AAPS Advances in the Pharmaceutical Sciences Series 42

Zoltan K. Nagy
Arwa El Hagrasy
Jim Litster *Editors*

Continuous Pharmaceutical Processing

aaps®

Springer

AAPS Advances in the Pharmaceutical Sciences Series

Volume 42

Series Editor

Yvonne Perrie
Strathclyde Institute of Pharmacy, University of Strathclyde,
Bearsden, Dunbartonshire, UK

The AAPS Advances in the Pharmaceutical Sciences Series, published in partnership with the American Association of Pharmaceutical Scientists, is designed to deliver volumes authored by opinion leaders and authorities from around the globe, addressing innovations in drug research and development, and best practice for scientists and industry professionals in the pharma and biotech industries.

More information about this series at http://www.springer.com/series/8825

Zoltan K. Nagy • Arwa El Hagrasy • Jim Litster
Editors

Continuous Pharmaceutical Processing

aaps®

Springer

Editors
Zoltan K. Nagy
Davidson School of Chemical Engineering
Purdue University
West Lafayette, IN, USA

Department of Chemical Engineering
Loughborough University
Loughborough, UK

Jim Litster
Department of Chemical and Biochemical Engineering
The University of Sheffield
Sheffield, UK

Arwa El Hagrasy
Faculty of Pharmacy
Cairo University
Cairo, Egypt

OPQ/CDER/FDA
Office of Pharmaceutical Manufacturing Assessment
Silver Spring, MD, USA

The opinions and conclusions expressed in this book are solely the views of the authors and do not necessarily reflect those of the Food and Drug Administration or the U.S. Government.

ISSN 2210-7371 ISSN 2210-738X (electronic)
AAPS Advances in the Pharmaceutical Sciences Series
ISBN 978-3-030-41523-5 ISBN 978-3-030-41524-2 (eBook)
https://doi.org/10.1007/978-3-030-41524-2

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The pharmaceutical industry for decades has been dominated by the batch manufacturing of pharmaceutical products. The high profitability characterizing the industry with a focus on blockbuster drugs hampered innovation or taking risk in developing new manufacturing technologies. More recently, however, it is increasingly recognized that material and manufacturing costs during drug development are significant, and the increasingly competitive environment with lower probabilities of discovering new blockbusters, and the continuously increasing market share of the generics industry, has incentivized all pharmaceutical industries, innovator companies, and generics alike, to innovate their manufacturing processes. It is now widely accepted that continuous manufacturing is a key enabling technology to implement process intensification in pharmaceutical manufacturing. Continuous manufacturing provides higher yields, lower utility consumption, reduced waste, and smaller footprint, enabling the pharmaceutical industry to move away from stepwise and time-consuming batch processing to fully integrated, environmentally friendlier, and closely controlled manufacturing systems with increased flexibility. Continuous manufacturing allows faster product development and can produce higher quality products, with excellent product consistency at lower cost, and enhance drug safety, providing significant advantages to governments, companies, and patients alike.

Because of its inherent advantages, continuous pharmaceutical manufacturing is currently receiving much interest from industry and regulatory authorities, with the joint aim of allowing rapid access of novel therapeutics and existing medications to the public, without compromising high quality.

Research groups from different academic institutions have significantly contributed to this field with an immense amount of published research addressing a variety of topics related to continuous processing. This book is structured to have individual chapters on the different continuous unit operations involved in drug substance and drug product manufacturing. A wide spectrum of topics are covered, including basic principles of continuous manufacturing, applications of continuous flow chemistry in drug synthesis, understanding residence time distribution, and

impact of surge vessels in continuous manufacturing, continuous crystallization, feeders and blenders, roll compaction, continuous wet granulation, and continuous drying. The underlying theme for each of these chapters is to present to the reader the recent advances in modeling, experimental investigations, and equipment design as they pertain to each individual unit operation. A separate chapter is also dedicated to practical considerations for continuous drug substance manufacturing, and the book also includes chapters on quality by design (QbD) and process analytical technology (PAT) for continuous processing, process control strategies including new concepts of quality-by-control (QbC), real-time optimization and process management, business and supply chain considerations related to continuous manufacturing, as well as safety guidelines related to continuous chemistry. A separate chapter is dedicated to discussing regulatory aspects of continuous manufacturing, with description of current regulatory environment quality/GMP aspects, as well as regulatory gaps and challenges.

Our aim in publishing this book is to make it a valuable reference for readers interested in continuous pharmaceutical manufacturing, with a desire to gain both fundamental knowledge of engineering principles and mechanistic studies utilized in understanding and developing continuous processes, but also to receiving practical guidelines that can be invaluable in the successful implementation and operation of robust continuous manufacturing systems aligned with regulatory requirements. In addition, our advanced readers and practitioners in this field will find that the technical content of *Continuous Pharmaceutical Processing* is at the forefront of recent technological advances, with coverage of future prospects and challenges for this technology.

West Lafayette, IN, USA — Zoltan K. Nagy
Silver Spring, MD, USA — Arwa El Hagrasy
Sheffield, UK — Jim Litster

Contents

1 Basic Principles of Continuous Manufacturing 1
Sudarshan Ganesh and Gintaras V. Reklaitis

2 Continuous Reactors for Pharmaceutical Manufacturing 23
Martin D. Johnson, Scott A. May, Michael E. Kopach, Jennifer Mc Clary Groh, Timothy Donald White, Kevin P. Cole, Timothy Braden, Luke P. Webster, and Vaidyaraman Shankarraman

3 Understanding Residence Time, Residence Time Distribution, and Impact of Surge Vessels 51
Martin D. Johnson, Scott A. May, Jennifer Mc Clary Groh, Luke P. Webster, Vaidyaraman Shankarraman, Richard D. Spencer, Carla Vanesa Luciani, Christopher S. Polster, and Timothy Braden

4 Intermittent Flow and Practical Considerations for Continuous Drug Substance Manufacturing 87
Martin D. Johnson, Scott A. May, Jennifer McClary Groh, Timothy Braden, and Richard D. Spencer

5 Continuous Crystallization: Equipment and Operation 129
Yiqing C. Liu and Zoltan K. Nagy

6 Continuous Feeding-Blending in Pharmaceutical Continuous Manufacturing 193
Qinglin Su, Gintaras V. Reklaitis, and Zoltan K. Nagy

7 Recent Progress in Roll Compaction Process Development for Pharmaceutical Solid Dosage Form Manufacture 227
Ariel R. Muliadi, Alamelu Banda, and Chen Mao

8 Continuous Wet Granulation 269
Arwa El Hagrasy, Li Ge Wang, and Jim Litster

9 Continuous Fluidized Bed Drying: Advanced Modeling and Experimental Investigations 301
Ibrahim Alaathar, Stefan Heinrich, and Ernst-Ulrich Hartge

10 Statistical Methods in Quality by Design and Process Analytical Technologies for Continuous Processes to Enable Real-Time Release 361
Vanessa Cárdenas, Juan G. Rosas, Carlos Pinzón, and Rodolfo J. Romañach

11 Active Process Control in Pharmaceutical Continuous Manufacturing – The Quality by Control (QbC) Paradigm 395
Qinglin Su, Sudarshan Ganesh, Gintaras V. Reklaitis, and Zoltan K. Nagy

12 Real-Time Optimization: How to Change Setpoints in Pharmaceutical Manufacturing 429
Arun Giridhar and Gintaras V. Reklaitis

13 Safety Guidelines for Continuous Processing 441
Martin D. Johnson and Jeffry Niemeier

14 Evaluating the Business Case for Continuous Manufacturing of Pharmaceuticals: A Supply Network Perspective 477
Jagjit Singh Srai, Ettore Settanni, and Parminder Kaur Aulakh

15 Regulatory Considerations for Continuous Manufacturing 513
Elaine Morefield

Index 537

Chapter 1
Basic Principles of Continuous Manufacturing

Sudarshan Ganesh and Gintaras V. Reklaitis

Abstract Continuous manufacturing in the pharmaceutical industry is an emerging technology, although it is widely practiced in industries such as petrochemical, bulk chemical, foods, and mineral processing. This chapter briefly discusses the characteristics of continuous manufacturing at the conceptual level, first, in its generic form, viewing the process as a unitary system, and then as a system composed of multiple manufacturing unit operations. Key requirements for implementing an effective continuous process are reviewed, while aspects specific to pharmaceutical applications are highlighted. The advantages and limitations of continuous manufacturing are discussed and compared to the advantages and limitations of the batch operating mode, which has been the mainstay of the pharmaceutical industry. Perspectives on advancing pharmaceutical manufacturing in the Industry 4.0 era are discussed.

Keywords Pharmaceutical manufacturing · Continuous processing · Batch processing · Fundamentals

1.1 Introduction

Continuous manufacturing has been receiving increasing attention in the pharmaceutical industry driven by the expectation of achieving reduced operating and capital costs, improved product quality, and increased reliability (Lee et al. 2015). While this mode of manufacture is new to the pharmaceutical industry, it is widely practiced in many industry sectors, such as refining and petrochemical, bulk chemical, and food and minerals processing. It most commonly involves the processing of

S. Ganesh · G. V. Reklaitis (✉)
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA
e-mail: reklaiti@purdue.edu

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_1

fluids, liquid or gases, although particulate and granular materials and suspensions are also handled. In these industries, the continuous manufacturing plant or line is usually dedicated to a specific product and is typically operated without interruption around the clock with only infrequent shutdown to perform maintenance functions or in case of emergency. A continuous manufacturing line is normally designed for a nominal production rate and while that rate can be reduced within a limited range, typically further reductions lead to unsatisfactory product outputs or damage to equipment. Generally, continuous manufacturing facilities enjoy economies of scale; that is, the investment and operating cost per unit of production decrease as the plant design capacity is increased.

The incentives for continuous manufacturing in the pharmaceutical industry are not the same in all aspects as they may be for the other industry sectors and, thus, it is important to understand what the essential elements of the continuous manufacturing mode are, and which aspects are really introduced to adapt to the needs of a specific industry sector. Thus, in this chapter, we will briefly discuss the characteristics of continuous manufacturing at the conceptual level, first, in its generic form, viewing the process as a unitary system, and then as a system composed of multiple manufacturing unit operations. Next, we will review the key requirements for implementing an effective continuous process and discuss some aspects that are specific to pharmaceutical applications. Lastly, we will conclude with a discussion of the advantages and limitations of continuous manufacturing and contrast those to the advantages and limitations of the batch operating mode, which has been the mainstay of the pharmaceutical industry.

1.2 General Characteristics of Continuous Processes

In this section, we will briefly review some of the basic concepts of relevance to the continuous mode of operating a pharmaceutical process. These concepts include: steady state, dynamic, and batch process; state of control; start-up and shutdown; nonconforming materials; batch/lot; residence time and residence time distribution; process time constant and gain.

Continuous manufacturing is a mode of operation in which the manufacturing system receives continuous inputs of mass and energy, transforms those inputs via a specific sequence of chemical and physical operations without interruption and produces continuous outputs of mass and energy. The ideal continuous plant is an open system that operates at *steady state*, that is, the input and output flows are constant over time and, thus, any accumulation of mass or energy in the system is likewise constant over time. By way of example, Fig. 1.1 depicts a continuous mixer in which components A and B are blended and the blend continuously withdrawn. To avoid overflow or depletion, the output flow must be chosen so as to keep the accumulation of material in the tank constant.

Unfortunately, in practice no real system is truly at steady state, rather input flows, external environmental conditions, and internal manufacturing parameters

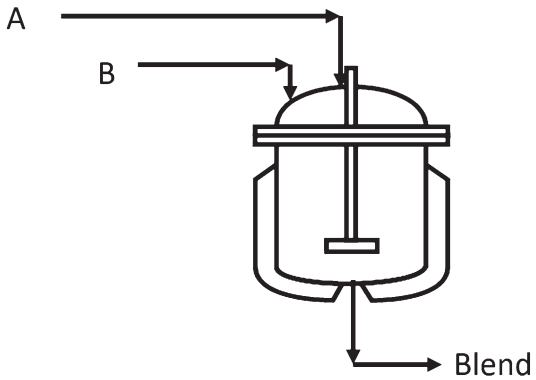


Fig. 1.1 Simple stirred tank mixing vessel

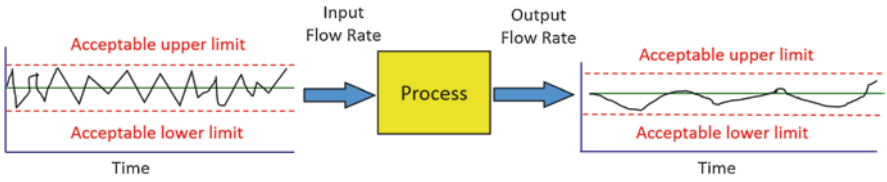


Fig. 1.2 Process in state of control

are continually subject to disturbances and thus fluctuate, in turn, causing deviations in the output flows, compositions, and possibly other properties. The continuous system is thus inherently in a dynamic state. If the disturbances are sufficiently small, then the process itself may dampen these disturbances sufficiently so that the deviations in the outputs are acceptable. In general, one cannot rely on the process to exhibit such stability and then the challenge is to equip the system with control strategies that will confine the fluctuations within limits such that product quality specifications are satisfied. If the process is operated so that all the important properties of the output streams can be confined within those acceptable limits, then the process is said to be in a *state of control* (CDER 2019). The process in Fig. 1.2 has disturbances in its input, but by virtue of a control logic the output is kept in state of control. That control logic may be passive (tank overflow outlet) or active (suitable manipulation of an outlet valve).

Any continuous process is thus inherently a dynamic process that is maintained in a state of control as a result of active intervention, typically by a suitably designed automation system. In general, a process can be in a state of control even if is operated in a cyclic or periodic fashion as long as the properties of the output are maintained within acceptable deviation limits around the nominal or "golden" periodic profile. With all continuous processes there are two situations in which large departures from a state of control can be expected: during start-up and shutdown. During *start-up*, a continuous process will be brought from an idle and empty state with no inputs and outputs to its nominal production rate, again in controlled fashion. The output material generated during start-up normally does not meet quality

specifications, is said to be *nonconforming*, and must be rejected. Likewise, during *shutdown*, the process is brought to a rest state with no input or outputs and in general at least a portion some of the output during that transition period will fail to meet quality specifications. Generally, the determination of safe and nonconforming material sparing start-up and shutdown strategies is an essential part of developing any continuous process control strategy.

By way of contrast, a pure batch operation involves charging the process unit with a specific amount of input, processing of that input while the system is closed and then removal of all of the output material at some point in time. During the period of time when processing occurs, the operation is in a dynamic state, with changing conditions within the unit. If the operation is carried out with fixed operating recipe and conditions, then any deviations in the input or any deviations from the operating recipe will translate to deviations in the output. In order to maintain quality specifications of the output, suitable changes in some of the recipe parameters must be undertaken. In practice, there are variations on the pure batch mode that occur. For instance, a batch operation can be *fed-batch*, that is, during the course of processing additional input is provided to the system, or *semi-batch*, that is, during processing some output component is removed over time, or both. A typical example of the former is a batch reactor in which one of the reactants is charged to the reactor and the other is fed to the reactor at some flow rate that may change over time so as to keep temperature rise within limits. An example of semi-batch is a batch centrifuge in which the cake is retained while the filtrate is removed as it is generated. One of the key differences between batch and continuous operations is that in the former, start and stop of the operation of the unit and the material transfers occur at discrete points in time, while in the latter, all inputs and outputs and processing occur continuously over time.

Another key difference between the batch and continuous modes is that in the batch case, the discrete amount of material produced inherently provides a convenient way of establishing an identity for the material produced. That discrete amount of material, called the *batch*, serves as a means of documenting and tracking material produced under FDA regulations. In the continuous case, the definition of *batch or lot* can be flexible and still satisfy the regulations. Thus, it can be defined in terms of a quantity of material processed, or based on a period of production time and can be flexible depending on the length of time, providing it is over a period of time during which conforming material was produced (CDER 2019).

By virtue of the fact that a batch operation is charged at a point in time, processing occurs over a defined period in time, and then all of the material discharged at a fixed point in time, all of the material in the batch spends the same amount of time in the unit. That length of time is called the *residence time*. When the process is continuous, it is not necessarily the case that all material flowing through the process will actually have the same residence time. In the ideal case of continuous flow of fluid through a pipe of uniform diameter, if the flow is ideal, that is, it involves no wall friction and no mixing in the axial direction, then the residence time of every element of fluid is the same and will be equal to the length of the pipe divided by the fluid velocity or L/u, as shown in Fig. 1.3.

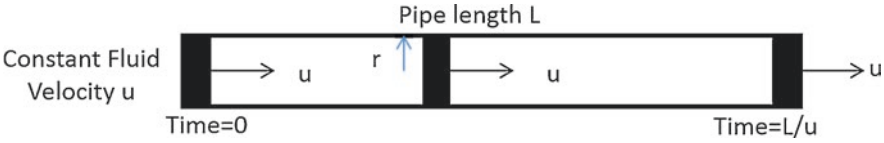


Fig. 1.3 Residence time for ideal flow in pipe

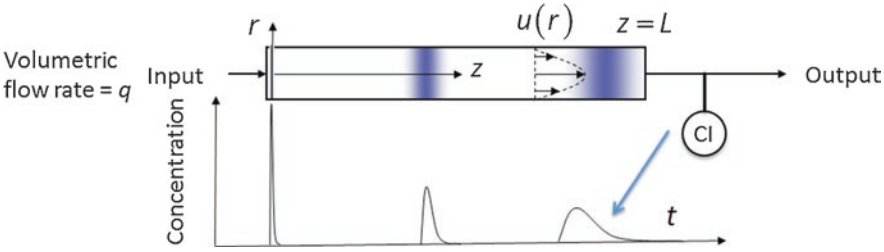


Fig. 1.4 Residence time in real flow in pipe

However, in the more realistic case in which mixing in the axial direction does occur and there is a wall friction effect creating a boundary layer, then the velocity will take on a parabolic shape with maximum value at the center line of the pipe. Some elements of the fluid (e.g., those along the pipe wall) will have a longer residence time in the pipe than others (e.g., those near the center line), as shown in Fig. 1.4.

This difference in residence time can be experimentally observed by conducting a tracer study that simply involves injecting at the inlet a small amount of a dye or other measurable component, which does not materially change the flow, and then measuring the concentration of dye at the exit over time. The *residence time distribution* (RTD) is simply a function of time, which indicates the fraction of fluid particles that experience a given residence time (Shinnar 1986). It can be computed from the dye concentration measurement normalized by the amount of dye injected, or

$$\mathrm{RTD} = E(t) = C(t) / \int_0^\infty C(t)$$

A very common unit operation is that of a stirred tank with continuous input and output. If the tank is perfectly mixed, then every fluid element entering the tank will be instantly mixed and have an equal probability of leaving the tank. It can be shown that the residence time distribution function, $E(t)$, for this ideal continuous operation is an exponential function parameterized by the *mean residence time,* which consists of the ratio of the volume of the vessel V divided by the (steady state) volumetric flow rate into the vessel, q, or $\theta = V/q$. Specifically, it is given by

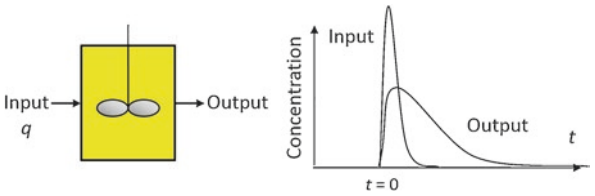


Fig. 1.5 RTD of continuous stirred tank

$$E(t) = (1/\theta)\exp(-t/\theta)$$

$$E(0) = 1/\theta$$

As shown in Fig. 1.5, in real stirred tank vessels there will be imperfections in the mixing performance, including a delay before an entering element of fluid enters the main mixing zone, bypassing of fluid elements to the output and stagnant zones where some fluid will be held back and thus the RTD will be distorted with some spread in shape. This will cause the mean residence time as well as the variance of the residence time to possibly be larger than the ideal. Moreover, in general the residence time distribution can also be affected by operating variables such as the impeller rpm.

