, 2018. **6**(5): p. 53.
67. Sampat, C., et al., *Enabling Energy-Efficient Manufacturing of Pharmaceutical Solid Oral Dosage Forms via Integrated Techno-Economic Analysis and Advanced Process Modeling*. Journal of Advanced Manufacturing and Processing, 2022.

68. Grieves, M. and J. Vickers, *Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems*, in *Transdisciplinary Perspectives on Complex Systems*. 2017, Springer, Cham. p. 85-113.
69. Jones, D., et al., *Characterising the Digital Twin: A systematic literature review*. CIRP Journal of Manufacturing Science and Technology, 2020.
70. Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., *Digital twin in manufacturing: a categorical literature review and classification*. IFAC Papers Online, 2018. **51**: p. 1016-1022.
71. Michael. Grieves, J.V., *Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems*, in *Transdisciplinary Perspectives on Complex Systems*, F.J.K.e. al., Editor. 2017, Springer: Switzerland. p. 85-113.
72. Nargund , P., Shilpa, P. Guenther , Kathrin, and K. Mauch, *The Move toward Biopharma 4.0*. Genetic Engineering & Biotechnology News, 2019. **39**(6): p. 53-55.
73. Steinwandter, V., Brochert, D. & Herwig, C. , *Data science tools and applications on the way to Pharma 4.0*. . Drug Discovery Today, 2019. **24**: p. 1795 - 1805.
74. Rogers, A.J. and M. Ierapetritou, *Challenges and opportunities in modeling pharmaceutical manufacturing processes*. 2015. **81**: p. 32-39.
75. Wang, Z.L., M.S. Escotet-Espinoza, and M. Ierapetritou, *Process analysis and optimization of continuous pharmaceutical manufacturing using flowsheet models*. Computers & Chemical Engineering, 2017. **107**: p. 77-91.
76. Ierapetritou, M., Sebastian Escotet-Espinoza, M., Singh, R., Khinast, J. & Rantanen, J. , *Process Simulation and Control for Continuous Pharmaceutical Manufacturing of Solid Drug Products*. *Continuous Manufacturing of Pharmaceuticals*. 2017: John Wiley & Sons, Ltd, 2017. 33–105
77. Yingjie Chen, P.B., Marianthi Ierapetritou, *Review of Recent Advancements Towards the Development of Digital Twin of Continuous Pharmaceutical Manufacturing in Continuous Pharmaceutical Processing and Process Analytical Technology* A.N. Atul Dubey, Editor. 2022, CRC Press.
78. Bin Yeom, S., et al., *Application of the Discrete Element Method for Manufacturing Process Simulation in the Pharmaceutical Industry*. *Pharmaceutics*, 2019. **11**(8): p. 414.
79. Ketterhagen, W.R., M.T.a. Ende, and B.C. Hancock, *Process modeling in the pharmaceutical industry using the discrete element method*. *JOURNAL OF PHARMACEUTICAL SCIENCES*, 2009. **98**(2): p. 442-470.
80. Bhalode, P. and M. Ierapetritou, *Discrete element modeling for continuous powder feeding operation: Calibration and system analysis*. *International Journal of Pharmaceutics*, 2020. **585**: p. 119427.
81. Bhalode, P. and M. Ierapetritou, *Hybrid multi-zonal compartment modeling for continuous powder blending processes*. *International Journal of Pharmaceutics*, 2021. **602**: p. 120643.
82. Marigo, M. and E.H. Stitt, *Discrete Element Method (DEM) for Industrial Applications: Comments on Calibration and Validation for the Modelling of Cylindrical Pellets*. *KONA Powder and Particle Journal*, 2015. **32**(0): p. 236-252.

83. *A compartmental model to describe hydraulics in a full-scale waste stabilization pond.* Water Research, 2012. **46**(2): p. 521-530.
84. Bezzo, F., S. Macchietto, and C.C. Pantelides, *General hybrid multizonal/CFD approach for bioreactor modeling.* Trans. AIChE 2003. **49**(8): p. 2133-2148.
85. Delafosse, A., et al., *CFD-based compartment model for description of mixing in bioreactors.* Chemical Engineering Science, 2014. **106**: p. 76-85.
86. Forgber, T., et al., *Extended validation and verification of XPS/AVL-Fire™, a computational CFD-DEM software platform.* Powder Technology, 2020. **361**: p. 880-893.
87. Gresch, M., et al., *Compartmental Models for Continuous Flow Reactors Derived from CFD Simulations.* Analytical Chemistry, 2009.
88. Guha, D., et al., *CFD-based compartmental modeling of single phase stirred-tank reactors.* Trans. AIChE 2006. **52**(5): p. 1836-1846.
89. Jajcevic, D., et al., *Large-scale CFD-DEM simulations of fluidized granular systems.* Chemical Engineering Science, 2013. **98**: p. 298-310.
90. Sen, M., et al., *A Multi-Scale Hybrid CFD-DEM-PBM Description of a Fluid-Bed Granulation Process.* Processes, 2014. **2**(1): p. 89-111.
91. Tobo, Y.M., J. Bartacek, and I. Nopens, *Linking CFD and Kinetic Models in Anaerobic Digestion Using a Compartmental Model Approach.* Processes, 2020. **8**(6): p. 703.
92. Weber, B., et al., *CFD based compartment-model for a multiphase loop-reactor.* Chemical Engineering Science: X, 2019. **2**: p. 100010.
93. Barrasso, D., A. Tamrakar, and R. Ramachandran, *A reduced order PBM-ANN model of a multi-scale PBM-DEM description of a wet granulation process.* Chemical Engineering Science, 2014. **119**: p. 319-329.