The residence time distribution has implications in terms of the definition of a lot or batch since it makes the boundary between adjacent lots less distinct—the adjacent lots will share material with a similar history. Of course, this boundary effect as a fraction of the material constituting the entire lot diminishes with increasing lot size. The residence time also has implications with regard to tracking nonconforming material since if at the input to the unit a deviation in the material properties occurs that lies outside of the specifications, then that nonconforming material will appear at the outlet of the unit, delayed by the mean residence time. Thus, it is only from that time point on that the output material needs to be rejected. Moreover, when the input returns to be within specification, the rejection of output material can be stopped after a time equal to the average residence time has passed. Depending upon the degree of risk that is accepted by the organization, one may wish to be more conservative and define the beginning and end points of nonconforming material using the mean residence time adjusted by some multiple of the variance of the residence time.

While the residence time distribution gives a very valuable indication of how material will track through the process, it assumes that the process is at steady state or that it is operating in a state of control. It does not explicitly reflect how a disturbance in flow or composition of the input streams or a process parameter change within the process will be transmitted to the outputs. This dynamic effect can be captured through the use of a dynamic model of the process and can be empirically observed and quantified through step response experiments. In particular, the simplest form of dynamic response of a process is a first-order response that corresponds to a system in which the dynamic model of the system is linear. As shown in Fig. 1.6, the response of a linear system to a step change in one of its inputs is

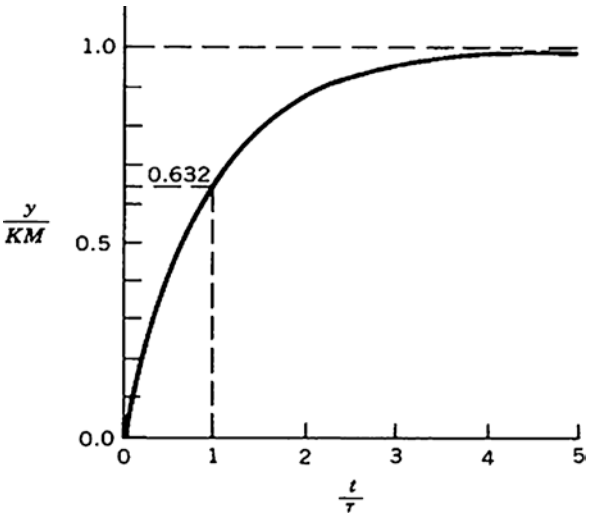


Fig. 1.6 Output response of first-order process to step change in input

represented by the classical exponential curve that is characterized by two parameters, the process *time constant* τ and the *process gain K*. As the linear dynamic process undergoes a step change in input of magnitude *M*, the dynamic lag of the process will cause a change in the output, which will only reach 0.632 of its final value in a period time equal to one time constant.

A typical process exhibiting first-order response is a well-mixed tank that is subjected to a change in composition or flow of one of its input streams and the exponential response is observed in the output. Moreover, the first-order model can be used as a reasonable approximation for any dynamic process undergoing small input changes. To capture the effects of larger deviations in inputs on a general process unit, the nonlinear effects will need to be suitably modeled and characterized. The dynamic characteristics of the continuous process serve as the basis for the design of active control strategies that can compensate for these characteristics in order to minimize the deviations in the output (Seborg et al. 2011).

1.3 Multiunit Continuous Processes

In general, a continuous process will consist of a sequence of unit operations linked by continuously flowing streams, for example, in the form of a fixed piping network. In this section, we will review some of the requirements and consequences of continuous operation of multiunit process trains. These include continuous material transfer between operations, the role of intermediate storage, limitations on process network structure, the impact on residence time distribution, and the implications of hybrid operations involving both batch and continuous subtrains. We will also note some comparisons to multi-operation batch campaigns.

First, the unit operations in a multiunit continuous process will span a wide range, including those typically used in production of small molecule active ingredients: continuous stirred tanks, continuous crystallizers, continuous filters and centrifuges, liquid–liquid extraction units, and distillation columns as well as those used in continuous production of dosage forms: loss in weight feeders, continuous powder blenders, roller compactors, mills, twin screw wet granulators, continuous dryers, and tablet presses. A key characteristic of continuous processing is that the transfer of materials between unit operations occurs continuously, without interruption. *Continuous material transfer* between units may be driven by various well-known means: gravity or pressure differences, pneumatic transport, pumps, or compressors. However, regardless of the driver selected, the requirement that the continuous flow is maintained can introduce challenges when the flow being transferred is a viscous fluid, a particle blend or a suspension. The properties of the material must be engineered to have reliable rheology and the flow of the stream must be monitored to assure that flow is consistent—in a state of control.

Another important characteristic of continuous processing is that generally holding or *intermediate storage* of material between unit operations is undesirable for multiple reasons.

Holding between process units raises the undesirable possibility of creating nonuniformities in the material being held—settling is the obvious example phenomena. Secondly, holding has an impact on the overall time constant of the process and thus causes responses to plant-level process control actions to be slower, which means corrections to deviations may be delayed. Moreover, holding also adds to the variance in the process residence time (as further elaborated later in this subsection), thus, increasing the boundaries defining successive lots. Finally, holding creates delays in start-up and shutdown as the contents of hold tanks have to be filled/emptied. The one positive aspect of intermediate storage is that it will dampen the fluctuations or surges in flows that are inputs to the hold tank and thus reduce the deviations that the unit downstream of the hold tank has to accommodate. This can be helpful to the control system since the reduced magnitude of flow deviations will reduce the frequency or need for control action.

The network of connected continuous processing units may in general take the form of a network that involves a variety of processing paths, including both bypassing of some unit operations as well as feedback loops or recycles from downstream units to upstream, as shown in Fig. 1.7.

In Fig. 1.7, the streams into the Interpass Tower Pump Tank are examples of bypass streams, while the split of the outlet from that unit that is returned to the dry tower pump tank is a recycle stream. Of course, there are multiple other streams in the flow sheet as well. By virtue of the need to be able to track lots through the manufacturing process, in pharmaceutical applications of continuous processing, *bypass* and *recycle* streams generally have to be avoided. Both structural forms result in the mixing of materials that have seen a different processing history and, therefore, violate the requirement for material traceability. Hence, continuous manufacture in the pharmaceutical domain generally has to follow a *serial flow sheet* structure.

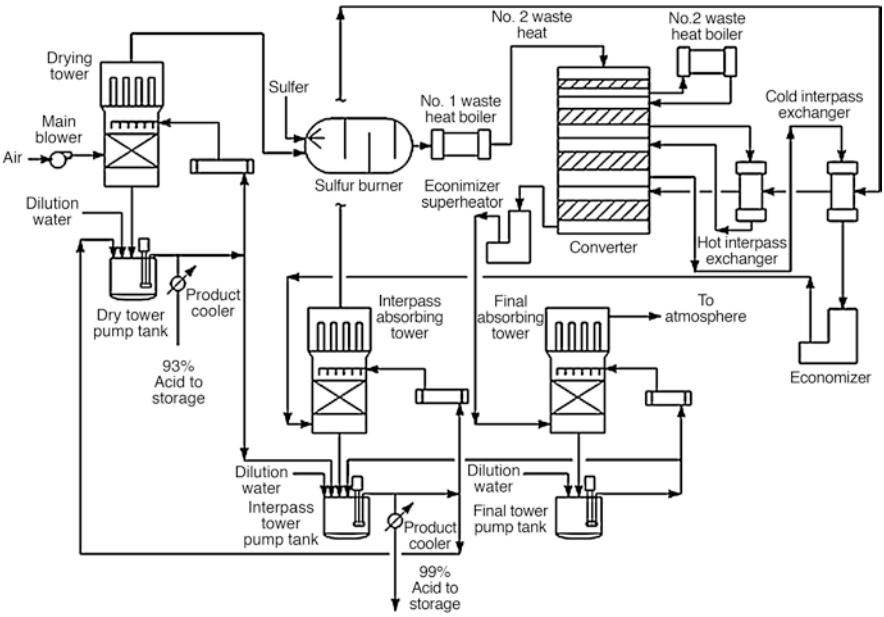


Fig. 1.7 Sulfuric acid process with bypass and recycle streams (Muller 2006)

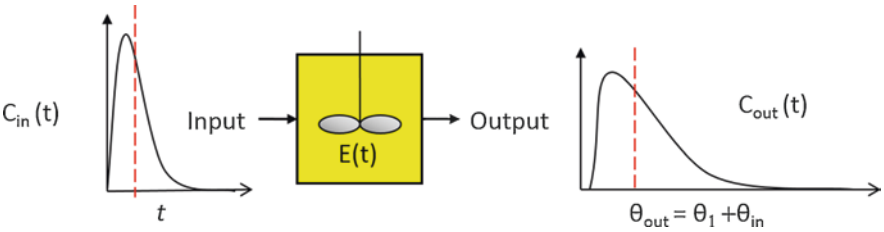


Fig. 1.8 Convolution of input and RTD of unit

An important implication of having multiple connected continuous unit operations is that the mean residence time of the sequence of units is the sum of those of the individual units and the variances are additive as well. Specifically, it can be shown that given a residence time distribution for a unit, $E(t)$, if that unit is subjected to a concentration wave represented by $C_{in}(t)$ then the output concentration wave will be given by the convolution integral

$$C_{out}(t) = \int_0^t C_{in}(t-\tau) E(\tau) d\tau$$

This mathematical operation can be represented graphically as shown in Fig. 1.8, where θ_1 is the mean residence time of the unit operation with RTD $E(t)$.

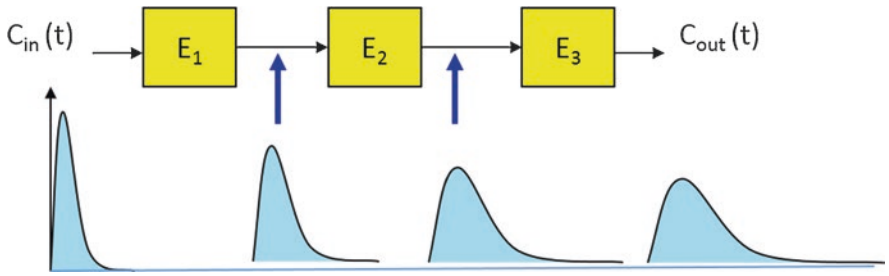


Fig. 1.9 Effect on RTD's of sequence of units

By successive application of the convolution integral it can readily be confirmed that for a sequence of unit operations, each with its RTD as shown in Fig. 1.9, the cumulative effect is additive in mean and variance. For example, for three units in series, mean and variance are as follows:

$$\theta_{\text{out}} = \theta_1 + \theta_2 + \theta_3 + \theta_{\text{in}} \qquad \sigma^2{}_{\text{out}} = \sigma^2{}_1 + \sigma^2{}_2 + \sigma^2{}_3 + \sigma^2{}_{\text{in}}$$

The net of effect of a series of unit operations is a broadening of the residence time distribution and thus in the case of nonconforming material, which arises upstream, the nonconforming material will be distributed across a larger portion of conforming materials. The consequently is that the total amount of material that must be rejected to meet quality specifications will be increased. The clear implication is that it is desirable to divert nonconforming material as close to its first observation as possible. It should be noted that in the above analysis the delay time due to transfer of material between unit operations is neglected. Of course, the effects of material transfer operations can be readily incorporated by simply treating the transfer as a unit operation with its own RTD and mean residence time. Such corrections may well be appropriate for operations such as pneumatic transport where significant back mixing may occur.

By contrast, for a series of batch operations, the total time of the material processed will again be the sum of the residence times in each operation. Of course, that total residence time must be increased by the transfer times between operations (or even more so by any quality control delays), which in general can be quite significant. However, since the material is transferred in discrete amounts between unit operations, any requirement for segregation of nonconforming material is confined to the amount corresponding to the batch size. However, while the broadening of RTD is not an issue for batch operations, rejection of nonconforming materials necessarily requires rejection of the entire batch. In the continuous case, rejection only applies to the portion of the material that is actually observed to be nonconforming, corrected for RTD effects noted above.

An additional contrasting feature of batch operations is that since normally manufacturing operations take place in campaigns, the hold times, quality control (QC) times or other delays between unit operations are not cumulative. Rather, since the

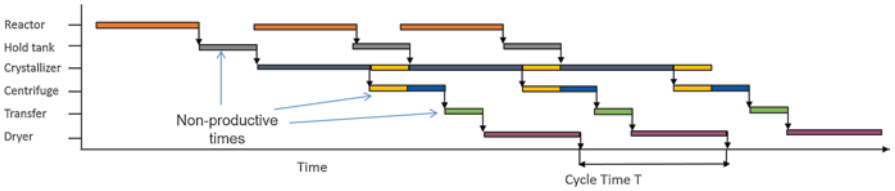


Fig. 1.10 Batch Campaign Cycle Time

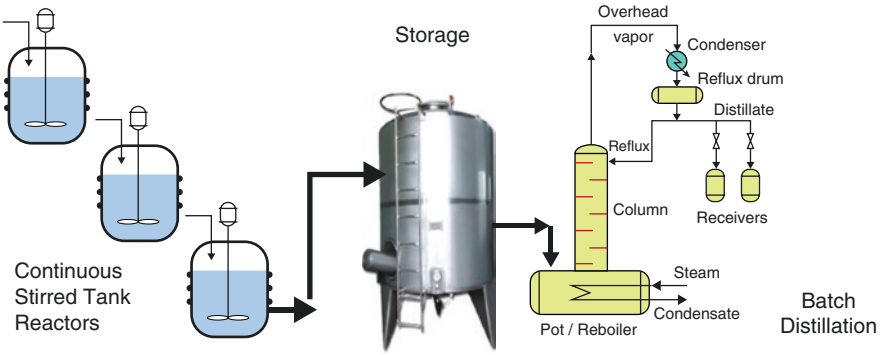


Fig. 1.11 Hybrid Operation

batch campaign is characterized by batch cycle time, it is the largest combined batch processing time at any unit operations that matters. This is illustrated in Fig. 1.10, which represents a lean operation in which operations are executed just in time. As shown, material is held in a drop tank after reaction for QC check; material is transferred from crystallizer to centrifuge over time and the cake from the centrifuge is transferred into an intermediate hold tank for transport and loading of the dryer.

In some cases, it can prove advantageous to create manufacturing lines in which some subtrains consist of batch operations and other subtrains consist of continuous operations that, in general, will need to operate in a semi-continuous fashion. Integrated operation of such mixed, or *hybrid* production lines, requires the use of intermediate buffer storage. In general, the batch size of the batch subtrain will impose the batch size on the continuous subtrain, although the material produced in the continuous subtrain could have further subdivision into lots. Moreover, to retain batch identify, the intermediate storage should not mix multiple batches; rather, batches must be stored individually. It is of course possible to operate such hybrid lines in a fully integrated "lean" fashion in which the production rate of the continuous subtrain is matched with the average production rate of the batch subtrain. While such operation will minimize the time interval that intermediate material will need to be held in storage, the tight integration will require use of predictive scheduling of these operations. A simple example of such a hybrid operation is illustrated in Fig. 1.11.

The conventional direct compression tablet production process, which consists of batch blending, manual transport of bins of the powder blend to the tablet press, tablet production, collection of tablets in a bin and finally tablet coating, is in fact a hybrid production line. Typically, blending and coating are batch operations while tableting is continuous. The batch size is typically defined by the blending stage.

1.4 Requirements for Effective Continuous Processing

The functions that are essential to carrying out continuous pharmaceutical manufacturing are process monitoring, deviation management, and materials tracking. These functions in turn are generally supported and executed using process analytical technologies, process control and intelligent alarm management systems, as well as process data and knowledge management systems. At a more advanced level, these basic functions can be further augmented with real-time release, real-time optimization capabilities, and operations management systems. In this section, we will briefly review these technologies at a conceptual level.

The most rudimentary capability that is required to operate a continuous manufacturing line is *process monitoring*, that is, the continuous assessment of the state of the line to determine whether or not it is in a state of control. This capability is essential because all real processes are continually subjected to disturbances, which cause deviations from desired product attributes and process variables and conditions, and clearly such deviations must be managed to assure product quality. In general, these disturbances can be of a random nature (often called common cause variations) or they can be disturbances that have a nonrandom component (often called special cause variations). An example of the former is deviations in the API composition of a powder blend, which arise due to the particulate nature of such blends. An example of the latter is fouling of an optical sensor due to the accumulation of process materials on the sensor lens. In Fig. 1.12, the process is undergoing random variations until the end of the time window, at which point the process variable, api content, undergoes a significant variation much larger than the variance of those observed previously.

Process Analytical Technology consists of in-line, at-line, and off-line measurement systems, which are used to provide the data of process monitoring. It includes actual sensors and their calibrations, as well as virtual sensors, models that can

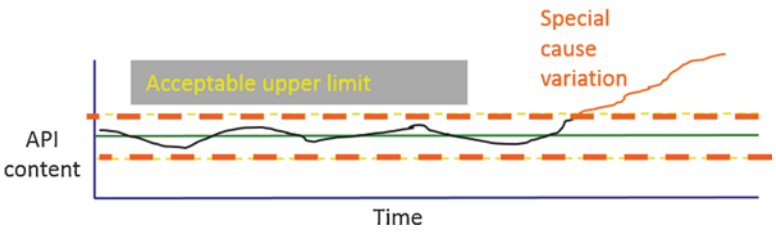


Fig. 1.12 Common and special cause variations

predict values of unmeasured or unmeasurable material attribute or process variables from process parameters or variables that can be measured. Examples of in-line sensors include those based on the use of NIR, Raman, ultrasound, X-ray, microwave and laser light scattering signals. At-line sensors normally involve technologies that require a longer period of time to arrive at a measurement relative to the dynamics of the process. For example, the time to conduct a 3D scan of a tablet to determine api distribution in the tablet is generally much longer than the time to produce a tablet in the press. Typically, the manufacturing line will be instrumented with a network of sensors, sufficient so that there is some degree of redundancy if and when sensors fail. An example of such a network is shown in Fig. 1.13 for a continuous dry granulation line (Ganesh et al. 2018).

The most rudimentary use of process monitoring capability is to detect deviations that require some form of intervention to assure product quality. Monitoring methods include classical univariate statistical quality control methods, multivariate statistical process control methods, and data reconciliation (DR) and gross error detection methods. The classical methods serve to track individual attributes or variables and to subject the time series of measurements to statistical tests to establish whether or not actionable deviations have occurred. MSPC methods are used to efficiently track multiple variables simultaneously and are particularly relevant when those variables have significant correlation. DR methods use all of the real-time measurement data, along with prior information about the statistically characterized error in those measurements and a model of the process to predict the most likely state of the process. Appropriate statistical tests are also used to detect gross errors, that is, identify measurement data that are outliers and thus indicators that a special cause variation has occurred (Moreno et al. 2018).

Common cause variations can lead to quality deviations if not corrected in real time. Generally, the correction is provided through suitably designed and tuned *process control systems* that employ feedback, feedforward, or multivariable control strategies that are based on the use of predictive process models. An example of a classical single-input/single-output feedback control structure is shown in Fig. 1.14. The difference or error between the actual measured value of a controlled variable and the target or *set point* of that variable is used in the control system logic to determine a correction or control action that will change the value of a suitable manipulated variable. The simplest possible logic might be to mulitply the error by a proportionality constant to determine the magnitude by which the manipulated variable should be adjusted. However, depending on the dynamics of the process and the performance required, more complex controller designs may need to be employed (Seborg et al. 2011).