94. Pokriefke, G.N.s.o.v.f.i.a.p.f.c.-r.t.s.e.l.P.P.-. and APA.
95. Eitzlmayr, A. and J. Khinast, *Co-rotating twin-screw extruders_ Detailed analysis of conveying elements based on smoothed particle hydrodynamics. Part 2_ Mixing.* Chemical Engineering Science, 2015. **134**(C): p. 880-886.
96. Metta, N., M. Ierapetritou, and R. Ramachandran, *A multiscale DEM-PBM approach for a continuous comilling process using a mechanistically developed breakage kernel.* Chemical Engineering Science, 2018. **178**: p. 211-221.
97. Metta, N., R. Ramachandran, and M. Ierapetritou, *A Computationally Efficient Surrogate-Based Reduction of a Multiscale Comill Process Model.* Journal of Pharmaceutical Innovation, 2019. **315**(3): p. 1-21.
98. Sampat, C., et al., *HPC enabled parallel, multi-scale & mechanistic model for high shear granulation using a coupled DEM-PBM framework.* Computer Aided Chemical Engineering, 2018. **44**: p. 1459-1464.
99. Sen, M., et al., *Mathematical Development and Comparison of a Hybrid PBM-DEM Description of a Continuous Powder Mixing Process.* Journal of Powder Technology, 2013. **2013**(7): p. 1-11.
100. R Ramachandran, P.B., *Effective parameter estimation within a multi-dimensional population balance model framework.* Chemical Engineering Science, 2010. **65**(16): p. 4884-4893.

101. Danckwerts, P.V., *Continuous flow systems: Distribution of residence times*. Chemical Engineering Science, 1953. **2**: p. 1-13.
102. Bhalode, P., Huayu Tian, Shashwat Gupta, Sonia M. Razavi, Andres Roman-Ospino, Shahrzad Talebian, Ravendra Singh, James V. Scicolone, Fernando J. Muzzio, and Marianthi Ierapetritou, *Using Residence time distribution in pharmaceutical solid dose manufacturing—a critical review*. International Journal of Pharmaceutics, 2021. **121248**.
103. Bhosekar, A. and M. Ierapetritou, *Advances in surrogate based modeling, feasibility analysis and optimization: A review*. Computers & Chemical Engineering, 2017. **108**: p. 250-267.
104. Chen, Y., and Marianthi Ierapetritou, *A framework of hybrid model development with identification of plant-model mismatch*. AIChE Journal, 2020. **66**(10).
105. Bascone, D., et al., *Hybrid Mechanistic-Empirical Approach to the Modeling of Twin Screw Feeders for Continuous Tablet Manufacturing*. Industrial & Engineering Chemistry Research, 2020.
106. Glassey, J. and M.v. Stosch, *Hybrid modeling in process industries*. 1 online resource (220 pages).
107. Portillo, P.M., F.J. Muzzio, and M. Ierapetritou, *Hybrid DEM-compartment modeling approach for granular mixing*. Trans. AIChE 2006. **53**(1): p. 119-128.
108. von Stosch, M., et al., *Hybrid semi-parametric modeling in process systems engineering: Past, present and future*. Computers and Chemical Engineering, 2014. **60**: p. 86-101.
109. Yang, S., et al., *Hybrid Modeling in the Era of Smart Manufacturing*. Computers and Chemical Engineering, 2020: p. 106874.
110. Zendehboudi, S., N. Rezaei, and A. Lohi, *Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review*. Applied Energy, 2018. **228**: p. 2539-2566.
111. Boukouvala, F., et al., *Computer-Aided Flowsheet Simulation of a Pharmaceutical Tablet Manufacturing Process Incorporating Wet Granulation*. Journal of Pharmaceutical Innovation, 2013. **8**(1): p. 11-27.
112. Boukouvala, F. and M. Ierapetritou, *Surrogate-based optimization of expensive flowsheet modeling for continuous pharmaceutical manufacturing*. Journal of Pharmaceutical Innovation, 2013. **8**(2): p. 131-145.
113. Boukouvala, F., et al., *An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process*. Computers & Chemical Engineering, 2012. **42**: p. 30-47.
114. Diab, S. and D.I. Gerogiorgis, *Design Space Identification and Visualization for Continuous Pharmaceutical Manufacturing*. Pharmaceutics, 2020. **12**(3): p. 235.
115. Escotet-Espinoza, M.S., et al., *Flowsheet models modernize pharmaceutical manufacturing design and risk assessment*. Pharmaceutical Technology, 2015. **39**(4): p. 34-42.
116. Galbraith, S.C., et al., *Integrated modeling of a continuous direct compression tablet manufacturing process: A production scale case study*. Powder Technology, 2019. **354**: p. 199-210.

117. Metta, N., et al., *Dynamic Flowsheet Model Development and Sensitivity Analysis of a Continuous Pharmaceutical Tablet Manufacturing Process Using the Wet Granulation Route*. Processes, 2019. **7**(4).
118. Rogers, A.J., C. Inamdar, and M. Ierapetritou, *An Integrated Approach to Simulation of Pharmaceutical Processes for Solid Drug Manufacture*. Industrial & Engineering Chemistry Research, 2013. **53**: p. 5128-5147.
119. Yoon, S., et al., *Flowsheet modeling of a continuous direct compression process*. Computer Aided Chemical Engineering, 2018. **41**: p. 121-139.
120. García-Muñoz, S., et al., *A flowsheet model for the development of a continuous process for pharmaceutical tablets: An industrial perspective*. AIChE Journal, 2017. **64**(2): p. 511-525.
121. Rogers, A. and M. Ierapetritou, *Challenges and Opportunities in Pharmaceutical Manufacturing Modeling and Optimization*. Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design, 2014. **34**: p. 144-149.