Moreover, the control system may be structured at multiple levels, for instance, at the unit operation level and at the plant-wide level (Su et al. 2019). Unit level control serves to maintain the operation of an individual unit operation at a desired set point, for example, the control of density of the ribbon produced by a roller compactor by manipulation of roll pressure. Plant-wide level control serves to control important plant operating variables, for example, the production rate of the plant at desired set point by manipulating the flow rates of input materials to the plant. Process control systems are normally implemented to function automatically,

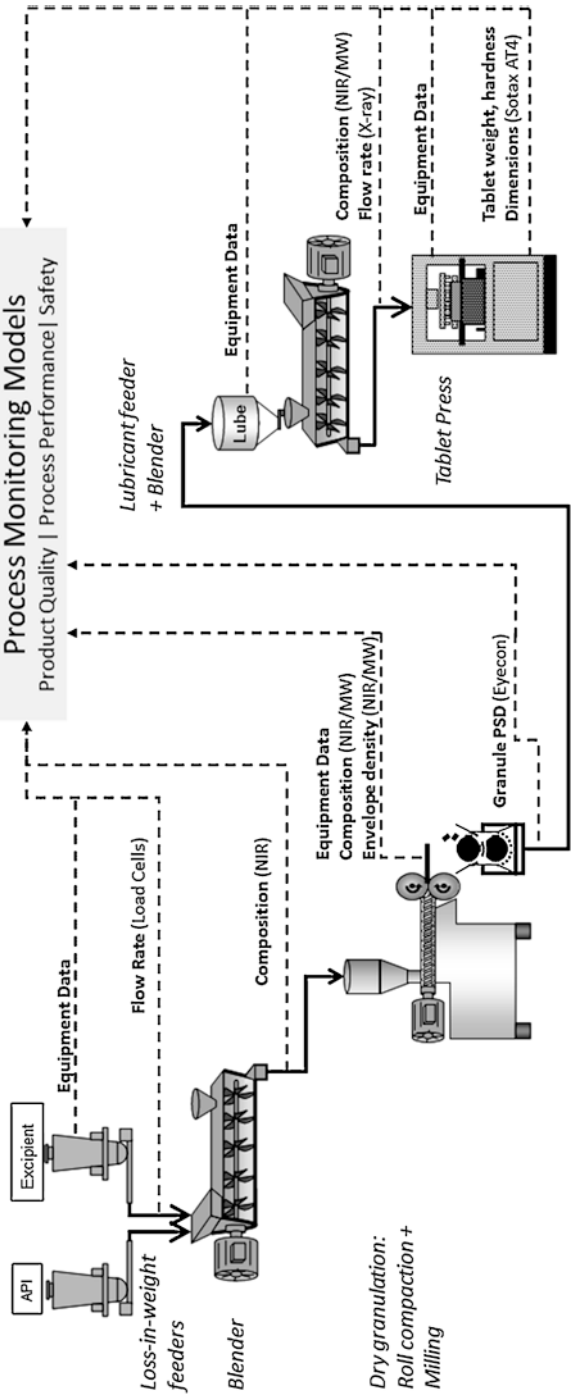


Fig. 1.13 Sensor network for continuous DG line (Ganesh et al. 2018)

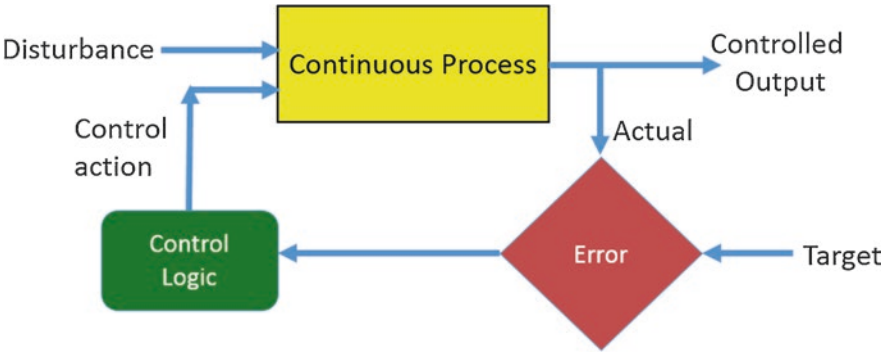


Fig. 1.14 Single-input/single- output feedback control

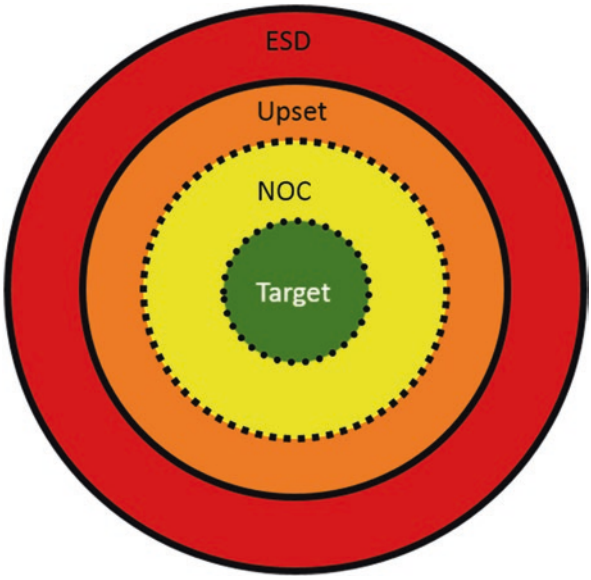


Fig. 1.15 Process condition model

without operator intervention. However, disturbances, which are nonrandom, cannot, in general, be handled by conventional process control systems. Such disturbances or faults could be caused by sensor or equipment degradation over time or outright failure of sensors, equipment, control system, or communication system. In general, process faults such as these require active intervention by the operator, although some degree of automation is possible. In order for the operator to be effective in intervening and returning the process to normal operation, it is important to provide the operator with as much help as possible in identifying or diagnosing the likely cause of the fault as well as guidance on actions to take to mitigate or correct the fault so as to avoid process shutdown.

In the continuous industries, the *process condition model*, represented in Fig. 1.15 for a typical process variable, is used to capture the way in which *intelligent*

operator advisory systems should be structured. In this model, the target operating condition constitutes the optimal operating condition, while the normal operating condition (NOC) constitutes acceptable operation. The process is in a state of control if it is in either the target or NOC region. The Upset condition indicates that the process is not under control and, thus, it is highly likely that the product being produced does not meet specifications. Finally, the Emergency Shut Down condition indicates that the process is operating in an unsafe region, with the potential of equipment damage, and, of course, producing a nonconforming product.

The International Society for Automation has issued models and guidance on the design of intelligent operator advisory systems (ISA-18.2, 2009). Ideally, such systems should have some form of diagnosis methodology imbedded that helps the operator determine what the likely cause of the fault is. There exists a large body of fault-diagnosis methods that provide the bridge between fault detection and fault diagnosis, that is, the determination of the cause of the observed disturbance (Venkatasubramanian et al. 2003). However, this methodology only applies to faults that have been previously observed so that both the sensor signature of the fault and the mitigation action are known and have been recorded. Given such a fault library, faults can be diagnosed by looking for a match of the observed sensor signature against those of the library of previously encountered faults. Of course, when a previously unobserved fault is encountered such advisory is not possible. In general, the integration of process control and intelligent alarm management constitutes the core of the deviation management capabilities of a continuous manufacturing line.

While deviation management is focused on keeping the process in a state of control, there remains the issue of how to deal with the consequences of these deviations, namely, the nonconforming material generated during these events. This material will be produced during the course of realization and mitigation of process faults as well as during start-up and shutdown procedures. A reliable system for tracking the nonconforming materials and segregating these materials so as not to contaminate materials that do meet critical quality attributes is an important element of the pharmaceutical quality system of a continuous process (Lee et al. 2015). Once the statistical testing methodology of the PAT system detects a fault, then that material must be traced as it progresses through the unit operations downstream of the point at which the deviation occurs. The time at which the process is returned to a state of control marks the end of nonconforming material production and thus the material associated with this transition time must also be tracked. The tracking logic must take into account the RTD of that process subtrain, as well as the process dynamics associated with the transition from the conforming to the nonconforming state and with the return of the process to state of control. Appropriate in-line sensor data can also be integrated into the tracking strategy to confirm the tracking predictions, which can be based on the use of dynamic models or on estimates based on residence time parameters and process time constants. The tracking will continue downstream until a selected point is reached at which the nonconforming flow can be diverted and removed from the process. The simplest approach is to carry out the diversion at the end of the manufacturing line, for example, to divert the tablets produced from the nonconforming materials to a reject bin. However, given the cumulative effects of the residence times of successive unit operations, diversion as

close as possible to the point of location of the first detection is desirable as it will minimize the amount of material that must be rejected. However, this may require shutdown or slow-down of the portion of the line downstream of the diversion point, a potentially complex process that must take into consideration the associated dynamics and control issues. While the tracking and segregation procedures do introduce complexities, the benefit is that the amount of material that is rejected will be much smaller than the rejection of an entire batch.

The process monitoring, deviation management, and materials tracking functions described above depend critically on data, models, and knowledge of considerable extent and diversity. Moreover, they rely on platforms for systems integration and implementation. The data sources include in-line and at-line process sensors for capturing in real time the properties of the streams in the process, equipment status, and process operating data obtained in real time from instrumentation integrated within the equipment, data from at-line sensors and off-line primary test methods and laboratory data from supporting laboratories. Much of this data is recorded and stored in the process historian associated with the Digital Control System that manages the process control functions. Other information is retained in complementary data repositories maintained by supervisory control systems. The deviation management systems may be implemented in a software system for configuring and executing multiple operations management decisions, which include nonconforming material tracking as well as systems for predictive condition-based maintenance of all of the manufacturing resources: the process instrumentation, equipment resources, the control systems, and the models supporting these functions. Comprehensive systems for integrating all of these functionalities are integral to the current wave of Industry 4.0 developments sweeping the manufacturing industries. Moreover, the decades of learnings in automation accumulated in the chemical processing industries have been captured in multiple ISA standards, as shown in Fig. 1.16. By following these standards, pharmaceutical manufacturing can leapfrog

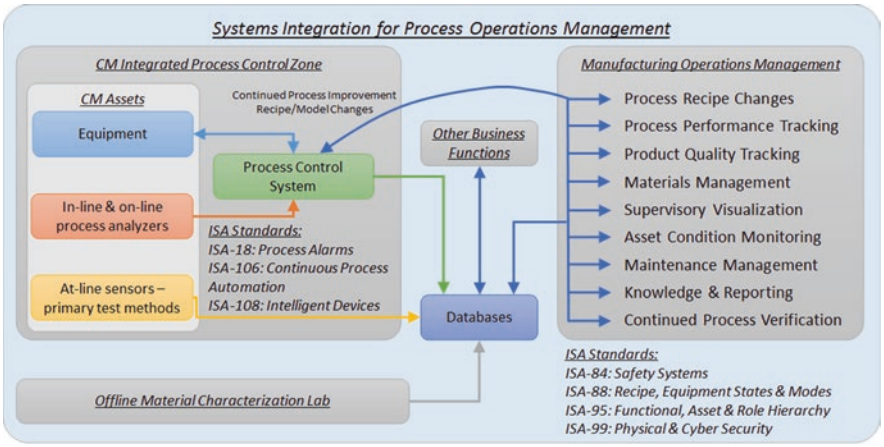


Fig. 1.16 Systems integration functions and standards

the old generation of process data management tools and structures with state-of-the-art implementation using new tools and architectures.

Whether at the unit operations level or at the plant-wide level, the key decision variables of the process control system of the continuous production line are the controller set points. As the operation of the process is impacted by changes, such as those in the properties of feed materials, a gradual decline in equipment performance, the occurrence of abnormal events, or production of noncompliant materials, the current set points may need to be changed. The new set points should of course be consistent and lead to feasible operation and, if a range of choices are feasible, to be optimal with respect to some appropriate performance criterion. The new set points may only be used for a short period or an extended period of time, depending upon the cause for the change. In the continuous processing domain, the new set of set points can be determined by solving a *real-time optimization* problem (RTO). Typically, the RTO problem is posed using a mathematical model of the manufacturing line, either steady state or dynamic, depending upon the situation, and a suitable objective or performance function, such as minimization of production of noncompliant materials production. If the expected changes in set points are relatively small—perhaps due to gradual changes in unit operation performance, a steady-state model may suffice, as the RTO will be carried out as needed to maintain the process at an efficient state. However, if the changes are large, such as may be the case when a step change in the desired plant production rate is to be initiated, then a dynamic model may be appropriate. In general, because the RTO problem has to be solved in a time period commensurate with the dynamics of the process, the process model will need to consist of reduced order models of the set of unit operations. One of the typical applications of RTO is to determine the optimal sequence of steps or trajectory that should be followed in starting-up and shutting down a continuous production line. Typically, the determination of start-up and shutdown trajectory is part of the overall control strategy of a continuous manufacturing line (Ierapetritou et al. 2016).

The release of a manufactured lot for distribution to the market constitutes an important decision point, which is designed to assure the customer (via the FDA) that the quality of the drug product is acceptably good and the producer that the manufacturing process is operating with acceptable process capability. It requires the execution of quality control tests to determine that CQAs are met within acceptable limits. In the case of tablets, the tests include standardized measurements, such as, api content, weight, hardness, physical dimensions, and dissolution behavior carried out on a specified number of tablets. Several of the CQAs essential to the release decision are monitored and controlled in the continuous process. However, testing for other CQAs requires longer duration or is destructive and thus either completed batches must be held for QC action or equivalent tests implemented, which would allow batches to be released immediately upon leaving the manufacturing line. *Real-Time Release Testing* is the term adopted to reflect that the objective is to implement the verification of the CQA as part of real-time monitoring (CDER 2019). One of the ways to implement RTRT of a CQA of the finished drug product that is not measurable in real time is to use a soft sensor or surrogate model

that relates the CQA of interest to other directly measured properties. In general, the model will take the form of a semiempirical relation or other reduced order model. The soft sensor employed in RTRT has to be systematically validated against reference measurement of the CQA of interest using at-line measurements or off-line tests. Although at present, lot release of continuously manufactured products follows the traditional off-line QC methodology, the development and evaluation of alternative RTRT methodology for continuous pharmaceutical products is the subject of ongoing research. A related issue is development of release criteria that take into account statistically sound measures of process capabilities based on previous batches of that product produced on that line along with real-time data from the current batch to arrive at estimates of the probability that all dosages of the batch to be released meet CQA requirements.

1.5 Comparative Assessment of Batch and Continuous Operating Modes

Traditional batch processing has served the industry for many years because of the flexibility it offers by virtue of using well-known multipurpose equipment that can accommodate a range of products and product recipes. A batch recipe can be developed and implemented empirically supported by design of experiments with a relatively limited detailed understanding of the chemical and physical phenomena. Once a recipe is found to be satisfactory, it can be repeated to produce batch after batch, using learning by doing, again with limited detailed fundamental understanding. Since processing occurs in discrete amounts of material—the batch size- material can be tracked readily from unit operation to unit operation to finished product and batches released based on well-established procedures. Moreover, traditional batch unit operations lend themselves to processing steps that require long residence times and involve multiphase phenomena.

However, batch processing has significant disadvantages. By their very nature, batch processes are not only dynamic but also cover a wide range of dynamic conditions from the start of the operation to its termination. The batch of material produced is the result of this entire range of conditions and thus the end state is very much path dependent. This makes the result of the operation very much affected by process, environmental and human factors. The human contribution to batch variability includes variation in the initial amount of material charged and variations in the timing of particular actions, including termination of the batch. Furthermore, by virtue of the discrete nature of the processing with starts and stops to fill, empty and transfer materials, equipment does spend quite a bit of time in nonproductive use. This translates to poor utilization of capital resources. Moreover, the scale-up from small batch size bench equipment to pilot plant size and eventually to manufacturing-scale equipment can be problematic, often requiring modification of the recipe to ensure that the same critical quality attribute targets are met. Similar scale-up challenges can arise when transferring recipes from one manufacturing facility to another.

By contrast, continuous manufacturing by its very nature offers high equipment utilization due to the elimination of idle times between operations as well as the time associated with transferring material between successive operations. High equipment utilization translates to smaller equipment capacity or footprint to achieve the same output per unit time. The smaller equipment size and possibility of increasing production by just running longer significantly reduces concerns with scale-up. The elimination of the need to transfer and hold material between operations also reduces the work-in-process inventory and also reduces material losses. Another important advantage is reduced product variability as a result of less material handling and other intervention by humans. While continuous manufacturing does require use of on-line measurement and process control, the on-line instruments see a more limited range of conditions—essentially only variations around nominal operating conditions, thus both required instrument range is reduced and calibration is simplified. Process control likewise involves control around a set point rather than tracking and controlling over an entire dynamic profile. Finally, for specific operations such as exothermic reactions, flow reactors provide much better heat transfer surface per unit volume ratios and thus much better temperature control and as a result safer operation. Flow through continuous units also generally provides improved micro-mixing.

However, continuous operations do require more detailed process understanding, especially of rates of reactions and transport processes in order to properly design equipment. Of course, continuous operation requires real-time measurement and control systems and, as we saw in the previous sections, the methodology for deviation management and material tracking that add complexity. In addition, because of longer operating runs, continuously operated equipment can encounter gradual fouling or degradation of performance over time, which may require special real-time strategies to avoid the need for unplanned shutdown. In general, continuous processes have less flexibility in accommodating different products because different rate phenomena directly impact equipment design capacity. However, in recent years there have been significant efforts to develop small-scale modularized designs of continuous processing equipment that can facilitate flexible assembly of continuous lines from an inventory of standardized components. While the limitations of dealing with long residence times and multiphase systems remain, significant progress is being made to address these limitations.

Given the relative strengths and weaknesses of these two operating modes, it can be expected that they will continue to not only coexist within the industry and even in the same manufacturing organization but will be combined to create hybrid configurations of process subtrains that effectively exploit their relative strengths. However, there remains much to be done to increase the penetration of continuous manufacturing into the pharmaceutical sector (Ierapetritou et al. 2016). This book aims to promote and accelerate this trend.

References

CDER US FDA. Quality considerations for continuous manufacturing guidance for industry (draft guidance). February, 2019.; Available from: https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.html.

Ganesh S, Moreno M, Liu J, Gonzalez M, Nagy Z, Reklaitis GV. Sensor network for continuous tablet manufacturing. Comput Aided Chem Eng. 2018;44:2149–54.

Ierapetritou M, Muzzio F, Reklaitis GV. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AICHE J. 2016;62:1846–62.

ISA-18.2. Management of alarm systems for the process industries. Research Triangle Park: available form ISA; 2009.

Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10(3):191–9.

Moreno M, Liu J, Su Q, Leach C, Giridhar A, Yazdanpanah N, O'Connor T, Nagy ZK, Reklaitis GV. Steady-state data reconciliation of a direct continuous tableting line. J Pharm Innov. (on line 9/2018., https://doi.org/10.1007/s12247-018-9354-9).

Muller TL. Sulfuric acid and sulfur trioxide. In: Kirk-Othmer encyclopedia of chemical technology. Hoboken: Wiley; 2006.

Seborg DE, Edgar TF, Mellichamp DA, Doyle FJ. Process dynamics and control. 3rd ed: Wiley; 2011. Chapt 2. Hoboken, NJ

Shinnar R. Use of residence- and contact-time distributions. In: Reactor design in chemical reactors and reaction engineering: Marcel-Dekker; 1986. Hoboken, NJ. p. 63–150.

Su Q, Ganesh S, Moreno M, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. Comput Chem Eng. 2019;125:216–31.

Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. A review of process fault detection and diagnosis part I: quantitative model-based methods. Comput Chem Eng. 2003;27:293–311.

Chapter 2
Continuous Reactors for Pharmaceutical Manufacturing

Martin D. Johnson, Scott A. May, Michael E. Kopach, Jennifer Mc Clary Groh, Timothy Donald White, Kevin P. Cole, Timothy Braden, Luke P. Webster, and Vaidyaraman Shankarraman

Abstract A variety of high-pressure and low-pressure plug flow reactors (PFRs) are described in this chapter with manufacturing examples for each. Coiled tube PFRs and vertical pipes-in-series PFRs were used for two-phase gas–liquid reactions. A pulsating flow coiled tube PFR was used for gas–liquid reaction with solids precipitate. Superheated PFRs were used for reactions involving homogeneous solutions, heated above the boiling point of the solvent. Disposable coiled tube PFRs were used with highly potent compounds. Continuous stirred tank reactors (CSTRs) were needed for heterogeneous continuous processes. CSTRs, CSTRs-in-series, and intermittent flow CSTRs were used for reactions with long reaction times, positive order kinetics, and multiple reaction phases, either solid/liquid or liquid/liquid. This chapter also explains how to calculate the actual internal temperature profile along the length of a PFR, which is often not practical to measure.