122. Tian, G., et al., *Development of an RTD-Based Flowsheet Modeling Framework for the Assessment of In-Process Control Strategies*. AAPS PharmSciTech, 2021. **22**(1): p. 1-10.
123. Yang, O., M. Qadan, and M. Ierapetritou, *Economic Analysis of Batch and Continuous Biopharmaceutical Antibody Production: a Review*. Journal of Pharmaceutical Innovation, 2020. **15**(1): p. 182-200.
124. Sen, M., et al., *Flowsheet optimization of an integrated continuous purification-processing pharmaceutical manufacturing operation*. Chemical Engineering Science, 2013. **102**: p. 56-66.
125. Ajmal, M., et al., *Calibration of cohesive DEM parameters under rapid flow conditions and low consolidation stresses*. Powder Technology, 2020. **374**: p. 22-32.
126. Boikov, A.V., R.V. Savelev, and V.A. Payor, *DEM Calibration Approach: Random Forest*. Journal of Physics: Conference Series, 2018. **1118**(1): p. 012009.
127. Boikov, A.V., R.V. Savelev, and V.A. Payor, *DEM Calibration Approach: design of experiment*. Journal of Physics: Conference Series, 2018. **1015**(3): p. 032017.
128. Boikov, A.V., et al., *DEM calibration approach: orthogonal experiment*. Journal of Physics: Conference Series, 2019. **1210**(1): p. 012025.
129. Coetzee, C., *Calibration of the discrete element method: Strategies for spherical and non-spherical particles*. Powder Technology, 2020. **364**: p. 851-878.
130. Coetzee, C.J., *Review: Calibration of the discrete element method*. Powder Technology, 2017. **310**: p. 104-142.
131. Coetzee, C.J. and D.N.J. Els, *Calibration of granular material parameters for DEM modelling and numerical verification by blade-granular material interaction*. Journal of Terramechanics, 2009. **46**(1): p. 15-26.
132. Do, H.Q., A.M. Aragon, and D.L. Schott, *A calibration framework for discrete element model parameters using genetic algorithms*. Advanced Powder Technology, 2018. **29**(6): p. 1393-1403.

133. Gao, Y., G. De Simone, and M. Koorapaty, *Calibration and verification of DEM parameters for the quantitative simulation of pharmaceutical powder compression process*. Powder Technology, 2021. **378**: p. 160-171.
134. Grobbel, J., et al., *Calibration of parameters for DEM simulations of solar particle receivers by bulk experiments and surrogate functions*. Powder Technology, 2020. **364**: p. 831-844.
135. Johnstone, M.W., *Calibration of DEM models for granular materials using bulk physical tests*. 2010, The University of Edinburgh.
136. Karkala, S., et al., *Calibration of Discrete-Element-Method Parameters for Cohesive Materials Using Dynamic-Yield-Strength and Shear-Cell Experiments*. Processes, 2019. **7**(5): p. 278.
137. Rackl, M., et al., *Efficient calibration of discrete element material model parameters using latin hypercube sampling and kriging*, in *ECCOMAS Congress 2016*. 2016: Crete Island, Greece.
138. Rackl, M. and K.J. Hanley, *A methodical calibration procedure for discrete element models*. 2017. **307**(C): p. 73-83.
139. Richter, C., et al., *Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials – Part II: Efficient optimization-based calibration*. Powder Technology, 2020. **360**: p. 967-976.
140. Roessler, T. and A. Katterfeld, *DEM parameter calibration of cohesive bulk materials using a simple angle of repose test*. Particuology, 2019. **45**: p. 105-115.
141. Roessler, T., et al., *Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials – part I: Solving the problem of ambiguous parameter combinations*. Powder Technology, 2019. **343**: p. 803-812.
142. Simons, T.A.H., et al., *A Ring Shear Tester as Calibration Experiment for DEM Simulations in Agitated Mixers – A Sensitivity Study*. Procedia Engineering, 2015. **102**: p. 741-748.
143. Abbott, B. and S.E. Lyshevski, *Signal Processing in MEMS Inertial Measurement Units for Dynamic Motion Control*. 2016 IEEE 36th International Conference on Electronics and Nanotechnology (Elnano), 2016: p. 309-314.
144. Roessler, T. and A. Katterfeld, *Scaling of the angle of repose test and its influence on the calibration of DEM parameters using upscaled particles*. Powder Technology, 2018. **330**: p. 58-66.
145. Jourdan, N., et al., *Compartmental Modelling in chemical engineering: A critical review*. Chemical Engineering Science, 2019. **210**.
146. Deen, N.G., et al., *Numerical Analysis of Solids Mixing in Pressurized Fluidized Beds*. Industrial & Engineering Chemistry Research, 2010. **49**(11): p. 5246-5253.
147. Godlieb, W., N.G. Deen, and J. Kuipers, *Characterizing solids mixing in DEM simulations*, in *6th International Journal on Multiphase Flow*. 2007: Leipzig, Germany.
148. Rogers, A. and M. Ierapetritou, *Feasibility and flexibility analysis of black-box processes Part 1: Surrogate-based feasibility analysis*. Chemical Engineering Science, 2015. **137**: p. 986-1004.

149. Singh, R., M. Ierapetritou, and R. Ramachandran, *System-wide hybrid MPC-PID control of a continuous pharmaceutical tablet manufacturing process via direct compaction*. *European Journal of Pharmaceutics and Biopharmaceutics*, 2013. **85**(3): p. 1164-1182.
150. Tian, G., et al., *Quality risk assessment and mitigation of pharmaceutical continuous manufacturing using flowsheet modeling approach*. *Computers and Chemical Engineering*, 2019. **129**: p. 106508.