Keywords Plug flow reactor · Continuous stirred tank reactor

Abbreviations

A/V	Surface area to volume ratio
API	Active pharmaceutical ingredient
CSTR	Continuous stirred tank reactor
dtbpf	1,1′-Bis(di-tert-butylphosphino)ferrocene
EE	Ethoxy ethyl
i.d.	Inside diameter
PFR	Plug flow reactor

M. D. Johnson (✉) · S. A. May · M. E. Kopach · J. M. C. Groh · T. D. White · K. P. Cole
T. Braden · L. P. Webster · V. Shankarraman
Eli Lilly and Company, Indianapolis, IN, USA
e-mail: johnson_martin_d@lilly.com; may_scott_a@lilly.com; kopach_michael@lilly.com; groh_jennifer_mcclary@lilly.com; white_timothy_donald@lilly.com; k_cole@lilly.com; braden_timothy@lilly.com; webster_luke_p@lilly.com; shankarraman_vaidyaraman@lilly.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_2

RTD	Residence time distribution
S/C	Substrate to catalyst ratio
τ	Mean residence time

2.1 Introduction

The purpose of this chapter is to describe different types of continuous flow and intermittent flow reactors that have been used in pharmaceutical production. It explains what types of reactions were accomplished with each type of reactor, the physical phases that were acceptable, and ranges for temperature, pressure, reaction time, and reactor volume. Seven of the examples show reactors that operated in laboratory fume hoods at 5–15 kg/day production rate, as part of multistep continuous process steps where the downstream separations were also done continuous in the lab hoods. The reason why the 5–15 kg/day production rate by fully continuous process is important is because many of the potential medicines in the pipeline are projected to require less than 1500 kg/year active pharmaceutical ingredient (API). Therefore, a portion of the portfolio may be delivered in a facility designed for fully continuous processing in laboratory fume hoods. The reactors were designed with practicality in mind. Heat and mass transfer rates were sufficient for the chemistry, but not extremely high like the heat and mass transfer rates that can be obtained in microreactors. The reactors shown in this chapter were designed to meet the needs of the chemistry in a cost-effective and practical way with minimal overdesign of heat and mass transfer characteristics. The benefits of inexpensive, dedicated, or consumable reactors are explained. The benefit of intermittent flow rather than truly continuous flow is explained for a Suzuki coupling reaction. Mean liquid residence time in the continuous reactors ranged from 1.5 to 24 h, and most of the reactions required longer than 3 h residence time in the continuous reactor. It is commonly assumed that only fast reactions are suitable candidates for continuous processing, but it is not the case, as demonstrated by examples in this chapter. Five of the reaction examples described in this chapter were run in PFRs, and three were run in CSTRs or CSTRs in series. Derivation of reactor design equations for PFRs, CSTRs, and CSTRs in series, calculation of required mean residence times (τ) for the different reactor types, comparison of residence time distributions (RTDs) for CSTRs in series versus PFRs with dispersion, modeling and quantifying axial dispersion number (*D/uL*), and numerical modeling of conversion as a function of *D/uL* are given in a different book chapter (Johnson et al. 2019).

2.2 Superheated PFR in Oven for Imidazole Cyclization

Plug flow reactors operating with homogeneous liquid solutions can be scaled down to very small volumes, for example, 0.1 mL, and scaled up to large volumes in manufacturing. Research scale PFRs have been increasingly used as laboratory tools by modern synthetic chemists due to improved control of reaction parameters and

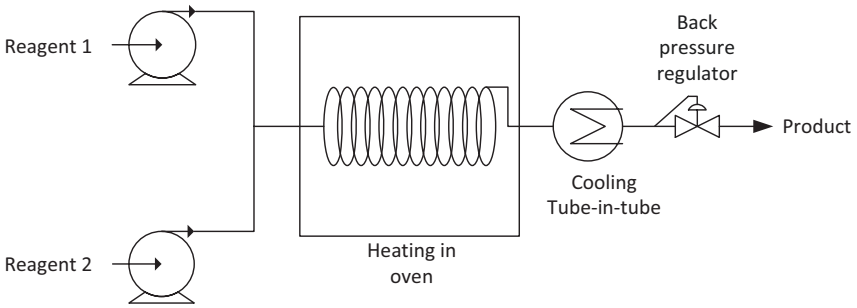
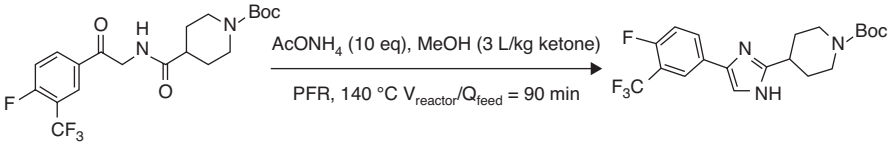


Fig. 2.1 Schematic of a thermal PFR system

the wide range of operating conditions that are possible (Wiles and Watts 2008; Lange et al. 2011). Successful application and scale-up requires control of mass flow rates and understanding the physical properties of the fluids including density, thermal expansion, heat capacity, reaction kinetics, thermodynamics, heat transfer rates, and RTD. A simplified schematic of a thermal PFR system is shown in Fig. 2.1.

Reagent solutions continuously flow from high-pressure pumps into the tube reactor. If there is more than one reagent solution, they are combined with some type of in-line static mixing. The static mixing could be as simple as a Tee-mixer, or as sophisticated as micro-structured split-recombine or multi-lamination device (Schwolow et al. 2012). The choice for mixing type depends on reaction kinetics and the mixing sensitivity of the reaction. Product solution flows out of the PFR and through a cooling heat exchanger and a back pressure regulation device. A 7.1 L high-pressure, high-temperature thermal tube PFR was used for a thermal cyclization for GMP production of 29 kg of an advanced intermediate as shown in the scheme (May et al. 2012).



The reaction was run with 90 min *V/Q* (*V* is reactor volume and *Q* is volumetric flow rate in the feed pumps at ambient temperature) using methanol as the solvent at 140 °C reaction temperature and 69 bar reaction pressure. These temperatures and pressures were beyond the limits of batch vessels in the pilot plant, therefore batch reaction was not an option in the chosen facility. Furthermore, the batch reaction did not scale up well because of conversion to undesired by-products during the slow heat-up to reaction temperature, but the continuous PFR reaction scaled-up with unchanged yield and selectivity because of consistently fast heat-up time at each scale. Heat-up from 20 to 140 °C in less than 2 min at the PFR inlet and cool-down from 140 to 20 °C in less than 2 min at the PFR outlet are easily accomplished in a continuous reactor but not possible in a pilot plant batch reactor. Faster heat-up

Fig. 2.2 Picture of a 7.1 L coiled tube thermal PFR (left) and forced convection oven for heating the reactor inside (right)

and cool-down times could have been achieved in continuous microreactors, but that was not necessary for this ~90 min reaction.

A picture of the 7.1 L thermal tube PFR is shown in Fig. 2.2, along with the forced-convection oven used to heat it. The PFR was made from 151 m long, 7.75 mm inside diameter (*i.d.*) coiled stainless steel tubing. The tubing was formed into four concentric cylindrical coils 0.43 m tall and ranging from 0.28 to 0.43 m diameter, connected by 7.75 mm *i.d.* jumpers from the top of one coil to the bottom of the next. It was designed so that most of the volume was in the uphill flow direction. A ketoamide solution and an ammonium acetate solution were mixed in a simple Tee mixer 2 mm *i.d.*, which was sufficient for this ~1 h reaction. The PFR was operated in a laboratory fume hood for the 29 kg GMP production run. The product solution was transferred to 1000 L pilot plant vessels for distillation and crystallization. Filtration and drying were done in a pilot scale agitated filter dryer. However, in a subsequent 77 kg production campaign, the downstream crystallization and filtration were also run continuously in the lab hoods at 8 kg/day throughput, eliminating the need for pilot plant vessels; 8 kg/day would have been commercial manufacturing scale for this product.

The *L/d* ratio for this reactor was about 20,000, which is unusually high compared to commercially available PFRs. This design kept axial dispersion low, even though reaction flow was in the laminar regime. An F-curve for this reactor is shown in Fig. 2.3, which was obtained by making a step change from 100% THF to a 3/2 (v/v) THF/toluene mixture flowing into the tube and monitoring the concentration of toluene in the reactor effluent with an on-line Raman probe. In the laminar flow regime, higher *L/d* is required to maintain low axial dispersion number as the reactor is scaled up to larger diameters, as described in another chapter in this book and in chemical engineering textbooks (Johnson 2020, Levenspiel 1962). Reynolds number was about 400 in this example with 97 min τ, and axial dispersion number *D/uL* was only 0.0009, which is very good plug flow behavior.

Another benefit to increasing the *L/d* is that the ratio of heat transfer surface area to unit volume (*A/V*) is higher. *A/V* was 516 m^2/m^3 in this 7.1 L PFR, but it would have been about 5 m^2/m^3 if a 1000 L batch reactor was used. *A/V* greater than

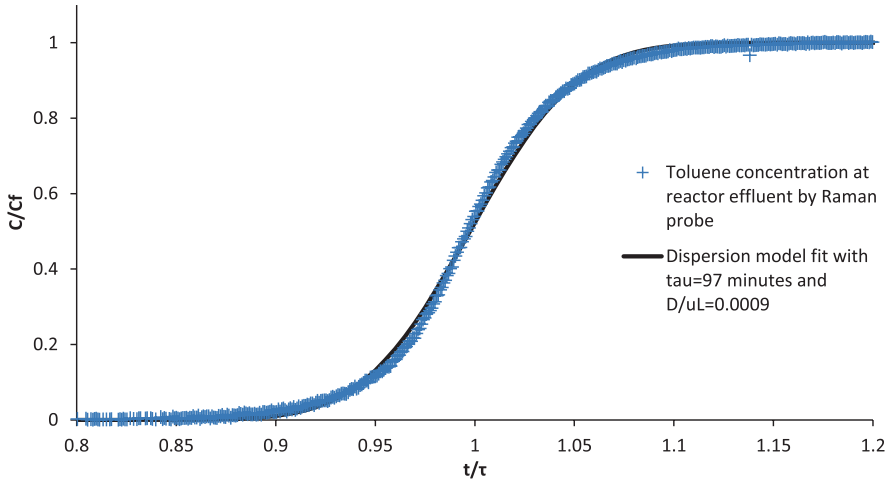


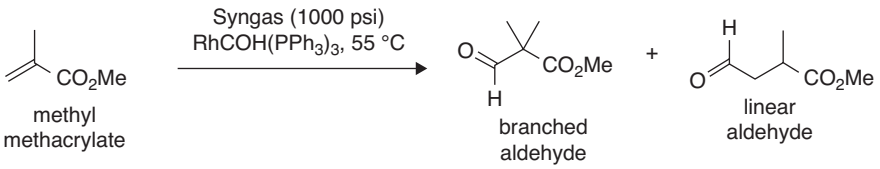
Fig. 2.3 Experimental F-curve data and model fit for solvent transition in a 7.1 L continuous thermal tube reactor

10,000 m^2/m^3 can be achieved in microreactors, but 516 m^2/m^3 was sufficient to achieve fast-enough heat-up time, cool-down time, and isothermal conditions for this 90 min reaction. A limit on the high side for *L/d* is pressure drop. However, the pressure drop from inlet to outlet during the flow reaction was only about 0.34 bar. Other limitations of high *L/d* were cost, ability to fabricate, and need to fit the reactor in an existing forced convection oven with an internal chamber that was a cube 46 cm on a side. The reactor was designed to be low cost so that it could be dedicated to this specific chemistry and then disposed when no longer needed for the project; in this case, the reactor cost $6000. While the oven cost about ten times more, the reactor was easily taken out of the oven and replaced with a new coiled tube PFR in the same oven for a subsequent product. For more details please see the publication (May et al. 2012).

2.3 Pulsating Flow Coiled Tube PFR for Hydroformylation with Solids Precipitate

A 32 L pulsating coiled tube reactor was used for a hydroformylation reaction where solids precipitated from the reaction mixture. The reaction used 1:1 $CO:H_2$ gas reagent at 68 bar pressure. Reaction temperature was 55 °C. There was no in-house batch scale-up option for this reaction because of safety restrictions regarding high-pressure CO. The continuous reaction used substrate to catalyst ratio (S/C) = 1000 with $RhCOH[PPh_3]_3$ catalyst. Mean reaction time (τ) was 24 h in the flow tube. Speeding up the reaction by increasing catalyst loading was not a viable option because of the cost of Rh. In addition, the catalyst precipitated from the

product solution. Increasing the catalyst loading would have exacerbated the potential solids clogging issues. Decreasing reaction time by increasing temperature was not a viable option because selectivity of branched (desired) to linear (undesired) aldehyde decreased at higher temperatures. The scheme is shown below.



The 32 L commercial manufacturing scale reactor was constructed from 152 m of 16.5 mm *i.d.* stainless steel tubing. The tubing was bent into five concentric cylindrical coils 0.53 m tall and ranging from 0.36 to 0.56 m diameter. The reactor and vessel for heating to 55 °C are pictured in Fig. 2.4. Vapor and liquid flowed cocurrently through each of the five coils in series in the uphill direction starting with the outside coil. These individual coils were linked by down-jumper tubes constructed from 316L stainless steel tubing with *o.d.* = 6.35 mm and *i.d.* = 4.57 mm. The reactor was designed with these narrow diameter down-jumpers to maintain maximum volume in the uphill flow direction (>99% of the total reactor volume), which helped the reactor run more liquid-filled in the forward direction, and subsequently made it easy to empty cleaning solvent from the reactor at the end of the campaign in the reverse direction. A picture of the reactor is shown in Fig. 2.4. The reactor cost was $12,000 and was kept deliberately low so that it could be dedicated to the process due to the inherent polymerization potential of the methyl methacrylate. In fact, the first time this chemistry was scaled up to an 8 L PFR, the tubes clogged with polymer and the reactor was discarded, which was not a significant issue because the 8 L tube reactor only cost about $7000. Therefore, the intention was to discard the 32 L reactor when no longer needed for this product because the polymer would be difficult to clean.

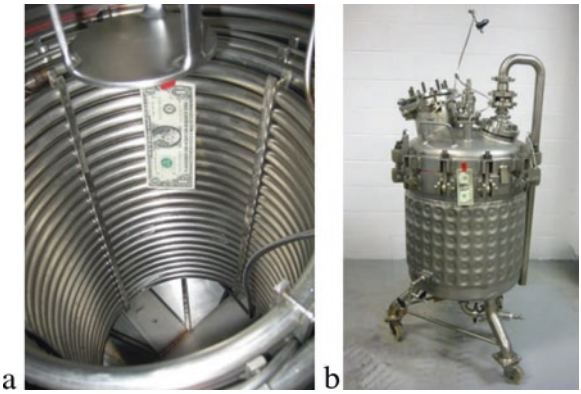


Fig. 2.4 Pictures of (**a**) 32 L coiled tube PFR and (**b**) constant temperature heating bath for submerging the reactor

The continuous production campaign generated 178 kg of aldehyde product for use in a GMP pilot plant campaign. This was an example of using continuous processing to enable speed to early phase material delivery. It was safer to scale up the process in flow because the gas supply was physically restricted, the reactor operated nearly completely liquid filled, and the reactor was smaller than batch for the same daily throughput. The reactor had pulsating flow in the forward and reverse direction. This was done to prevent solids clogging, because the catalyst precipitated in the aldehyde product solution and because some of the methyl methacrylate polymerized. Flow was forced to surge back and forth by about 1 m, which kept the solids from accumulating in the tube. The pulsating flow was generated by repeated reversal of pressure differential using sequenced automated block valves (Braden et al. 2009). After continuous filtration, the precipitated solids were quantified. It was found that the reaction product slurry was about 1% solids. The pulsating flow also increased gas/liquid mixing. The high selectivity of the desired branched to undesired linear aldehyde was evidence that gas/liquid mixing rate was high. Small-scale batch autoclave experiments had previously proven that selectivity was mixing sensitive. Figure 2.5 shows the high quality of product over the entire 320 h continuous production run. Desired branched aldehyde was >99% for the majority of the continuous run. The one low data point in Fig. 2.5, at 98.8% desired, resulted when the τ was deliberately reduced from 24 to 12 h for a day. The data in Fig. 2.5 was measured for samples taken after continuous fractional distillation. Area% desired branched aldehyde averaged about 97% before distillation. This fully

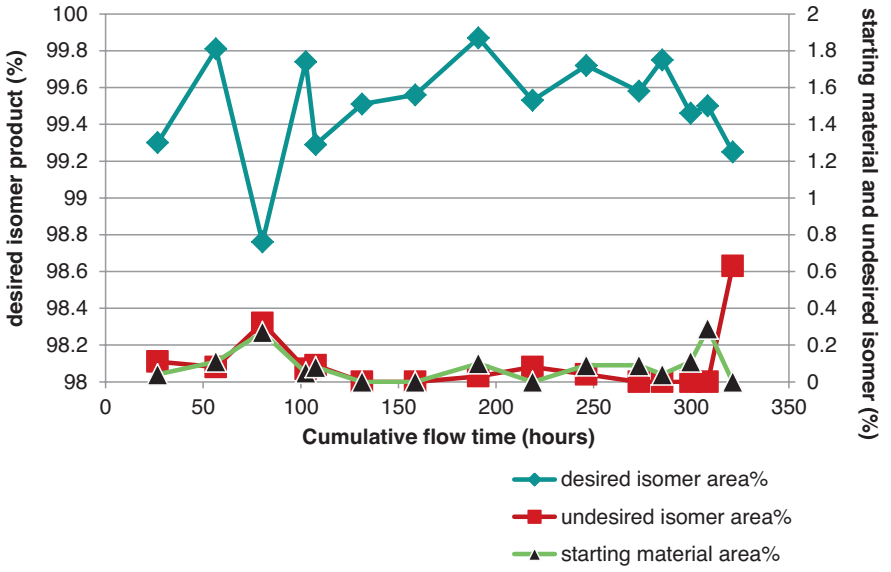


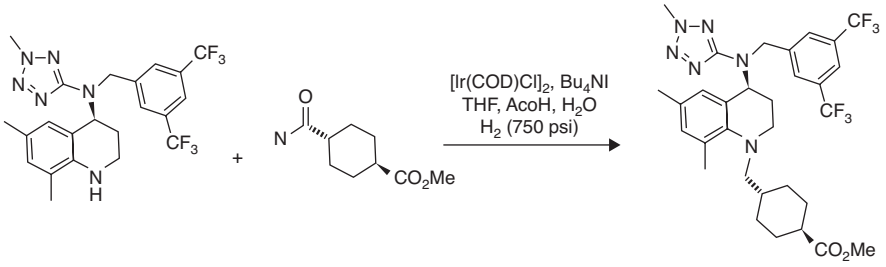
Fig. 2.5 Hydroformylation results for 320 h continuous production campaign in pulsating flow tube reactor

continuous synthetic route step consisted of continuous reaction, filtration, and multistage fractional distillation at 13 kg/day production rate, which was commercial manufacturing scale for this material.

Product quality from the continuous process was superior when compared to the batch option, which consisted of a TEMPO-catalyzed oxidation in methylene chloride solvent. The batch process generated 128 kg waste per 1 kg product. In contrast, the continuous process including purification by fractional distillation ran neat in methyl methacrylate without additional solvent and only generated 0.8 kg waste per kg product.

2.4 Vertical Pipes-in-Series PFR for Reductive Amination

A vertical pipes-in-series bubble flow reactor was designed and used for a two-phase gas–liquid reaction. This type of PFR design was scalable to larger volumes than the coiled tubes. A 360 L vertical pipes-in-series continuous reactor was used for a reductive amination to make >2000 kg GMP intermediate in a manufacturing plant (May et al. 2016). The reaction was run with dissolved $[Ir(Cod)Cl]_2$ catalyst. Reaction temperature was 20 °C, pressure was 50 bar, and mean reaction time τ was 12 h. Throughput for this system was 100 kg per day of penultimate, and the process ran nonstop for 24 days. Workup and isolation was done in batch 2000 gallon vessels, as well as the feed solution preparation for the continuous reactor. The chemistry is shown below.



The high cost of homogeneous catalysts and ligands, as well as the need to remove the metals from the product, drove the decision to a larger reactor and longer reaction τ in favor of minimizing catalyst loading. Speeding up the reaction to run with shorter τ and smaller reactor volume was limited because of a difficult to reject isomeric impurity that increased at higher temperatures. For these reasons, the reactor was designed for long τ and large volumes. Reagent gas and reagent solutions flowed through the reactor co-currently, in the upward direction through large diameter pipes and downward through small diameter tubing. A simplified schematic of a vertical pipe gas–liquid PFR system is shown in Fig. 2.6.

The product solution and excess reagent gas flowed continuously out the end of the reactor through a back pressure regulator and a vapor–liquid separator. A

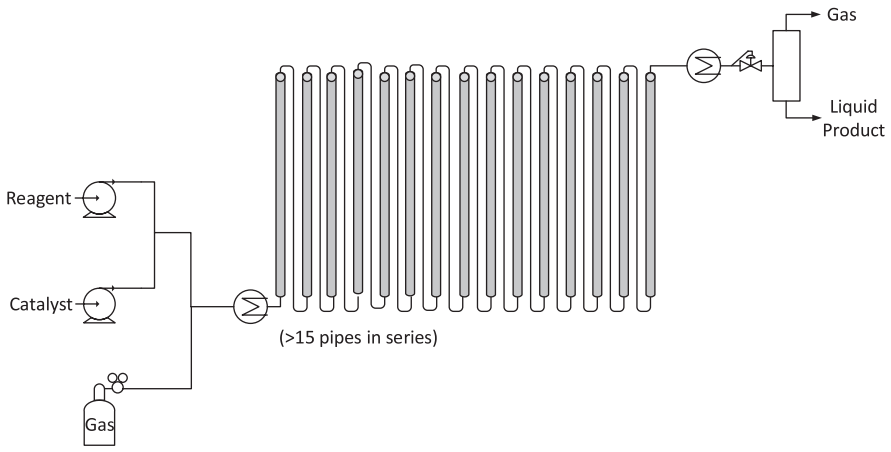


Fig. 2.6 Schematic diagram of a vertical pipes-in-series PFR system for two-phase reactions with reagent gas

vapor–liquid separator was required to keep liquid product from misting or spraying out the vapor vent when it depressurized. The gas residence time was much less than liquid because the reagent gas was orders of magnitude less dense than the liquid and therefore bubbled up through the liquid in the pipes. The reactor was temperature-controlled by enclosing in a shell with heat transfer fluid (water). The construction was similar to a large shell and tube heat exchanger; however, the main differences compared to a standard heat exchanger were that the pipes were connected in series not parallel, and the connecting tubes were of much smaller diameter than the pipes. Connecting the pipes in series resulted in higher gas/liquid mass transfer rates because it increased linear velocities through the tubing, for a given total mean residence time. It also decreased overall axial dispersion for the entire vessel, because increasing number of pipes in series decreases overall axial dispersion number, as described in the publication. (Johnson et al. 2016). The vertical down jumpers were of smaller diameter than the upflow bubble pipes to allow the reactor to remain about 98% liquid filled, increase gas/liquid mass transfer rates in the jumpers, and eliminate surging. A picture of the bottom of the 360 L PFR used for GMP manufacturing is shown in Fig. 2.7.

The stainless steel reactor was constructed from 45 vertical pipes in series, each 3.7 m tall and 53 mm *i.d.*, connected by down-jumper tubes with 4.6 mm *i.d.* At least 15 vertical pipes-in-series are recommended for keeping the overall axial dispersion number small and the overall vapor–liquid mass transfer and heat transfer rates high (Johnson et al. 2016) Other reactor designs such as falling film reactors and microchannel reactors provide much higher gas–liquid mass transfer rates, but mixing rates were sufficiently high such that the 12 h reaction was not mass transfer rate limited. The 360 L reactor was designed to operate >98% liquid filled and about 7 bar overall pressure drop when operating at 12 h τ. Another chapter in this book shows the start-up transition curve for this 360 L reactor in the manufacturing plant

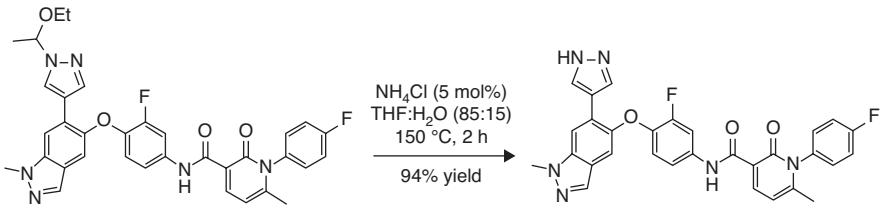
Fig. 2.7 Picture of a bottom of pipes and jumpers for a 360 L vertical pipes-in-series PFR for two-phase reactions with reagent gas used in GMP production

and numerically models the data, revealing that *D/uL* was 0.001. There were significant safety and capital cost advantages for this continuous reaction compared to the batch alternative. The main safety advantages were that the amount of hydrogen vapor space in the reactor at any time was low (about 2% of reactor volume) and that the reactor system was physically located outside of the building. The combination of these two factors would significantly diminish the consequences of a reactor leak. All hydrogen was outside the building at all times. Hydrogen supply cylinders were outside, the reactor was outside, and the vapor–liquid separator and gas stripping downstream from the reactor were outside. This is something that is feasible continuous but not batch. A batch hydrogenation autoclave must be inside because it is regularly opened to charge reagents and catalyst, but a continuous PFR is always sealed. An advantage of the vertical pipes-in-series instead of the coiled tubes was that it was easier to inspect surfaces after cleaning. In addition, the vertical pipes-in-series were scalable to larger reactor volumes than coiled tubes. For more details please see the publication (May et al. 2016).

2.5 Superheated PFR in Steam Shell for Thermal EE Deprotection

The thermal deprotection reaction shown in the following scheme was run in a thermal tube reactor because of the high temperature and pressure requirements (150–170 °C and 250–300 psig). These operating temperatures and pressures were beyond the batch reactor capabilities in the chosen manufacturing facility. The reaction thermolysed an ethoxyethyl (EE) protecting group from the starting material (Frederick et al. 2015) (Scheme 2.1).

The first time this continuous reaction was run in manufacturing, a 12 L reactor was used inside a steam shell, shown in Fig. 2.8. The Hastelloy reactor tube was



Scheme 2.1 Thermal deprotection of an N-Ethoxy Ethyl group achieved in a superheated PFR. (Reprinted with permission from Frederick et al. (2015). Copyright (2015) American Chemical Society)

Fig. 2.8 Pictures of (**a**) 12 L Hastelloy coiled tube PFR, and (**b**) steam jacket pipe for temperature control

255 m long and 7.75 mm *i.d.* The reaction was run at 150 °C, 17 bar, and 100 min τ (120 min V/Q). It was run as a hybrid batch/flow process, where the workup and isolation steps were run in 250 gallon batch vessels. Throughput in the continuous reactor was 36 kg/day, and the GMP campaign produced 150 kg API.

The second time that this same thermal EE deprotection ran in a scaled-up campaign, it was part of a multistep continuous process with reactions and separations unit operations run continuously. The last four synthetic route steps of the process were modified so that the entire process could run continuous in laboratory hoods. The reactor was a 2.5 L PFR made from 53 m long, 7.75 mm *i.d.*, Hastelloy C276 coiled tube, as shown in Fig. 2.9. It was used for the production of 20 kg API in a 100 h non-GMP continuous run (Cole et al. 2019). This 2.5 L Hastelloy PFR only cost $6800. It accomplished 8 kg/day throughput, which was commercial

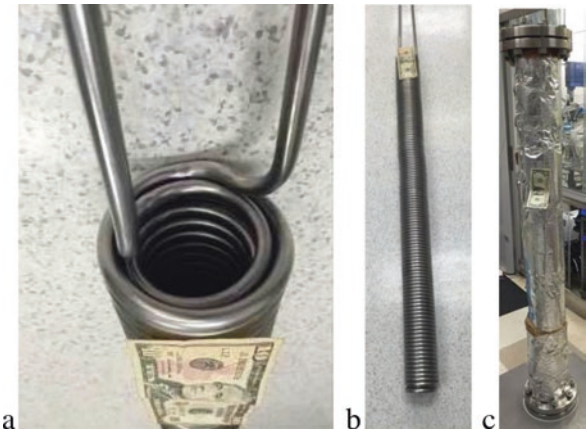


Fig. 2.9 Pictures of 2.5 L Hastelloy coiled tube PFR (**a**, **b**), and (**c**) insulated steam jacket pipe for temperature control

manufacturing scale for this product. The reactor was so inexpensive that it was considered dedicated and then disposable when no longer needed for this product. Figure 2.9 shows that the coils were much smaller diameter than the 12 L PFR; therefore, the reactor fit inside a 10 cm inside diameter 1.2 m tall steam heating pipe. This reduced the cost and footprint of the steam heating shell as well compared to the 12 L PFR. The smaller diameter steam shell had higher pressure rating than the steam shell used for the 12 L PFR; therefore, it was possible to operate with 7.8 bar steam pressure on the shell side and consequently increase reaction temperature to 170 °C. Experiments verified that there was not significant hydrolysis of the amide bond at this temperature. Reaction pressure was 22 bar to maintain liquid phase in the PFR, and mean residence time (τ) was 29 min. Amide feed solution was heated to 30 °C to maintain solubility; therefore, the feed solution was pumped from a heated vessel through heat-traced tubing and a heated dual syringe pump. For more information, please see the publications, especially supporting information for the Cole et al. publication (Frederick et al. 2015; Cole et al. 2019). The heat transfer to and from the reactor is much better for steam jacketing than for a forced convection oven like the one shown in the imidazole cyclization example.

Another benefit of the smaller radius coils was that it reduced the overall axial dispersion number compared to tubular PFRs with larger radius bends. The pushout F-curve was measured with an on-line refractive index probe during the switch from product to solvent at the end of the production time. The F-curve and axial dispersion model fit are shown in Fig. 2.10. Once again, flow was laminar; nevertheless, the axial dispersion number *D/uL* was only 0.00075.

The third time that this same thermal EE deprotection was conducted in a scaled-up campaign, it was part of a multistep continuous process with reactions and separations unit operations in a 183 kg GMP production campaign (Reizman et al. 2019). The PFR was 7.2 L, made of 7.75 mm *i.d.* coiled tubing heated in a forced convection oven, very similar to the PFR described previously for the imidazole cyclization reaction. The geometry of both reactors was almost identical, but the PFR used for the thermal EE deprotection was Hastelloy C276 rather than 316L

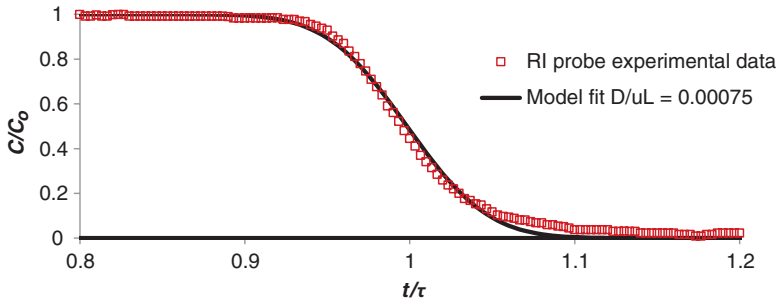
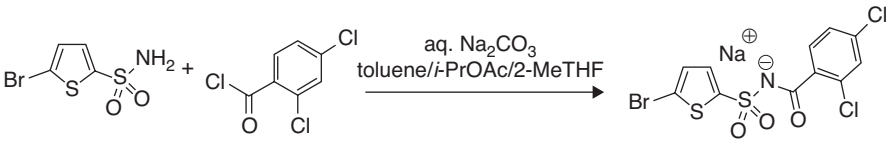


Fig. 2.10 Pushout F-curve and axial dispersion model fit for thermal deprotection in coiled tube PFR

stainless steel. The deprotection reaction was run at 170 °C temperature, 25 bar backpressure, and 40 min τ. The reaction was complete in about 25 min, but excess residence time was used to provide margin for throughput changes. For more detailed information please see the publication (Reizman et al. 2019). Two countercurrent heat exchangers in series were used to simultaneously preheat the reagent feed solution entering the PFR and cool the product solution exiting the PFR. In other words, the feed solution cooled reaction product, while product solution heated the reagent feed by flowing in opposite directions through tube in tube heat exchangers. This type of heat integration demonstrates an energy efficiency advantage of a continuous process. In batch, a reaction is typically heated by plant utilities at the start of a reaction, then held at constant temperature for a designated reaction time, and finally cooled by plant utilities at the end of reaction. In contrast, the continuous PFR process with steady-state countercurrent heat exchange accomplishes about half the PFR heating and cooling requirements at no utility cost, thus lowering the environmental footprint.

2.6 CSTRs-in-Series for Schotten-Baumann with Two Liquid Phases

Continuous stirred tank reactors (CSTRs) in series were used for a Schotten-Baumann reaction that produced a cytotoxic API. The chemistry is shown in the following scheme. Operating temperature was 65 °C, and τ was 1 h in each CSTR (3 h total τ).



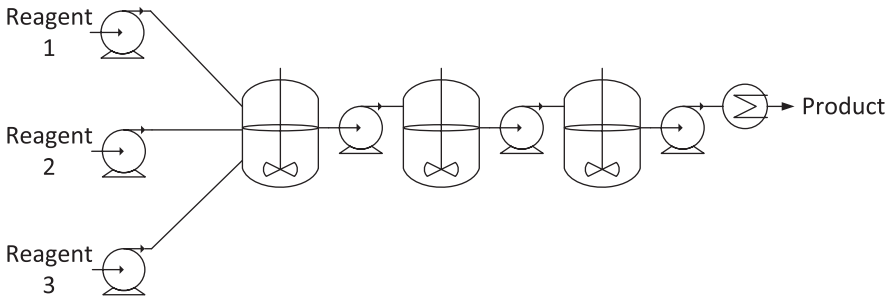


Fig. 2.11 Schematic diagram of CSTRs-in-series reactor system for heterogeneous continuous reactions

Fig. 2.12 Picture of three 12 L CSTRs-in-series used for a continuous Schotten-Baumann reaction

The reaction required 45 min to reach 99% conversion batch. The two-phase liquid–liquid mixture in the reactor required fast mixing to minimize an amidine impurity (White et al. 2012). A PFR was not a practical option for this chemistry because of the requirement for fast liquid–liquid mixing over long τ, but CSTRs in series were capable. Additionally, CO_2 was evolved during the reaction. This introduced a third phase, which was easily tolerated in the CSTRs without impacting τ. A simplified schematic of a reactor system with CSTRs-in-series is shown in Fig. 2.11. Three feeds continuously pumped into the first CSTR, and the two-phase mixture was pumped continuously to the subsequent CSTRs, maintaining constant liquid levels.

A photograph of the three CSTRs-in-series reactor system used for 5 kg/day production of 20 kg cytotoxic API is shown in Fig. 2.12. Each stirred tank was a 12 L glass reactor, which was commercial production scale for this product. These reactors were very inexpensive, only about $500 each, and they were disposable at the end of the API production campaign. Cross-contamination potential of one

cytotoxic compound in another from the reactor is not possible when the reactor is dedicated to only a single cytotoxic product.

One of the main advantages of continuous instead of batch for this process was containment of the cytotoxic API in fume hoods. Reaction, extraction, solvent exchange, and crystallization were all run in fully continuous, inexpensive, portable, disposable equipment, dedicated to a single API, removing cross-contamination potential with other APIs.

Given that the batch reaction was 45 min to reach 99% conversion, calculations can be made to estimate what residence time is required in CSTRs to achieve the same conversion. Assuming a pseudo first-order reaction, the rate constant for a 45 min batch reaction can be calculated using Eq. 2.1:

$$\frac{C_{\mathrm{AF}}}{C_{\mathrm{Ao}}} = e^{-kt} \tag{2.1}$$

If the reaction was first order, then $C_{\mathrm{AF}}/C_{\mathrm{Ao}} = 0.01$ and $t = 45$ min corresponds to rate constant $k = 0.1$/min.

The design equation for three equal-sized CSTRs-in-series for irreversible first-order reaction is given in Eq. 2.2 (Levenspiel 1962):

$$\frac{C_{\mathrm{AF}}}{C_{\mathrm{Ao}}} = \left(\frac{1}{1+k\tau}\right)^3 \tag{2.2}$$

Calculating for 99% conversion $C_{\mathrm{AF}}/C_{\mathrm{Ao}} = 0.01$ and rate constant $k = 0.1$/min, τ is calculated to be 36 min.

With three CSTRs-in-series, τ is 36 min in each CSTR, and thus $\tau = 108$ min total in all three. This is about 2.4 times longer than batch or ideal PFR to achieve 99% conversion, but the reactors are still much smaller than batch because they are able to produce continuously. At steady-state operation, the first CSTR in series ran at 88% conversion of the sulfonamide starting material, the second at 98% conversion, and the third at 99% conversion. Thus, there was 88% conversion in CSTR1, 83% conversion of the remaining unreacted starting material in CSTR2, and 50% conversion of the remaining unreacted starting material in CSTR3. These would all be the same, for example 88% conversion of remaining SM in each CSTR, if the kinetics were really simple first order. This is evidence that the reaction was more complex; nevertheless three CSTRs in series each with 1 h τ achieved the goal of 99% conversion. For more information please see the publication (White et al. 2012).

As the number of CSTRs-in-series increases, the behavior of conversion versus time approaches that of an ideal PFR. This is shown by the simulated RTD trends in Fig. 2.13. Here, C is concentration at time t, C_f is final concentration of non-reacting tracer at steady state. At time = 0, a step change is made from 0 to C_f entering the first reactor in the numerical simulation. The squares represent the F-curve response to a step change in feed for three equal-sized CSTRs in series. The circles represent an F-curve response to a step change for 100 equal-sized CSTRs in series, and the

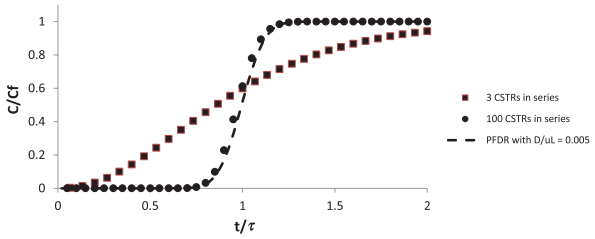


Fig. 2.13 F-curve transitions comparing RTDs for CSTRs-in-series versus typical plug flow with dispersion reactor

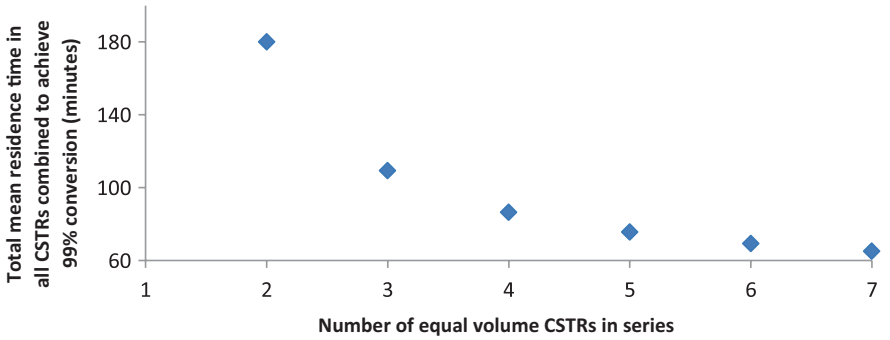


Fig. 2.14 Total τ in CSTRs in series as a function of number of CSTRs-in-series

dotted line represents an F-curve response to a step change for a plug flow with dispersion reactor having axial dispersion number $D/uL = 0.005$.

In most situations, three CSTRs-in-series is a practical number, and more than three in series gives diminishing returns regarding performance versus the complexity and cost of additional equipment.

Figure 2.14 shows the diminishing returns for increasing number of CSTRs in series for a reaction with 0.1/min first-order rate constant.