151. Escotet-Espinoza, M.S., et al., *Effect of tracer material properties on the residence time distribution (RTD) of continuous powder blending operations. Part I of II: Experimental evaluation*. *Powder Technology*, 2019. **342**: p. 744-763.
152. Bérard, A., B. Blais, and G.S. Patience, *Experimental methods in chemical engineering: Residence time distribution—RTD*. *The Canadian Journal of Chemical Engineering*, 2019. **98**(4): p. 848-867.
153. Blackshields, C.A. and A.M. Crean, *Continuous powder feeding for pharmaceutical solid dosage form manufacture: a short review*. *Pharmaceutical Development and Technology*, 2017.
154. Boukouvala, F., Niotis, V., Ramachandran, R., Muzzio, F. J. & Ierapetritou, M. , *An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process*. . *Computers & Chemical Engineering*, 2012. **42**: p. 30–47.
155. Rogers, A.J., A. Hashemi, and M. Ierapetritou, *Modeling of Particulate Processes for the Continuous Manufacture of Solid-Based Pharmaceutical Dosage Forms*. *Processes*, 2013. **1**(2): p. 67-127.
156. Toson, P., et al., *Model-based approach to the design of pharmaceutical roller-compaction processes*. *International Journal of Pharmaceutics: X*, 2019. **1**: p. 100005.
157. Hopkins, M., *Loss in weight feeder systems*. *Measurement and Control*, 2006. **39**(8): p. 237-240.
158. Engisch, W.E. and F.J. Muzzio, *Feedrate deviations caused by hopper refill of loss-in-weight feeders*. *Powder Technology*, 2015. **283**: p. 389-400.
159. Engisch, W.E. and F.J. Muzzio, *Method for characterization of loss-in-weight feeder equipment*. *Powder Technology*, 2012. **228**: p. 395-403.
160. Heckel, R.W., *Density pressure relationships in powder compaction*. *Trans. Metall. Soc. AIME*, 1961(221): p. 671-675.
161. Cundall, P.A. and O.D.L. Strack, *A discrete numerical model for granular assemblies*. *Géotechnique*, 1980. **30**(3): p. 331-336.
162. Boukouvala, F., Dubey, A., Vanarase A., Ramachandran, R., Muzzio, F., Ierapetritou, M., *Computational Approaches for Studying the Granular Dynamics of Continuous Blending Processes, 2 – Population Balance and Data-Based Methods*. . *Macromolecular Materials and Engineering*, 2012. **297**: p. 9-19.
163. Boukouvala, F., et al., *Reduced-order discrete element method modeling*. *International Journal of Pharmaceutics*, 2013. **95**: p. 12-26.
164. Deng, X., et al., *Discrete element method simulation of a conical screen mill: A continuous dry coating device*. *Chemical Engineering Science*, 2015. **125**: p. 58-74.

165. Gao, Y., M. Ierapetritou, and F.J. Muzzio, *Periodic section modeling of convective continuous powder mixing processes*. Trans. AIChE 2012. **58**(1): p. 69-78.
166. Kureck, H., et al., *Industrial Scale Simulations of Tablet Coating using GPU based DEM: A Validation Study*. Chemical Engineering Science, 2019.
167. Portillo, P.M., F.J. Muzzio, and M. Ierapetritou, *Using Compartment Modeling to Investigate Mixing Behavior of a Continuous Mixer*. Journal of Pharmaceutical Innovation, 2008. **3**(3): p. 161-174.
168. Sen, M., et al., *Analyzing the Mixing Dynamics of an Industrial Batch Bin Blender via Discrete Element Modeling Method*. Processes, 2017. **5**(4): p. 22.
169. Toson, P. and J. Khinast, *Particle-level residence time data in a twin-screw feeder*. Data in Brief, 2019. **27**: p. 104672.
170. Toson, P., et al., *Detailed modeling and process design of an advanced continuous powder mixer*. International Journal of Pharmaceutics, 2018. **552**(1-2): p. 288-300.
171. Simons, T.A.H., et al., *A ring shear tester as calibration experiment for DEM simulations in agitated mixers—a sensitivity study*. Procedia Engineering, 2015. **102**: p. 741-748.
172. Derakhshani, S.M., D.L. Schott, and G. Lodewijks, *Micro–macro properties of quartz sand: Experimental investigation and DEM simulation*. Powder Technology, 2015. **269**: p. 127-138.
173. Wilkinson, S.K., et al., *A parametric evaluation of powder flowability using a Freeman rheometer through statistical and sensitivity analysis: A discrete element method (DEM) study*. Computers and Chemical Engineering, 2017. **97**: p. 161-174.
174. Zhou, H., et al., *Calibration of DEM models for irregular particles based on experimental design method and bulk experiments*. 2018. **332**: p. 210-223.
175. Yan, Z., et al., *Discrete element modelling (DEM) input parameters: understanding their impact on model predictions using statistical analysis*. Computational Particle Mechanics, 2015. **2**(3): p. 283-299.
176. Chung, Y.C., et al., *Mechanical behaviour of a granular solid and its contacting deformable structure under uni-axial compression – Part I: Joint DEM–FEM modelling and experimental validation*. Chemical Engineering Science, 2016. **144**: p. 404-420.
177. Ucgul, M., J.M. Fielke, and C. Saunders, *Three-dimensional discrete element modelling of tillage: Determination of a suitable contact model and parameters for a cohesionless soil*. Biosystems Engineering, 2014. **121**: p. 105-117.