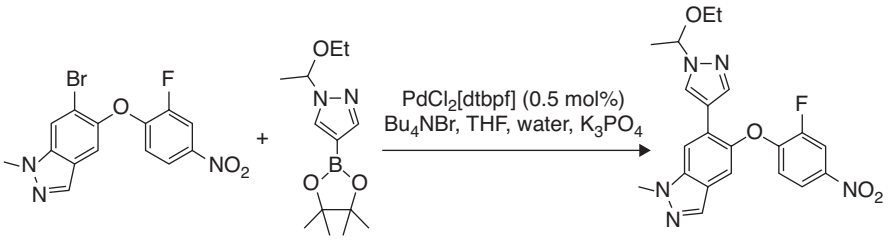
Compared to a single CSTR, using three CSTRs in series drastically reduces the overall total reactor residence time required to achieve a given target % conversion of reagents to products, if the reaction has positive order; however, total time to the same target conversion is still less in batch or ideal PFR. For example, consider an elementary first-order reaction. If the reaction is 99% converted in 1 min batch, then elementary first-order rate constant is 4.56/min. in a single CSTR this same reaction would reach 99% conversion in 21 minutes mean residence time, and it would reach 99% conversion in 2.4 munites mean residence time in 3CSTRs in series (0.8 min in each of the three CSTRs). If the elementary first-order reaction is 99% converted in 30 min batch, then first-order rate constant is 0.152/min. in a single CSTR this same reaction would reach 99% conversion in 650 minutes mean residence time, and it would reach 99% conversion in 72 minutes mean residence time in 3CSTRs in series (24 min in each of the three CSTRs). The comparison is summarized in Table 2.1. The simple calculations were done using the Eq. 2.1 and Eq. 2.2 for batch and CSTR reactors.

Table 2.1 Hypothetical calculations for reaction conversion time as a function of rate constant and reactor type for an elementary first-order reaction

k (1/min)	τ (min)	C/C_o batch or PFR	C/C_o, 1 CSTR	C/C_o, 3 CSTRs-in-series
4.56	1	0.01	0.18	0.06
4.56	21	<0.0000	0.01	<0.0001
4.56	2.4	<0.0000	0.08	0.01
0.152	30	0.01	0.18	0.06
0.152	650	<0.0001	0.01	<0.0001
0.152	72	<0.0001	0.08	0.01

2.7 Intermittent Flow Stirred Tank Reactor for Suzuki Cross-Coupling with Two Liquid Phases

An intermittent flow stirred tank reactor was used as an alternative to CSTRs-in-series for the Suzuki–Miyaura cross-coupling chemistry shown in the scheme below (dtbpf = 1,1′-Bis(di-tert-butylphosphino)ferrocene]) (Cole et al. 2016).



Reaction conditions were 64 °C temperature, 25 min reaction time, and 0.5 mol% catalyst. The reactor had three individual feeds from three automated intermittent flow feed pumps. The organic feed solution contained 7 wt% indazole, 6 wt% boronic ester, and 0.3 wt% TBABr in THF solvent. The aqueous feed solution contained 20 wt% K3PO4 in water. The catalyst feed solution contained 4 wt% PdCl2[dtbpf] in dichloromethane solvent. The resulting two-phase liquid–liquid reaction mixture required high agitation because the reaction rate was mixing sensitive. A stirred tank reactor was used rather than a PFR because mixing rate is independent of reaction residence time in a stirred tank, while mixing rate is a function of linear velocity in a PFR with static mixing elements. Reaction time batch was about 20 min and the reaction rate was positive order, therefore a truly continuous single CSTR would not be suitable for full conversion in a reasonable time. A residence time (τ) of 44 h would have been required to reach 99.9% conversion in true CSTR, while only 20 min was required to reach 99.9% conversion in this intermittent flow stirred tank reactor. CSTRs-in-series were considered, but the single intermittent flow reactor was selected because it was better for catalyst performance and impurity profile to do a controlled addition of catalyst solution. For more information please see the publications (Cole et al. 2016, 2019).

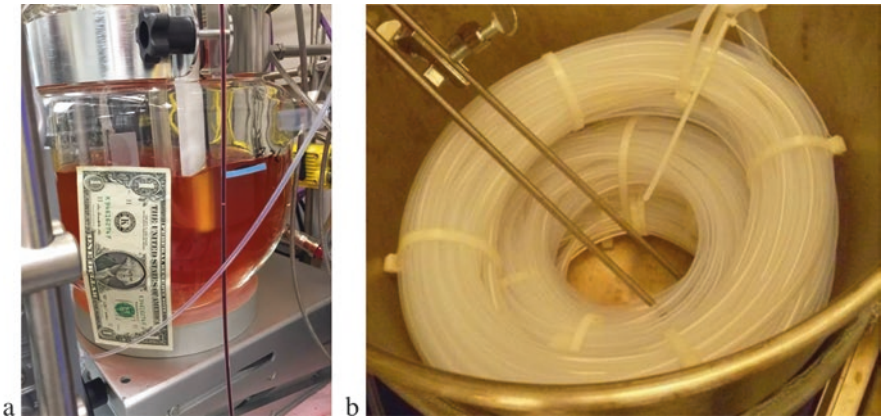


Fig. 2.15 Pictures of (**a**) the 6 L stirred tank that was used for an intermittent flow reactor, and (**b**) the coiled tube heat exchangers to warm the reagent solutions to 60 °C before they entered the reactor

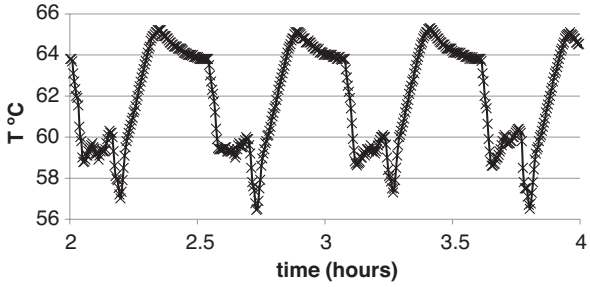


Fig. 2.16 Temperature versus time trend plot for a 2 h segment of the 4-day continuous production campaign with intermittent stirred tank reactor

Figure 2.15 shows a picture of the 6 L stirred tank that was used for a 100 h continuous run to make 22 kg of product. The reactor operated in automated fill-empty mode. It filled quickly, held at reaction conditions for 20 min, and then emptied quickly, in a repeating fashion.

The reactor was inexpensive ($2600), so that it could be dedicated to this product and disposed when no longer needed for this API intermediate. Furthermore, this 6 L reactor with 3 L preheat coils was about 1/2 commercial scale for this low volume API.

Figure 2.16 shows the temperature profile inside the reactor for 2 h of the continuous campaign, and illustrates the reproducibility of the system.

This reactor was very productive. It turned over automatically about 45 times per day. The nonproductive reaction time for filling and emptying was brief. The reagents entered the reactor at 60 °C because they flowed through heat exchangers on the way into the reactor. Heat of reaction increased temperature up to 66 °C in 8 min. The reaction occurred at about 64–65 °C for about 18 more minutes. The sharp drop-off in temperature seen for each cycle in Fig. 2.16 was due to the reactor being rapidly emptied, just prior to refilling automatically. The cycle was fully automated and very repeatable. The heat exchangers were inexpensive perfluoroalkoxy

(PFA) coiled tubes submerged in a constant temperature water bath, shown in Fig. 2.15b. The organic and aqueous feeds each flowed through 1.95 L heat exchange tubes that were 6.37 mm *i.d.* and 61 m long. The coils were large enough to hold the feed volume for the next reaction iteration; therefore, the materials had 32 min heat-up time in the coils. These PFA preheat tubes added about $2000 to the cost. It is possible to design shell and tube heat exchangers with much higher surface area to volume ratio, and higher heat transfer coefficients on both the process side and shell side, but higher heat transfer rates were not needed for this process given the 32 min heat-up time in the coiled tubes

This was not a truly continuous process, but it operated continuously, and it integrated well with an otherwise fully continuous process in laboratory fume hoods. The first stirred tank in a continuous mixer-settler extraction downstream served as surge capacity for the intermittent flow, as the liquid level in the mixer oscillated up and down each reaction cycle. In automated fill-empty mode, an intermittent flow stirred tank reactor like this can be operated with either controlled addition of one reagent to the other or co-addition of reagents. The choice depends on what is best for impurity profile at the end of reaction. The pumps are automated by the control system, and the user decides the order of reagent addition, the rate of reagent addition, and the total reaction time. This is something that is not possible in truly continuous PFRs or CSTRs.

Intermittent flow stirred tank reactors are also used for heterogeneous reactions with solid/liquid/gas at elevated pressure like hydrogenations (Cole et al. 2017a). They are not truly continuous, but they have a large number of reactor volume turnovers per day and integrate well with otherwise continuous multistep processes.

2.8 CSTRs-in-Series for Barbier Grignard, Quench, and Neutralization to Minimize Racemization

A Barbier Grignard formation and coupling, along with quench and neutralization, were run fully continuous in CSTRs-in-series in order to minimize racemization of an unstable tetrahedral intermediate, shown in the following scheme (Kopach et al. 2012; Braden et al. 2017). Racemization was minimized by reducing the time to quench with HOAc and time to neutralize with Na_2CO_3. The product was stable after neutralization.

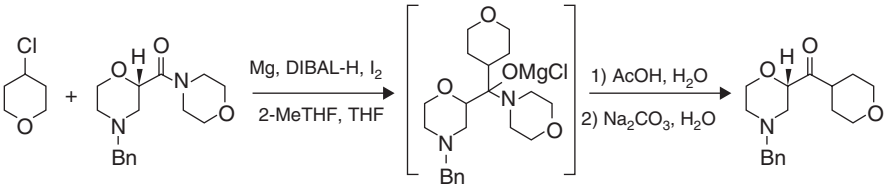


Figure 2.17 shows a picture of the CSTR train used for the 5 kg/day production in laboratory fume hoods.

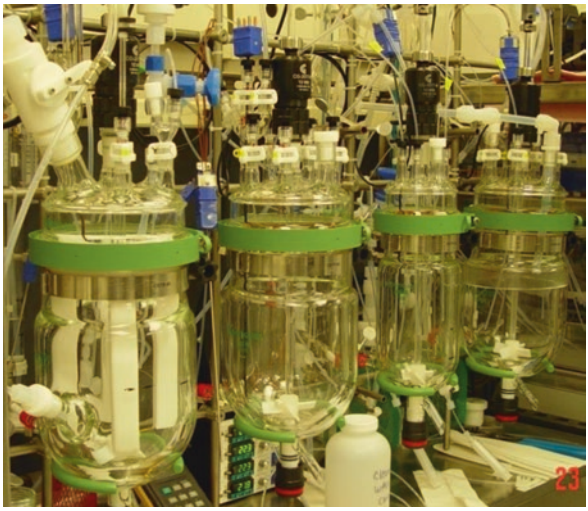


Fig. 2.17 Picture of 2 L CSTRs in series used for continuous Barbier Grignard, quench, and neutralization (the third vessel from the left was not used). (Reprinted with permission from Braden et al. (2017). Copyright 2017 American Chemical Society)

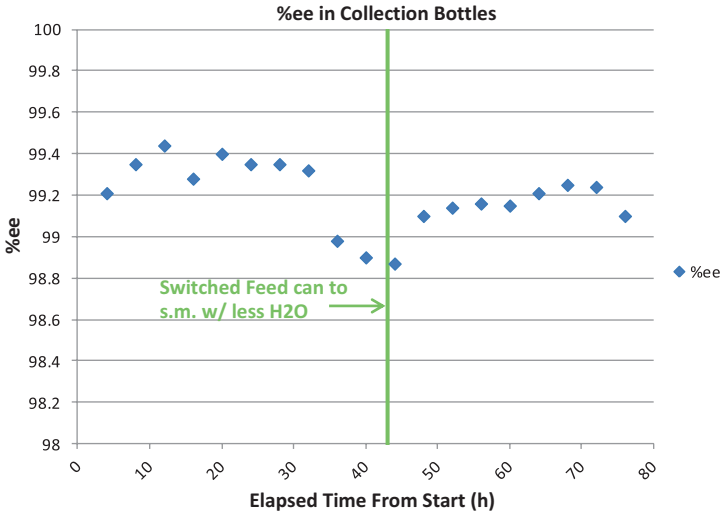


Fig. 2.18 *ee* versus time for neutralized reaction product solution from continuous Barbier Grignard, quench, and neutralization. (Reprinted with permission from Braden et al. (2017). Copyright 2017 American Chemical Society)

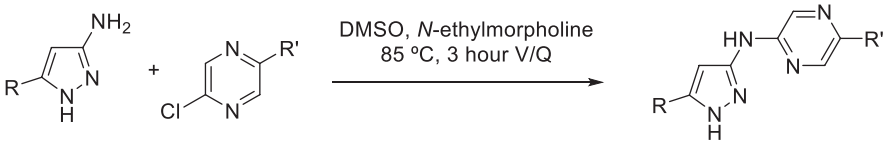
The reason why a CSTR was used for the Grignard reaction, instead of a PFR, was because of the solid Mg reagent. Mg was sequestered in the CSTR by an internal settling pipe. Liquid continuously flowed into and out of the CSTR, but the solid Mg remained until it was used up. Mg solids were intermittently manually charged to the CSTR once every 4–8 h. The freshly added Mg was initiated by the stirring Grignard mixture. Any Mg fines that escaped the internal settling pipe were caught in a downstream Mg solids trap and returned to the CSTR. Figure 2.18 shows the high *ee* maintained for the entire 75 h continuous run, which averaged >99%.

In batch scale-up, the best that was typically achieved was about 98% *ee* after neutralization, and it was inconsistent batch to batch. Chrial purity was maintained more reliably in the CSTRs because time prior to quench and neutralization was minimized. Batch crystallization of this intermediate achieved *ee* greater than 99%. However, continuous processing afforded the opportunity to eliminate the crystallization and telescope directly into the next synthetic route step.

Furthermore, this Grignard formation reaction had runaway potential because it was highly exothermic and difficult to initiate. Operating batch mode in 100 L equipment would have provided about the same overall throughput as the 2 L continuous reactor. Thus, continuous reactor volume was 50 times smaller than batch reactor volume for the same overall throughput, which was a significant safety advantage of the continuous process. Another safety advantage of the continuous reaction versus batch is that much less excess Mg required quenching at the end. Quenching activated Mg generates hydrogen gas; therefore, the quantity of hydrogen gas generated from the continuous process was less than batch. For more information please see the publications (Kopach et al. 2012; Braden et al. 2017). Calculation of real-time energy balance for a Grignard formation reaction is given in a different book chapter (Johnson et al. 2019).

2.9 Low-Pressure Disposable Coiled Tube PFR for Highly Potent Compound

A nucleophilic aromatic substitution reaction (S_NAr) coupling a pyrazine and a pyrazole is shown in the following scheme (Cole et al. 2017b):



The reaction was run continuously in a 2.9 L PFR, which was constructed from 91 m 6.38 mm *i.d.* PFA tubing, coiled to fit inside a cylindrical constant temperature bath that was about 0.5 m diameter. Liquid flow was in the bottom and out the top of the horizontally coiled tube so that it would operate liquid filled. The continuous reaction process produced >20 kg API during a 200 h GMP production campaign (Cole et al. 2017b). Two feed solutions were pumped into the reactor and mixed in a simple Tee mixer and tubing section with 1.6 mm *i.d.* The first feed solution was 20.4% pyrazole and 8.3% N-Ethylmorpholine in DMSO solvent, and the second feed solution was 16.8% pyrazine in DMSO solvent. N-Ethylmorpholine was selected as the base partly because of its solubility in the reaction solvent. Flow rates for the two feed solutions were 10.2 ml/min and 5.4 ml/min, respectively, providing about 2.9 kg/day of the pyrazole product after continuous crystallization. Reaction temperature was 70 °C and τ was 180 min. The reaction was actually done in about 120 min, but product was stable to end of reaction conditions for the extra

hour, which provided some margin in favor of full conversion. The S_NAr coupling reaction could have run well batch and yielded the same result. Because temperature was only 70 °C and solvent was DMSO, the reaction operated at atmospheric pressure, therefore it did not have any of the extreme pressure/temperature conditions that would normally drive the decision to use a continuous reactor. However, a continuous reactor option was selected in this case because this chemical transformation fell in the middle of an otherwise fully continuous process in laboratory fume hoods. There were several advantages to running the entire process continuously at small scale, as described in the publication. Fume hoods provided extra containment for the cytotoxic compound, a hydrazine reaction was run more safely and efficiently in a PFR, countercurrent multistage extraction removed excess hydrazine while minimizing product loss to an aqueous phase, and two isolations were eliminated in the continuous process compared to batch. The product was highly potent and had a low projected annual demand, therefore the 2.9 L PFR was sufficiently large for commercial manufacturing scale. The reactor tube itself only cost $1371. It was designed to be disposable at the end of an API production campaign. As stated in a previous example, cross-contamination potential of one cytotoxic compound in another from the reactor is not possible if the reactor is dedicated to only a single cytotoxic product. A picture of the reactor and the constant temperature heating bath is shown in Fig. 2.19.

In this case, the reactor did not need to be designed for exceptional heat transfer properties. Calculated heat-up time to 70 °C required about 300 s (Fig. 2.20). That is not very fast heat-up time for a PFR, but since it was a 2 h reaction in a reactor with 3 h τ, the 5 min heat-up time was negligible. An array of microreactors with higher surface area to unit volume ratio could give a much faster heat-up time, but it was not needed for this chemistry, and it was more valuable to keep the reactor inexpensive and disposable because of the cytotoxic product.

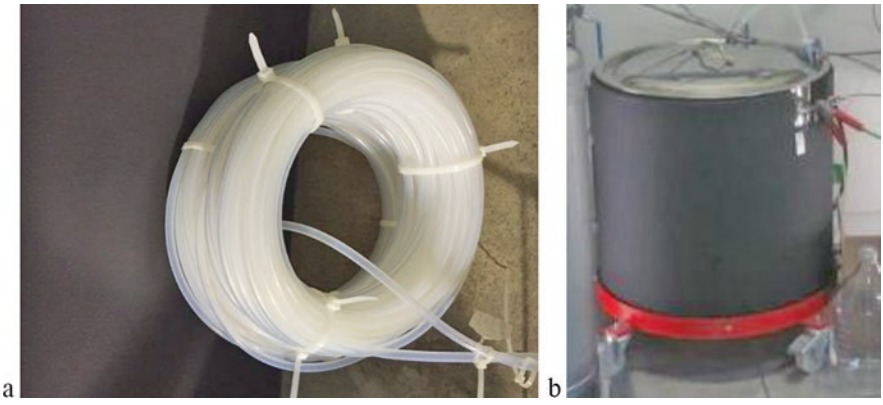


Fig. 2.19 Pictures of (**a**) disposable coiled tube low-pressure PFR for highly potent compound, and (**b**) constant temperature water bath for submerging the PFR

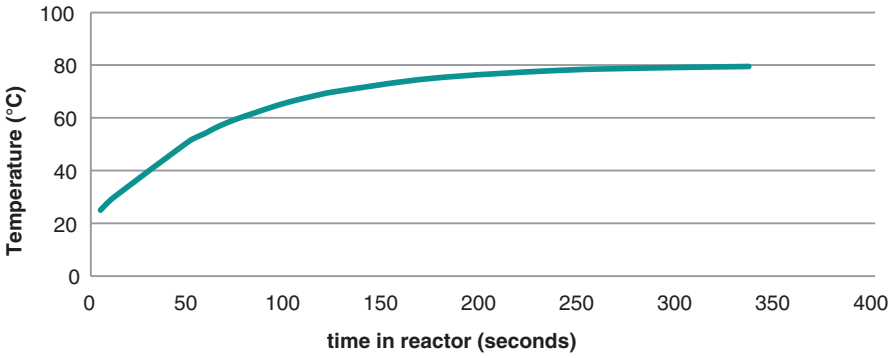


Fig. 2.20 Simulated heat-up time for SNAr reaction in 2.9 L PFA coiled tube reactor

2.10 Understanding the Real Reaction Temperature in a PFR

It is often difficult to measure in-process temperature at the desired location inside the PFR, for example, at the "hot spot." Furthermore, it is usually not feasible to measure temperature at all points along the length of a PFR. If heat transfer is happening, then process temperature is not the same as jacket temperature, therefore it is important to understand the limitations. Packed catalyst bed reactors are particularly challenging because it is difficult to install thermocouples in the middle, heat transfer surface area to volume ratio (*A/V*) is usually low because of larger diameters, and it is difficult to remove heat from the "hot spot" because of low mixing in the radial direction. Highly exothermic reactions are also challenging because of hot spots. Microreactors are known for exceptional heat transfer, maintaining isothermal reactions and minimizing hot spots, because of the <1 mm, and often <0.1 mm, characteristic dimension (Fogler 2010). However, not all PFRs have high heat transfer coefficients, for example, a coiled tube PFR heated inside a forced-convection oven. Scaled-up PFRs may not have such high *A/V* for heat transfer. Because it is not possible to measure real in-process temperature at all points along the length of the PFR, this is best modeled numerically. Numerical calculations require reaction energetics and kinetics, fluid physical properties, reactor *A/V*, and measured heat transfer coefficients.

The temperature profile along a reactor affects both reaction rate and fluid residence time. As the temperature of process fluids within a reactor changes, the reaction rate will change. Changing reaction rate will change the rate at which heat is consumed or evolved by the reaction, which will in turn affect the fluid and reactor temperature. These thermal effects due to the reaction are coupled to the heat exchanged between the process fluid, the reactor, and the environment. This non-isothermal operation is captured by an energy balance around the tubular

reactor. Fogler (2010) expressed the energy balance of a non-isothermal tubular reactor as Eq. 2.3, which is integrated along the length of the reactor to give the temperature profile.

$$\frac{dT}{dV_{\text{reactor}}} = \frac{r_A(\mathrm{T})\Delta H_{\text{rxn}}(T) - Ua(T - T_a)}{\sum F_i C_{p_i}} \tag{2.3}$$

where $\frac{dT}{dV_{\text{reactor}}}$ is the change in temperature of the process fluids in a segment of the reactor; $r_A(T)$ is the reaction rate of species A at temperature T; $\Delta H_{\text{rxn}}(T)$ is the heat of reaction at temperature T; U is the overall heat transfer coefficient between the heating media, the reactor, and the process fluids; a is the surface area per unit volume of the reactor; T is the temperature of the process fluids; T_a is the temperature of the heating or cooling media; and $\bullet\ F_i C_{p_i}$ is the heat capacity of all species within the reactor, weighted according to composition.

In addition to affecting reaction rate, temperature changes increase or decrease fluid residence time within a reactor due to thermal expansion or contraction of the process fluids. As a fluid is heated, it will generally expand, creating a larger volumetric flow rate within the reactor and decreasing residence time. Conversely, for a cooled fluid, the residence time within a tubular reactor will increase. Thermal expansion is described by Eq. 2.4 (Bridgman 1914).

$$\left.\frac{\delta V_{\text{fluid}}}{\delta T}\right|_P \frac{1}{V_{\text{fluid}}} = \beta \tag{2.4}$$

where $\left.\frac{\delta V_{\text{fluid}}}{\delta T}\right|_P$ is the change in fluid volume with temperature at constant pressure, V_{fluid} is the volume of the fluid, and β is the coefficient of volumetric thermal expansion. In order to completely model a non-isothermal tubular reactor, the thermal effects on reaction rate and residence time must be solved simultaneously with a material balance. This gives a system of first-order nonlinear differential equations.

Some parameters within this model such as flow rates and jacket temperature set points are generally known. Other parameters used in this model require prior experimentation or can be estimated from calculations. For example, the overall heat transfer coefficient for the tubular reactor can be experimentally characterized, or it can be predicted using Nusselt number correlations for internal and external flows (Bergman et al. 2011). Thermal expansion coefficients have been experimentally measured for numerous fluids (Riddick et al. 1986; Lide and Kehiaian 1994). The heat of reaction for a molecule can be experimentally measured or it can be estimated via Hess' Law (Fogler 2010). Finally, the reaction kinetics and activation energies must be experimentally determined for each reaction.

This approach was used to model a homogeneous, cryogenic lithiation in a tubular reactor. The model was a custom Euler method solution in MATLAB® that solved Eqs. 1 and 2 along with a differential mass balance.

Euler method solution in MATLAB:

$$\Delta T_i = \frac{2\pi r U\left(T_{\text{ext}} - T_i\right) + \Delta H_{\text{rxn}} r_{A_i} A_{\text{cross}}}{\dot{m} C_p} \Delta L_i, \quad r_{A_i} = f\left(T_i\right) \tag{1}$$

$$T_{i+1} = T_i + \Delta T_i$$

$$\Delta L_{i+1} = \Delta L_0 \left[1 + \beta\left(T_{\text{ref}} - T_i\right)\right] \tag{2}$$

Material balance:

$$\left[A\right]_{i+1} = \left[A\right]_i + r_{A_i} \Delta t$$

The cryogenic lithiation reaction was also run experimentally. The reaction was mixing rate limited and performed within a helical ribbon static mixer. This allowed for temperature measurement only at the inlet and outlet of a 22.9 cm long, 3.2 mm inside diameter (*i.d.*) static mix tube. The reactor feeds flowed through cooling heat exchangers, combined in a mixing tee, and then flowed through the static mixer. As the feeds rapidly mixed, the heat of reaction created a hot spot within the static mixer, accelerating the reaction rate. The temperature of this hotspot increased until the heat generated by the reaction at that point in the reactor is matched by the heat removed by the cryogenic bath. This hot spot, however, was not physically measurable. The thermocouples immediately before and after the static mixing tube reported −76 °C and −74 °C, respectively, at steady state. These temperatures corroborate the model results depicted in Fig. 2.21. However, the model results indicate

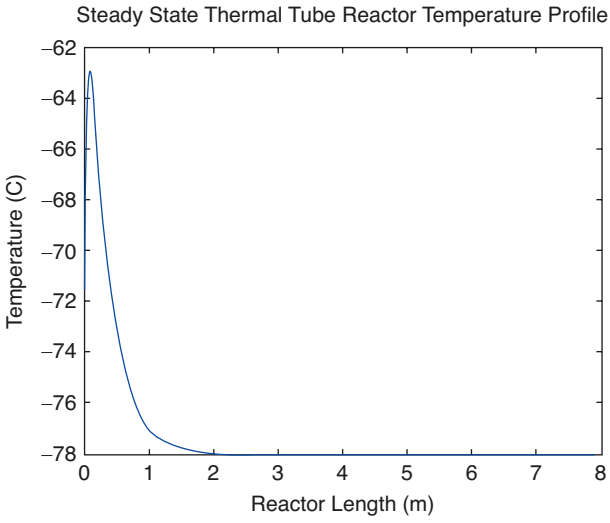


Fig. 2.21 Steady-state temperature profile along the length of a tubular reactor for a homogeneous, cryogenic lithiation reaction, calculated by numerical model using MATLAB®

that the reactor hot spot, where the reaction temperature was about −63 °C, was not monitored by the thermocouples because it was not physically possible to put a tee with thermocouple at this length. This underlines the importance of temperature profile modeling in understanding actual reaction temperature and illustrates the inaccuracy of reporting a single reaction temperature at the inlet and outlet of the static mixing section.

Nevertheless, temperature must be measured at a specific location in the GMP manufacturing plant. In this example, the best practical choice may be to insert in-process thermocouples into a Tee right after the reagents first mix, understanding that this does not necessarily match the reaction hot spot.

The best practical measurement choice for GMP temperature measurement may be different for different types of PFRs. If the PFR is a shell and tube heat exchanger with live steam condensing on the shell side, then steam temperature on the shell side may be the best indicator of reaction temperature. In fact, if condensate is not building up and if steam is actively condensing on all surfaces of the reactor, then steam *pressure* may be the best indicator of reaction temperature. If the PFR is a packed catalyst bed reactor, and if adiabatic temperature rise is known, then it may be best to preheat the feed solutions to desired reaction temperature minus adiabatic heat rise and let the heat of reaction bring the process temperature up to target value. In this scenario, the packed bed reactor would be insulated so that would operate adiabatically. This technique minimizes the temperature gradient radially. If insulation is efficient, then the wall temperature will be close to the temperature in the middle of the packed bed. Thus, temperature measurement at the wall may be sufficient for an adiabatic packed catalyst bed reactor. In cGMP processing, a clear temperature measurement point will need to be defined, and manufacturing will need to prove they maintained a temperature in a range to satisfy a proven acceptable range target. Whatever was measured in the lab during development should be the starting point for what is measured in the plant.

Acknowledgments Doug Kjell and Chris Doecke developed the process for continuous hydroformylation. Brian Haeberle, Philip Hoffman, Randy Lambertus, and Tim White developed the Ir-catalyzed homogeneous reductive amination process, and Declan Hurley, Niall G. Kerrigan, and Richard Spencer scaled it up to manufacturing. Brandon Reizman and Richard Cope were engineers for the thermal EE deprotection in superheated PFR, and Mike Frederick and Joel Calvin developed the chemistry. Brandon Reizman led the thermal EE deprotection scale up to GMP manufacturing at a CMO. Ed Deweese, Paul Milenbaugh, and Rick Spears of D and M Continuous Solutions constructed and operated most of the continuous reactor systems. Matt Yates was the lead chemist for development of the continuous Schotten-Baumann reaction. Brandon Reizman and Chris Polster were engineers for the intermittent flow stirred tank reactor for 2-phase Suzuki cross-coupling. Rick Spencer, Jeff Lewis, and Mike Heller scaled up the CSTRs in series for Barbier Grignard, quench, and neutralization. Wei-Ming Sun was the deltaV automation engineer for most of these projects. We thank Bret Huff for initiating, leading, and sponsoring the continuous reaction design and development work at Eli Lilly and Company.

References

Bergman TL, Incropera FP, DP DW, Lavine AS. Fundamentals of heat and mass transfer: Wiley; Hoboken, NJ, 2011.

Braden TM, Gonzalez MA, Jines AR, Johnson MD, Sun W-M, inventors; Eli Lilly And Company, assignee. Reactors and methods for processing reactants therein. United States patent application WO 2009023515 A2 Feb 19, 2009.

Braden TM, Johnson MD, Kopach ME, Groh JM, Spencer RD, Lewis J, et al. Development of a commercial flow Barbier process for a pharmaceutical intermediate. Org Process Res Dev. 2017;21(3):317–26.

Bridgman PW. A complete collection of thermodynamic formulas. Phys Rev. 1914;3(4):273.

Cole KP, Campbell BM, Forst MB, McClary Groh J, Hess M, Johnson MD, Miller RD, Mitchell D, Polster CS, Reizman BJ, Rosemeyer M. An automated intermittent flow approach to continuous Suzuki coupling. Org Process Res Dev. 2016;20(4):820–30.

Cole KP, Johnson MD, Laurila ME, Stout JR. An automated repeating batch with catalyst recycle approach to nitro group hydrogenolysis. React Chem Eng. 2017a;2:288–94.

Cole KP, Groh JM, Johnson MD, Burcham CL, Campbell BM, Diseroad WD, et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow cGMP conditions. Science. 2017b;356:1144–50.

Cole KP, et al. Small-volume continuous manufacturing of merestinib. Part 1. Process development and demonstration. Org Process Res Dev. 2019;23:858–69.

Fogler HS. Essentials of chemical reaction engineering: Pearson Education; Upper Saddle River, NJ, 2010.

Frederick MO, Calvin JR, Cope RF, LeTourneau ME, Lorenz KT, Johnson MD, et al. Development of an NH 4 cl-catalyzed ethoxy ethyl deprotection in flow for the synthesis of merestinib. Org Process Res Dev. 2015;19(10):1411–7.

Johnson MD, May SA, Haeberle BD, Lambertus GR, Pulley SR, Stout JR. Design and comparison of tubular and pipes-in-series continuous reactors for direct asymmetric reductive amination. Org Process Res Dev. 2016;20(7):1305–20.

Johnson MD, May SA, Kopach ME, Groh JM, Cole KP, Braden TM, Shankarraman V, Merritt JM. Chapter 16: Design and selection of continuous reactors for pharmaceutical manufacturing. In: am Ende DJ, am Ende MT, editors. Chemical engineering in the pharmaceutical industry. 2nd ed: Drug Substance/API; Hoboken, NJ, 2019.

Kopach ME, Roberts DJ, Johnson MD, Groh JM, Adler JJ, Schafer JP, Kobierski ME, Trankle WG. The continuous flow Barbier reaction: an improved environmental alternative to the Grignard reaction? Green Chem. 2012;14(5):1524–36.

Lange H, Carter CF, Hopkin MD, Burke A, Goode JG, Baxendale IR, et al. A breakthrough method for the accurate addition of reagents in multi-step segmented flow processing. Chem Sci. 2011;2(4):765–9.

Levenspiel OS. Chemical reaction engineering. An introduction to the design of chemical reactors. New York, NY: Wiley; 1962.

Lide DR, Kehiaian HV. CRC handbook of thermophysical and thermochemical data. Boca Raton, FL: CRC Press; 1994.

May SA, Johnson MD, Braden TM, Calvin JR, Haeberle BD, Jines AR, et al. Rapid development and scale-up of a 1 H -4-substituted imidazole intermediate enabled by chemistry in continuous plug flow reactors. Org Process Res Dev. 2012;16(5):982–1002.

May SA, Johnson MD, Buser JY, Campbell AN, Frank SA, Haeberle BD, et al. Development and manufacturing GMP scale up of a continuous Ir-catalyzed homogeneous reductive amination reaction. Org Process Res Dev. 2016;20(11):1870–98.

Reizman BJ, et al. Small-volume continuous manufacturing of merestinib. Part 2. Technology transfer and cGMP manufacturing. Org Process Res Dev. 2019;23(5):870–81.

Riddick JA, Bunger WB, Sakano TK. Organic solvents: physical properties and methods of purification. New York, NY: Wiley; 1986.
Schwolow S, Hollmann J, Schenkel B, Röder T. Application-oriented analysis of mixing performance in microreactors. Org Process Res Dev. 2012;16(9):1513–22.
White TD, Berglund KD, Groh JM, Johnson MD, Miller RD, Yates MH. Development of a continuous Schotten–Baumann route to an acyl sulfonamide. Org Process Res Dev. 2012;16(5):939–57.
Wiles C, Watts P. Continuous flow reactors, a tool for the modern synthetic chemist. Eur J Org Chem. 2008;10:1655–71.

Chapter 3
Understanding Residence Time, Residence Time Distribution, and Impact of Surge Vessels

Martin D. Johnson, Scott A. May, Jennifer Mc Clary Groh, Luke P. Webster, Vaidyaraman Shankarraman, Richard D. Spencer, Carla Vanesa Luciani, Christopher S. Polster, and Timothy Braden

Abstract Axial dispersion and residence time distribution (RTD) have significant process implications in the design of continuous processes. Low dispersion is better for minimizing time for complete reaction and minimizing deviation boundaries. High dispersion is better for dampening out process fluctuations or disturbances. In typical processes with surge vessels, axial dispersion may not have a significant impact in lot genealogy. Axial dispersion number (*D/uL*) and mean residence time (τ) can be calculated by fitting the analytical solution to axial dispersion model to experimental F-curve transition data. Numerical modeling tools can be used to quantify the dampening of disturbances (e.g., mass flow oscillations) by axial dispersion and reactor configuration. The analysis herein includes plug flow reactors (PFRs) with different values for axial dispersion number, and continuous stirred tank reactor (CSTR) followed by PFRs in series. For reactor design, RTD for actual product and reagents can be much different than RTD for bulk solvent flow, and therefore it must be measured experimentally for the actual substrate to quantify it precisely. Numerical predictions of RTD based on Peclet number can be highly inaccurate predictors of axial dispersion for larger PFRs, even for bulk flow solvents. τ in PFRs must be corrected for thermal expansion. For lot genealogy, surge vessels can be the major contributors to overall RTD spread. Surge vessels decouple operations, simplify automation, reduce the need for redundancy, allow for interme-

M. D. Johnson (✉) · S. A. May · J. M. C. Groh · L. P. Webster · V. Shankarraman
R. D. Spencer · C. V. Luciani · C. S. Polster · T. Braden
Eli Lilly and Company, Indianapolis, IN, USA
e-mail: johnson_martin_d@lilly.com; may_scott_a@lilly.com;
groh_jennifer_mcclary@lilly.com; webster_luke_p@lilly.com;
shankarraman_vaidyaraman@lilly.com; spencer_richard_d@lilly.com;
luciani_carla_vanesa@lilly.com; polster_christopher_s@lilly.com;
braden_timothy@lilly.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_3

diate quality assessment, dampen out process disturbances, and help eliminate start-up and shutdown transition waste. Numerical modeling tools can be used to quantify overall RTD for a fully continuous process with multiple unit operations linked in series, predicting how disturbances are dampened out downstream and determining lot genealogy.

Keywords Residence time · Residence time distribution · Continuous processing · Drug substance

Abbreviations

CSTR	Continuous stirred tank reactor
D or $D_{\_eff}$	Dispersion coefficient
D/uL	Axial dispersion number
i.d.	Inside diameter
L	Characteristic length
MSMPR	Mixed-suspension mixed-product removal
PAT	Process analytical technology
PFDR	Plug flow with dispersion reactor
PFR	Plug flow reactor
RI	Refractive index
RTD	Residence time distribution
SM	Starting material
u	Average velocity within the tube
τ	Mean residence time

3.1 Residence Time Distribution in PFRs Is a Consequence of Axial Dispersion

Axial dispersion occurs in laminar flow PFRs due to both diffusion and advection. Velocity profile is parabolic, a phenomenon that was first characterized by Taylor (Taylor 1953). In his investigation, Taylor made injections of a soluble tracer into a flowing stream and measured the dispersion of the tracer downstream. Taylor quantified and reported this dispersion via the vessel dispersion number, in the following equation, which accounts for the dispersive effects of both advection and diffusion. Vessel Dispersion Number = $D_{\_eff}/uL$, where $D_{\_eff}$ is the dispersion coefficient, u is the average velocity within the tube, and L is the characteristic length. Levenspiel applied vessel dispersion number ($D_{\_eff}/uL$ and hereafter simply called D/uL) to

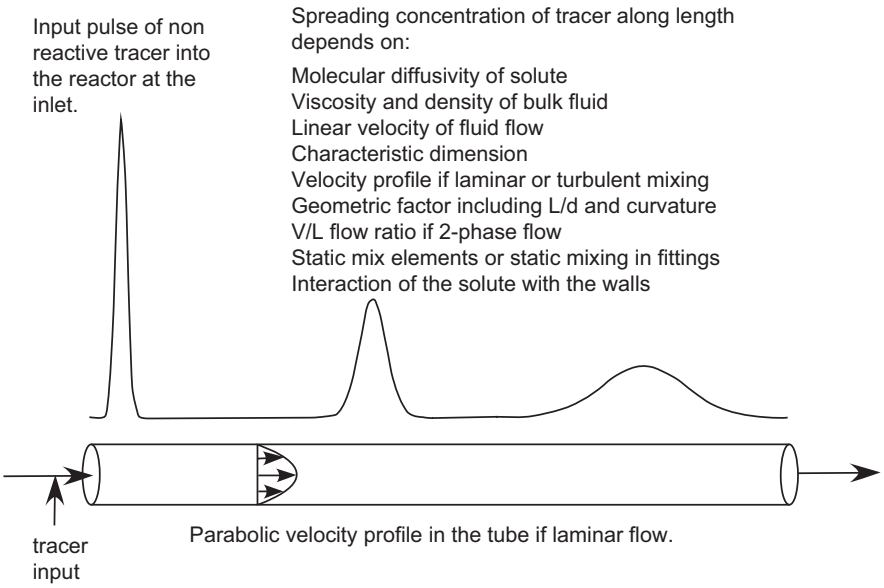


Fig. 3.1 C-curve illustrating a nonreactive pulse tracer spreading in concentration as it flows along the length of a PFR

nonideal continuous reactors (Levenspiel 1962). There are many literature sources on this topic (Levenspiel 1962, 1979; Fogler 1999; Weber Jr and DiGiano 1996; Aris 1956; Taylor 1953, 1954a, b). A perfect plug flow reactor would have a $D/uL = 0$, while a perfect continuous stirred tank reactor would have $D/uL \rightarrow \infty$. Using the guidance of Levenspiel, $D/uL = 0.025$ represents an intermediate amount of dispersion, while $D/uL = 0.002$ represents a low amount of dispersion (Levenspiel 1962). Figure 3.1 illustrates how a nonreactive pulse tracer spreads in concentration by forward mixing and back mixing as it flows along the length of a PFR. This is known as the Dankwerts C-curve (Danckwerts 1954).