178. Ketterhagen, W.R., Ende, M. T. A. & Hancock, B. C. , *Process modeling in the pharmaceutical industry using the discrete element method*. Journal of Pharmaceutical Sciences 2009. **98**: p. 442–470.
179. Hertz, H., *On the contact of rigid elastic solids*. Journal fur die Reine und Angewandte Mathematik, 1882: p. 156-171.
180. Mindlin, R.D., *Compliance of elastic bodies in contact*. ASME Journal of Applied Mechanics, 1949(16): p. 259-268.
181. Mindlin, R.D. and H. Deresiewicz, *Elastic Spheres in Contact under Varying Oblique Force*. ASME Journal of Applied Mechanics, 1953(20): p. 327-344.

182. Hare, C., et al., *Analysis of the dynamics of the FT4 powder rheometer*. Powder Technology, 2015. **285**: p. 123-127.
183. Yan, Z., et al., *Investigating mixing and segregation using discrete element modelling (DEM) in the Freeman FT4 rheometer*. International Journal of Pharmaceutics, 2016. **513**(1-2): p. 38-48.
184. Bharadwaj, R., W.R. Ketterhagen, and B.C. Hancock, *Discrete element simulation study of a Freeman powder rheometer*. Chemical Engineering Science, 2010. **65**(21): p. 5747-5756.
185. Benvenuti, L., C. Kloss, and S. Pirker, *Identification of DEM simulation parameters by Artificial Neural Networks and bulk experiments*. Powder Technology, 2016. **291**: p. 456-465.
186. Bhalode, P. and M. Ierapetritou, *Discrete Element Modeling (DEM) Parametric Study of Feeder Unit in Continuous Pharmaceutical Industry*. Computer Aided Chemical Engineering, 2019. **47**: p. 341-346.
187. Dubey, A., et al., *Computational Approaches for Studying the Granular Dynamics of Continuous Blending Processes, 1 – DEM Based Methods*. Macromolecular Materials and Engineering, 2011. **296**(3-4): p. 290-307.
188. Radeke, C., B.J. Glasser, and J. Khinast, *Large-scale powder mixer simulations using massively parallel GPU architectures*. Chemical Engineering Science, 2010. **65**(24): p. 6435-6442.
189. Remy, B., J.G. Khinast, and B.J. Glasser, *Discrete element simulation of free flowing grains in a four-bladed mixer*. Trans. AIChE 2009. **55**(8): p. 2035-2048.
190. Sarkar, A. and C. Wassgren, *Comparison of flow microdynamics for a continuous granular mixer with predictions from periodic slice DEM simulations*. Powder Technology, 2012(221): p. 325-336.
191. Jones, B., *21st century screening experiments: What, why, and how*. Quality Engineering, 2016. **28**(1): p. 98-106.
192. Jones, B. and C.J. Nachtsheim, *Effective Design-Based Model Selection for Definitive Screening Designs*. Technometrics, 2017.
193. Jones, B. and C.J. Nachtsheim, *A class of Three-level Designs for Definitive Screening in the Presence of Second-Order Effects*. Analytical Chemistry, 2011. **43**(1).
194. Akaike, H., *A new look at the statistical model identification*. IEEE Transactions on Automatic Control, 1974. **19**(6): p. 716-723.
195. Schwarz, G., *Estimating the Dimension of a Model*. The Annals of Statistics, 1978. **6**(2): p. 461-464.
196. Tawarmalani, M. and N.V. Sahinidis, *A polyhedral branch-and-cut approach to global optimization*. Mathematical Programming, 2005. **103**(2): p. 225-249.
197. Fogler, H.S., *Essentials of Chemical Reaction Engineering*. 2010: Pearson Education.
198. Rodrigues, A.E., *Residence time distribution (RTD) revisited*. Chemical Engineering Science, 2021. **230**: p. 116188.

199. Davis, M.E. and R.J. Davis, *Fundamentals of chemical reaction engineering. McGraw-Hill chemical engineering series*. 2003, New York, NY: McGraw-Hill Higher Education.
200. Levenspiel, O., *Chemical reaction engineering*. 1999: John Wiley & Sons.
201. Gao, Y.J., F.J. Muzzio, and M.G. Ierapetritou, *A review of the Residence Time Distribution (RTD) applications in solid unit operations*. Powder Technology, 2012. **228**: p. 416-423.
202. Martin, A.D., *Interpretation of residence time distribution data*. Chemical Engineering Science, 2000. **55**(23): p. 5907-5917.
203. MacMullin, R.B. and M. Weber, *The Theory of Short-Circuiting in Continuous-Flow Mixing Vessels in Series and Kinetics of Chemical Reactions in Such Systems*. . Trans. AIChE 1935. **31**(2): p. 409-458.
204. Nauman, E.B., *Residence Time Theory*. Industrial & Engineering Chemistry Research, 2008. **47**(10): p. 3752-3766.
205. Nauman, E.B., *Residence time distribution theory for unsteady stirred tank reactors*. Chemical Engineering Science, 1969. **24**(9): p. 1461-1470.
206. Todd, D.B., *Residence time distribution in twin-screw extruders*. Polymer Engineering & Science, 1975. **15**(6): p. 437-443.
207. Trachsel, F., Günther, A., Khan, S., & Jensen, K. F., *Measurement of residence time distribution in microfluidic systems*. Chemical Engineering Science, 2005. **60**(21): p. 5729-5737.
208. Torres, A.P., & Oliveira, F. A. R., *Residence time distribution studies in continuous thermal processing of liquid foods: a review*. Journal of Food Engineering, 1998. **36**(1): p. 1-30.
209. Le Moullec, Y., Potier, O., Gentric, C., & Leclerc, J. P. , *Flow field and residence time distribution simulation of a cross-flow gas-liquid wastewater treatment reactor using CFD*. Chemical Engineering Science, 2008. **63**(9): p. 2436-2449.
210. Engisch, W. and F. Muzzio, *Using Residence Time Distributions (RTDs) to Address the Traceability of Raw Materials in Continuous Pharmaceutical Manufacturing*. Journal of Pharmaceutical Innovation, 2016. **11**(1): p. 64-81.
211. Mangal, H. and P. Kleinebudde, *Experimental determination of residence time distribution in continuous dry granulation*. International Journal of Pharmaceutics, 2017. **524**(1-2): p. 91-100.
212. Gao, Y., F. Muzzio, and M.G. Ierapetritou, *A review of the Residence Time Distribution (RTD) applications in solid unit operations*. Powder Technology, 2012. **228**: p. 416-423.
213. Karttunen, A.P., et al., *Measurement of residence time distributions and material tracking on three continuous manufacturing lines*. International Journal of Pharmaceutics, 2019. **563**: p. 184-197.
214. Pauli, V., P. Kleinebudde, and M. Krumme, *From powder to tablets: Investigation of residence time distributions in a continuous manufacturing process train as basis for continuous process verification*. European Journal of Pharmaceutics and Biopharmaceutics, 2020. **153**: p. 200-210.

215. U.S. Food and Drug Administration, 2004. *Pharmaceutical CGMPs for the 21st Century-A Risk-Based Approach*.
216. Fisher, A.C., et al., *Advancing pharmaceutical quality: An overview of science and research in the U.S. FDA's Office of Pharmaceutical Quality*. International Journal of Pharmaceutics, 2016. **515**(1-2): p. 390-402.
217. J., W., *The concept of pharmaceutical quality*. . Am Pharm Rev, 2004: p. 1-3.
218. Yu, L.X., et al., *Understanding Pharmaceutical Quality by Design*. Aaps Journal, 2014. **16**(4): p. 771-783.
219. Garcia, T., G. Cook, and R. Nosal, *PQLI Key Topics - Criticality, Design Space, and Control Strategy*. Journal of Pharmaceutical Innovation, 2008. **3**(2): p. 60-68.
220. Potter, C., *PQLI Application of Science- and Risk-based Approaches (ICH Q8, Q9, and Q10) to Existing Products*. Journal of Pharmaceutical Innovation, 2009. **4**(1): p. 4-23.
221. Berridge, J.C., *PQLI ®: Current Status and Future Plans*. Journal of Pharmaceutical Innovation, 2009. **4**(1): p. 1-3.
222. Escotet-Espinoza, M.S., et al., *Using a material property library to find surrogate materials for pharmaceutical process development*. Powder Technology, 2018. **339**: p. 659-676.
223. Wang, Y., *Using Multivariate Analysis for Pharmaceutical Drug Product Development*. 2017(April).
224. Wang, Y., et al., *Development and applications of a material library for pharmaceutical continuous manufacturing of solid dosage forms*. International Journal of Pharmaceutics, 2019. **569**: p. 118551.
225. Razavi, S.M.R.-O., Andres D. ; Bhalode, Pooja; Scicolone, James ; Callegari, Gerardo; Dubey, Atul ; Ierapetritou, Marianthi; Muzzio, Fernando, *Implementation of a Methodology for Selection of an Appropriate Tracer to Measure the Residence Time Distribution (RTD) of Continuous Powder Blending Operations*. To be submitted, 2022.
226. Escotet-Espinoza, M.S., et al., *Effect of material properties on the residence time distribution (RTD) characterization of powder blending unit operations. Part II of II: Application of models*. Powder Technology, 2019. **344**: p. 525-544.
227. Berntsson, O., *Characterization and application of near infrared reflection spectroscopy for quantitative process analysis of powder mixtures*. 2001: Kemi.
228. Andrés D Román-Ospino, Y.B., Jingzhe Li, Jenny Vargas, Benoît Igne, Simon Bate, Davinia Brouckaert, Fabien Chauchard, Douglas Hausner, Rohit Ramachandran, Ravendra Singh, Fernando J Muzzio, *Sampling optimization for blend monitoring of a low dose formulation in a tablet press feed frame using spatially resolved near-infrared spectroscopy*. International Journal of Pharmaceutics, 2021. **602**: p. 120594.
229. Román-Ospino, A.D., et al., *Characterization of NIR interfaces for the feeding and in-line monitoring of a continuous granulation process*. International Journal of Pharmaceutics, 2020. **574**: p. 118848.

230. Meng, W., et al., *Advanced process design and understanding of continuous twin-screw granulation via implementation of in-line process analytical technologies*. *Advanced Powder Technology*, 2019. **30**(4): p. 879-894.
231. Hetrick, E.M., et al., *Development of Near Infrared Spectroscopy-based Process Monitoring Methodology for Pharmaceutical Continuous Manufacturing Using an Offline Calibration Approach*. *Analytical Chemistry*, 2017. **89**: p. 9175-9183.
232. Berntsson, O., et al., *Quantitative in-line monitoring of powder blending by near infrared reflection spectroscopy*. *Powder Technology*, 2002. **123**(2-3): p. 185-193.
233. O'Haver, T., *A pragmatic introduction to signal processing* 1997, University of Maryland at College Park.
234. Rabiner, L.R., and Bernard Gold., *Theory and application of digital signal processing*. 1975, Englewood Cliffs: Prentice-Hall.
235. Orfanidis, S.J., *Introduction to signal processing*. 1995: Prentice-Hall, Inc.
236. Oppenheim, A.V., *Applications of digital signal processing*. 1978, Englewood Cliffs
237. T. Huang, G.Y., G. Tang, *A fast two-dimensional median filtering algorithm*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1979. **27**(1): p. 13 - 18.
238. Betts, J.A., *Signal processing, modulation and noise*. 1970: English universities Press.
239. Tuzlukov, V., *Signal Processing Noise*. *Electrical Engineering and Applied Signal Processing Series*. Vol. 688. 2010: CRC Press.
240. McKinney, W., & others., *Data structures for statistical computing in python*. In *Proceedings of the 9th Python in Science Conference*, 2010. **445**: p. 51-56.
241. *MathWorks, (2021). Signal Processing Toolbox: User's Guide (R2021a)*. Retrieved Oct 18, 2021 from https://www.mathworks.com/help/pdf_doc/signal/signal.pdf. 2021.
242. Abraham, S.M.J.G., *Smoothing and Differentiation of Data by Simplified Least Square Procedures*. *Analytical Chemistry*, 1964. **36**(8): p. 1627-1639.
243. Press, W.H. and S.A. Teukolsky, *Savitzky-Golay Smoothing Filters*. *Computers in Physics*, 1990. **4**(6).
244. Schafer, R.W., *What Is a Savitzky-Golay Filter? [Lecture Notes]*. *IEEE Signal Processing Magazine*, 2011. **28**(4): p. 111-117.
245. **Bellanger, M.**, *Digital processing of signals* 2000: Wiley.
246. Pauli Virtanen, R.G., Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E.A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors., *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*. *Nature Methods*, 2020. **17**(3): p. 261-272.
247. Heideman, M.T.J., Don H.; Burrus, Charles Sidney *Gauss and the history of the fast Fourier transform*. *IEEE ASSP Magazine*, 1984. **1**(4): p. 14-21.
248. Strang, G., *Wavelets*. *American Scientist*, 1994. **82**(3): p. 250-255.

249. Cochran, W.T., James W. Cooley, David L. Favon, Howard D. Helms, Reginald A. Kaenel, William W. Lang, George C. Maling, David E. Nelson, Charles M. Rader, and Peter D. Welch., *What is the fast Fourier transform?* Proceedings of the IEEE. **55**(10): p. 1664-1674.
250. Nussbaumer, H.J., *The fast Fourier transform*, in *Fast Fourier Transform and Convolution Algorithms*. 1981, Springer: Berlin, Heidelberg.
251. Cooley JW, L.P., Welch PD., *The fast Fourier transform and its applications*. IEEE Transactions on Education, 1969. **12**(1): p. 27-34.
252. Addison, P.S., *Wavelet transforms and the ECG: a review*. Physiological measurement, 2005. **26**(5).
253. Alfaouri M, D.K., *ECG signal denoising by wavelet transform thresholding*. American Journal of applied sciences, 2008. **5**(3): p. 276-281.
254. L. Angrisani, P.D., M. D'Apuzzo and A. Testa, *A measurement method based on the wavelet transform for power quality analysis*. IEEE Transactions on Power Delivery, 1998. **13**(4): p. 990-998.
255. Peng, Z.K. and F.L. Chu, *Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography*. Mechanical Systems and Signal Processing, 2004. **18**(2): p. 199-221.
256. Madaleno M, P.C., *International stock market indices comovements: A new look*. International Journal of Finance & Economics, 2012. **17**(1): p. 89-102.
257. MJ, S., *The discrete wavelet transform: wedding the a trous and Mallat algorithms*. IEEE Transactions on signal processing, 1992. **40**(10): p. 2464-2482.
258. Rioul O, D.P., *Fast algorithms for discrete and continuous wavelet transforms*. IEEE Transactions on Information Theory, 1992. **38**(2): p. 569-586.
259. *Wavelet Toolbox - Mathworks*. Available from: <https://www.mathworks.com/products/wavelet.html>.
260. *Introduction to Wavelet Families - Mathworks*. Available from: <https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html>.
261. Lee, G., Gommers, R., Waselewski, F., Wohlfahrt, K., & O'Leary, A., *PyWavelets: A Python package for wavelet analysis*. Journal of Open Source Software, 2019. **4**(36): p. 12-37.
262. *Wavelet Browser by Pywavelets*. Available from: <http://wavelets.pybytes.com>.
263. Downie, T.R. and B.W. Silverman, *The discrete multiple wavelet transform and thresholding methods*. IEEE Transactions on Signal Processing, 1998. **46**(9): p. 2558 - 2561.
264. Abramovich, F., Y. Benjamini, D. L. Donoho, and I. M. Johnstone, *Adapting to Unknown Sparsity by Controlling the False Discovery Rate*. Annals of Statistics, 2006. **34**(2): p. 584-653.
265. Donoho, D.L., *De-noising by Soft-Thresholding*. IEEE Transactions on Information Theory, 1995. **42**(3): p. 613-627.
266. *MathWorks, (2021). Wavelet Toolbox: User's Guide (R2021b)*. Retrieved Oct 18, 2021 from https://www.mathworks.com/help/pdf_doc/wavelet/wavelet Ug.pdf.

267. Patrick Egan (2021). fwhm (<https://www.mathworks.com/matlabcentral/fileexchange/10590-fwhm>), M.C.F.E.R.O., 2021.
268. O'Connor, T.F., Yu, L. X. & Lee, S. L. , *Emerging technology: A key enabler for modernizing pharmaceutical manufacturing and advancing product quality*. . International Journal of Pharmaceutics, 2016. **509**: p. 492-498.
269. Tian, G., et al., *A dimensionless analysis of residence time distributions for continuous powder mixing*. Powder Technology, 2017. **315**: p. 332-338.
270. V. Vanhoorne, C.V., *Recent progress in continuous manufacturing of oral solid dosage forms*. International Journal of Pharmaceutics, 2020. **579**: p. 119-194.
271. Ismail, H.Y., et al., *Developing ANN-Kriging hybrid model based on process parameters for prediction of mean residence time distribution in twin-screw wet granulation*. Powder Technology, 2019. **343**: p. 568-577.
272. Wesholowski, J., A. Berghaus, and M. Thommes, *Inline Determination of Residence Time Distribution in Hot-Melt-Extrusion*. Pharmaceutics, 2018. **10**(2): p. 49.
273. Kumar, A., et al., *Linking granulation performance with residence time and granulation liquid distributions in twin-screw granulation: An experimental investigation*. European Journal of Pharmaceutical Sciences, 2016. **90**: p. 25-37.
274. Kumar, A., et al., *Conceptual framework for model-based analysis of residence time distribution in twin-screw granulation*. European Journal of Pharmaceutical Sciences, 2015. **71**: p. 25-34.
275. Chen, H., Diep, E., Langrish, T.A. and Glasser, B.J.,, *Continuous fluidized bed drying: Residence time distribution characterization and effluent moisture content prediction*. AIChE Journal, 2020. **66**(5): p. 16902.
276. Gyurkes, M., et al., *Process Design of Continuous Powder Blending Using Residence Time Distribution and Feeding Models*. Pharmaceutics, 2020. **12**(11): p. 1119.
277. Osorio, J.G. and F.J. Muzzio, *Effects of processing parameters and blade patterns on continuous pharmaceutical powder mixing*. 2016.
278. Palmer, J., et al., *Mapping key process parameters to the performance of a continuous dry powder blender in a continuous direct compression system*. Powder Technology, 2020. **362**: p. 659-670.
279. Vanarase, A.U. and F.J. Muzzio, *Effect of operating conditions and design parameters in a continuous powder mixer*. Powder Technology, 2011. **208**(1): p. 26-36.
280. Van Snick, B., et al., *Impact of material properties and process variables on the residence time distribution in twin screw feeding equipment*. International Journal of Pharmaceutics, 2019. **556**: p. 200-216.
281. Dulle, M., H. Ozcoban, and C.S. Leopold, *Analysis of the powder behavior and the residence time distribution within a production scale rotary tablet press*. European Journal of Pharmaceutical Sciences, 2018. **125**: p. 205-214.
282. Dulle, M., H. Ozcoban, and C.S. Leopold, *Influence of the feed frame design on the powder behavior and the residence time distribution*. International Journal of Pharmaceutics, 2019. **565**: p. 523-532.

283. Furukawa, R., R. Singh, and M. Ierapetritou, *Effect of material properties on the residence time distribution (RTD) of a tablet press feed frame*. International Journal of Pharmaceutics, 2020. **591**: p. 119961.
284. Puckhaber, D., et al., *Impact of Particle and Equipment Properties on Residence Time Distribution of Pharmaceutical Excipients in Rotary Tablet Presses*. Pharmaceutics, 2020. **12**(3): p. 283.
285. Tanimura, S., et al., *Residence time distribution modeling and in line monitoring of drug concentration in a tablet press feed frame containing dead zones*. Chemical Engineering Research and Design, 2020.
286. Galbraith, S.C., et al., *Linking process variables to residence time distribution in a hybrid flowsheet model for continuous direct compression*. Chemical Engineering Research and Design, 2020(153): p. 85-95.
287. Karttunen, A.-P., et al., *Measurement of residence time distributions and material tracking on three continuous manufacturing lines*. International Journal of Pharmaceutics, 2019. **563**: p. 184-197.
288. Kruisz, J., et al., *Material tracking in a continuous direct capsule-filling process via residence time distribution measurements*. International Journal of Pharmaceutics, 2018. **550**(1-2): p. 347-358.
289. Billups, M. and R. Singh, *Systematic Framework for Implementation of Material Traceability into Continuous Pharmaceutical Tablet Manufacturing Process*. Journal of Pharmaceutical Innovation, 2020. **15**(1): p. 51-65.
290. Peterwitz, M. and G. Schembecker, *Evaluating the potential for optimization of axial back-mixing in continuous pharmaceutical manufacturing*. Computers and Chemical Engineering, 2021. **147**: p. 107251.
291. Bhalode, P.R., Sonia M.; Muzzio, Fernando J.; Ierapetritou, Marianthi, *Statistical data pre-treatment and noise removal for residence time distribution studies in pharmaceutical manufacturing*. To be submitted, 2022.
292. Fonteyne, M., et al., *Process analytical technology for continuous manufacturing of solid-dosage forms*. TrAC Trends in Analytical Chemistry, 2015. **67**: p. 157-166.
293. De Leersnyder, F., et al., *Evaluation of an in-line NIR spectroscopic method for the determination of the residence time in a tablet press*. International Journal of Pharmaceutics, 2019. **565**: p. 358-366.
294. Vargas, J.M., et al., *Process analytical technology in continuous manufacturing of a commercial pharmaceutical product*. International Journal of Pharmaceutics, 2018. **538**(1-2): p. 167-178.
295. Román-Ospino, A.D., Ravendra Singh, Marianthi Ierapetritou, Rohit Ramachandran, Rafael Méndez, Carlos Ortega-Zuñiga, Fernando J. Muzzio, and Rodolfo J. Románach. , *Near infrared spectroscopic calibration models for real time monitoring of powder density*. International journal of pharmaceutics, 2016. **512**(1): p. 61-74.
296. Luypaert, J., D.L. Massart, and Y.V. Heyden, *Near-infrared spectroscopy applications in pharmaceutical analysis*. Talanta, 2007. **72**(3): p. 865-883.