Furthermore, plotting C/C_0 versus t/τ after a step change in nonreactive tracer flowing into a tube produces an s-shaped curve, where C is nonreactive tracer concentration at exit and C_0 is tracer concentration at the inlet. Dankwerts defined this as the F-curve (Danckwerts 1954). It is mathematically equal to the integral of the Dankwerts "C-curve," which is the C/C_0 response curve to a pulse injection of nonreactive tracer. The C-curve and the F-curve are functions of D/uL mathematically; therefore, either may be used to calculate D/uL. C-curves and F-curves for different values of D/uL are shown in Fig. 3.2.

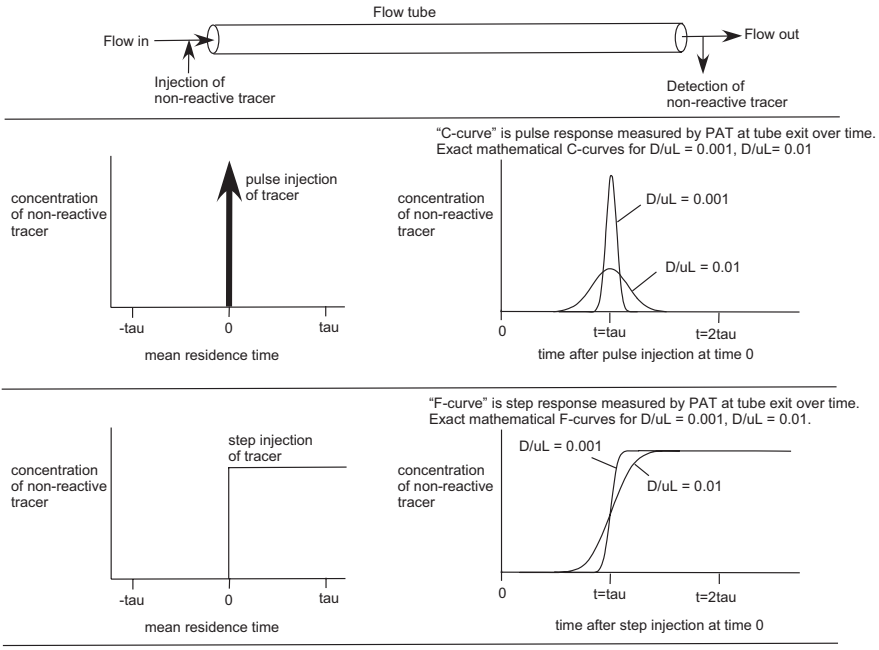


Fig. 3.2 C-curves and F-curves for different values of *D/uL*

3.2 Advantages and Disadvantages of Low Axial Dispersion

Lower axial dispersion is better for minimizing time to full conversion in PFRs for positive order reactions, minimizing lot-to-lot carryover, time to steady state, time for recovery after a process upset, and theoretical amount of material diverted as a result of a process upset. On the other hand, higher axial dispersion is better for dampening out oscillations and disturbances, which can help to prevent out-of-spec material in the first place. One approach is to design for the narrowest possible residence time distribution (RTD) in the entire system so that the amount of material diverted to waste is minimized in the event of a process disturbance. Another approach is to design for broad RTD so that the same disturbance may not lead to any nonconforming material, thus eliminating the need to divert any material. Both approaches are valid. The best design depends on the control strategy for each given product.

3.3 Quantifying Axial Dispersion in Plug Flow Tube Reactors

Using the plug flow with dispersion reactor (PFDR) model, *D/uL* is calculated by fitting experimental F-curve data to the basic differential equation representing dispersion, which is shown in the following equation (Levenspiel 1962):

$$\frac{\partial C}{\partial \theta} = \left(\frac{D}{uL}\right)\frac{\partial^2 C}{\partial z^2} - \frac{\partial C}{\partial z}$$

where D/uL = vessel dispersion number

C = concentration
θ = dimensionless time (t/τ), where τ is mean residence time
Z = dimensionless reactor length (x/L)

The equation is numerically solved for D/uL and θ. For instance, A BOC deprotection reaction was run in a 12 L coiled tube gas/liquid PFR (Cole et al. 2017). The start-up transition curve during the GMP manufacturing of prexasertib was measured in the plant by on-line refractive index (RI) probe. The following calculations illustrate why a PFR is superior to CSTRs-in-series for achieving narrow RTD. The dispersion model was fit to the experimental data to determine D/uL (Fig. 3.3). Calculations were done using a simple Excel® spreadsheet using the C-curve analytical solution given for nonreactive tracer in Levenspiel (Levenspiel 1962), and integrating over time.

$$C = \frac{1}{2\sqrt{\pi\left(\frac{D}{uL}\right)}}\exp\left[-\frac{(1-\theta)^2}{4\left(\frac{D}{uL}\right)}\right]$$

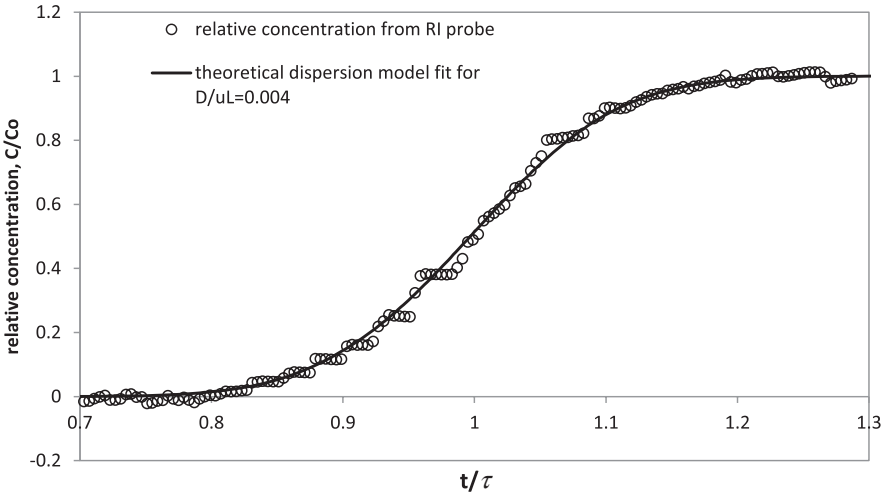


Fig. 3.3 Experimental F-curve reactor start-up data and dispersion model fit for the investigated BOC-deprotection reaction. C_o is the concentration entering the reactor when the reagent starts feeding at time = 0; therefore, at the new steady state $C/C_o = 1$

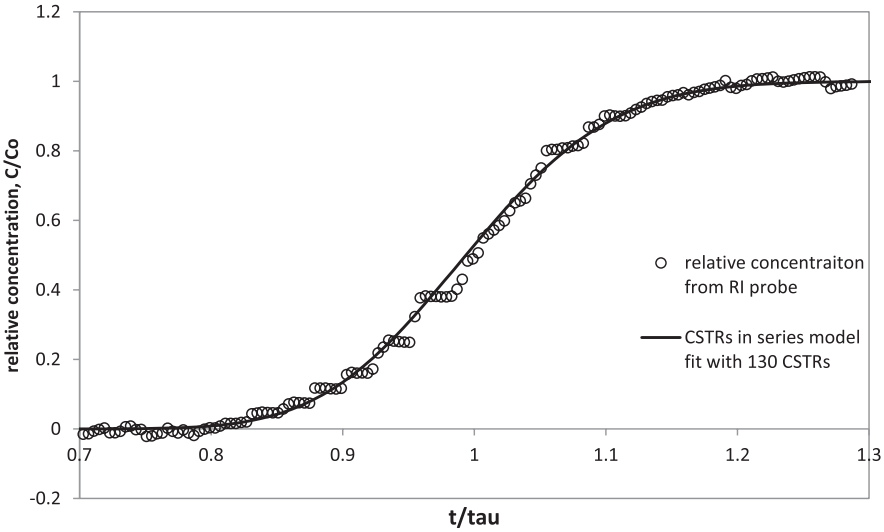


Fig. 3.4 Experimental F-curve reactor start-up data and CSTRs-in-series model fit

This analytical solution applies only for small extents of dispersion, $D/uL < 0.01$. The spreadsheet analysis determined that $D/uL = 0.004$. Practically speaking, this means that reactor effluent concentration was 99% of full strength after 1.2 reactor volume turnovers during start-up transition.

The same experimental F-curve data was also fit to the CSTRs-in-series reactor model. Process fluid volume in each theoretical CSTR unit is total reactor volume divided by the number of CSTRs. The E-curve represents the spread of nonreactive tracer as it moves along the reactor. The E-curve of CSTRs-in-series is mathematically equal to (Weber Jr and DiGiano 1996):

$$E(t) = \frac{e^{-t/\tau}}{\tau(n-1)!}\left(\frac{nt}{\tau}\right)^{n-1}$$

where n equals number of CSTRs-in-series of equal τ.

The F-curve can be generated because it is the time integral of the E-curve:

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

The model was fit to the data, to solve for n (Fig. 3.4). Calculations were done using a simple Excel® spreadsheet.

The model fit revealed that the RTD was the equivalent of 130 CSTRs-in-series, which is clearly not practical.

As the axial dispersion goes down, the equivalent number of CSTRs-in-series goes up. This can be seen in Fig. 3.5. A single CSTR has infinite axial dispersion.

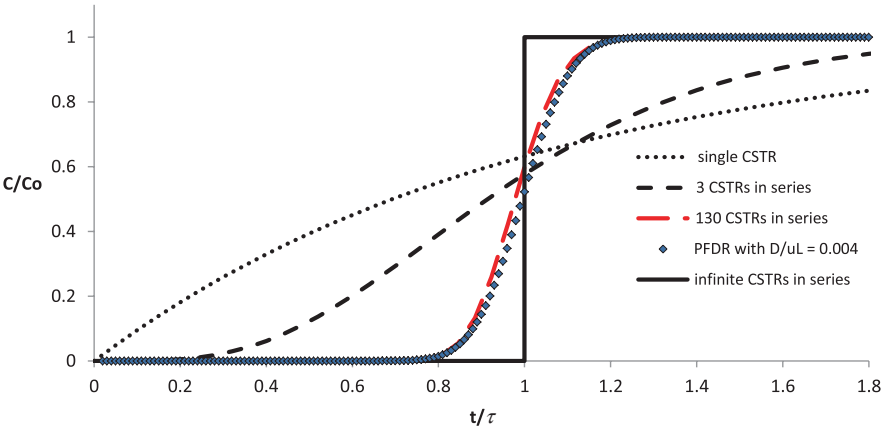


Fig. 3.5 RTD as a function of number of CSTRs-in-series. C_o is the concentration entering the reactor when the step change occurs at time = 0; therefore, at the new steady state $C/C_o = 1$

The single CSTR has RTD with exponential decay of starting concentration after a step change in feed composition ($C/Co = \exp(-t/\tau)$). With only three CSTRs-in-series, the RTD starts looking more like an F-curve. With 130 CSTRs-in-series, RTD is about the same as PFDR with axial dispersion number 0.004, as described above. A perfect plug flow reactor ($D/uL = 0$) corresponds to an infinite number of CSTRs-in-series of infinitesimally small volume.

3.4 Quantifying Impact of Axial Dispersion on Reactor Performance

As axial dispersion increases, so does the time required for desired reaction conversion. Fractional conversion of reactant A, as it passes through a tubular reactor, is governed by axial dispersion, advection, reaction rate constant (k), and reaction order (n), as shown in the following dimensionless form of the steady-state material balance equation for component A. (Levenspiel 1962)

$$\frac{D}{uL}\frac{d^2X_A}{dz^2} - \frac{dX_A}{dz} + k\tau C_{A0}^{n-1}\left(1-X_A\right)^n = \frac{dX_A}{dt} = 0$$

where X_A = Fractional conversion

C_{A0} = Initial concentration
z = Fraction of total reactor length (l/L)

Steady state means that the differential change in concentration with time equals zero in the material balance equation. Thus, the dX_A/dt term on the right-hand side

is zero. The first term represents flux of A in and out of the differential volume by dispersion; the second term represents flux of A in and out by bulk flow; and the third term represents change in concentration of A by reaction. The material balance equation was solved analytically for first-order reaction (Levenspiel 1962).

$$\frac{C_A}{C_{A0}} = 1 - X_A = \frac{4ae^{\left(\frac{1}{2}\frac{uL}{D}\right)}}{(1+a)^2 e^{\left(\frac{a}{2}\frac{uL}{D}\right)} - (1-a)^2 e^{\left(-\frac{a}{2}\frac{uL}{D}\right)}}$$

$$a = \sqrt{1 + 4k\tau\left(D/uL\right)}$$

where $k\tau$ = reaction Damkohler Number.

This is an analytical solution for a simplified scenario. For more complex mathematical scenarios, a more sophisticated numerical modeling tool like gPROMS can be utilized, as described elsewhere (Johnson et al. 2019).

Alternatively, a tubular reactor can be modeled as CSTRs-in-series. The reactor design equation for first-order irreversible reaction in equal-sized CSTRs-in-series is (Levenspiel 1962)

$$\frac{C_{An}}{C_{A0}} = \frac{1}{(1+k\tau)^n}$$

where C_{An} = concentration of reagent A in the nth CSTR-in-series.

n = number of CSTRs-in-series.

Both of these reactor models were used to calculate conversion versus time curves for a first-order reaction with reaction rate constant $k = 1.2$/h in the 12 L PFR from the previous example ($D/uL = 0.004$ or 130 CSTRs-in-series). The ideal PFR model was included for comparison, as shown in the next equation.

$$\frac{C_A}{C_{Ao}} = e^{-kt}$$

The results are shown in Fig. 3.6. Time in the reactor is the same as distance along the length of the reactor. The calculations were done with a simple Excel® spreadsheet because they used analytical solutions to the design equations. These curves practically lay on top of each other; therefore, there is not a significant conversion versus time difference between ideal PFR versus a tube reactor with dispersion number 0.004. This means that the 12 L reactor can be modeled as ideal plug flow with minimal error in conversion versus time prediction, because axial dispersion is sufficiently low.

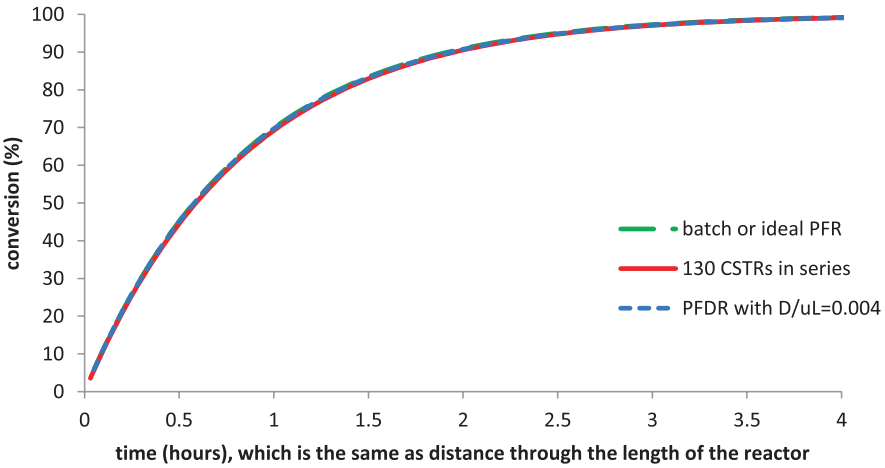


Fig. 3.6 Numerical models of conversion versus time (distance) for reaction in the 12 L PFR

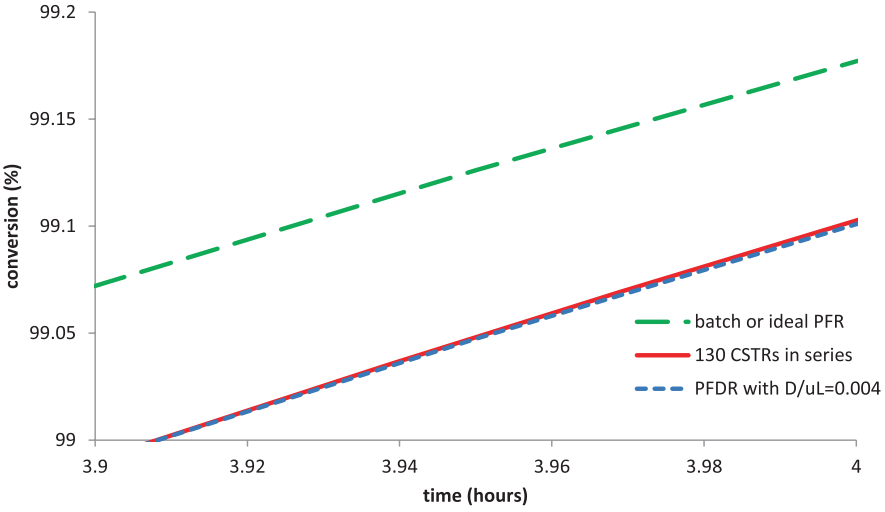


Fig. 3.7 Numerical models of conversion versus time (distance) for reaction in the 12 L PFR

Figure 3.7 shows the same plot but zoomed in on the end of reaction. Conversion versus time (distance) is slightly slower for a reactor with axial dispersion than for an ideal PFR. Model predictions show that this reaction would reach 99.17% conversion in an ideal PFR after 4 h, 99.10% in a PFDR with $D/uL = 0.004$ after 4 h, and 99.10% in 130 equal volume CSTRs-in-series after 4 h total time equally divided between the small CSTRs-in-series. For practical purposes, the reactor can be modeled as an ideal PFR when axial dispersion is this low.

3.5 Axial Dispersion in PFRs Dampens Out Flow Disturbances

Axial dispersion is not necessarily considered a negative reactor attribute. Higher axial dispersion is better for dampening out flow fluctuations, disruptions, or pulsations. A hypothetical example is used to illustrate this concept. The hypothetical catalytic reaction has the form

$$A + B \rightarrow C + B \qquad k = 0.99 \frac{L}{\text{mol} \cdot s}$$

where A is the substrate, B is the catalyst and C is the product. Suppose that the kinetics are such that 95% conversion of A is achieved in a PFR with inlet concentration of A at 1 M, 1% B (mole/mole relative to A at inlet) and residence time τ of 5 min. The PFR has separate feeds for substrate A and catalyst B. If catalyst solution flow rate oscillates, then concentration of catalyst entering the reactor oscillates up and down. Since B is a catalyst and used in small amounts relative to substrate, this will not significantly impact the residence time in the PFR.

The governing equation for PFR with dispersion is shown in the equation

$$\frac{\partial c_j}{\partial t} = -\frac{\partial c_j}{\partial z} + \left(\frac{D}{uL} \right) \frac{\partial^2 c_j}{\partial z^2} + \tau r_j$$

where

- z is dimensionless length along reactor (fraction of reactor length L),
- c_j is the concentration of species j at length z along the reactor
- $\frac{D}{uL}$ is the dispersion number
- r_j is the rate of reaction of species j, which in this example is

$$r_A = -kc_A c_B,\ r_B = 0,\ r_C = kc_A c_B$$

Initial condition is based on starting with no reactants in the reactor

$$c_j \left(t = 0, z \right) = 0$$

- t is dimensionless time (time divided by τ)

Boundary conditions impose flux balance at the inlet and outlet of the reactor $c_j^{in}\left(t\right)$ is the concentration at inlet to PFR and has the fluctuations for species B:

$$c_j^{in}\left(t\right) = c_j\left(t, z = 0\right) - \left(\frac{D}{uL} \right) \frac{\partial c_j}{\partial z}\left(t, z = 0\right)$$

$$\frac{\partial c_j}{\partial z}\left(t, z = 1\right) = 0$$

For a scenario where a CSTR precedes the PFR, $c_j^{in}(t)$ to PFR is the outlet concentration from CSTR, which is obtained by solving the following set of equations for CSTR:

$$\frac{dc_j}{dt} = c_j^{in} - c_j + \tau r_j$$

$$c_j(t=0) = 0$$

where

- $c_j^{in}(t)$ is the inlet concentration to CSTR (which has the fluctuations for B)
- c_j is the concentration of species j in CSTR
- r_j is the rate of reaction of species j

The simulation is implemented as starting the reactor with no reactants, only solvent. Reactants A and B flow into the reactor without any fluctuation for twice the residence time (to reach steady state in the absence of fluctuations). Then, a fluctuation in concentration of B is imposed. The impact of fluctuations is based on a 5 min snapshot closer to six τ (when all the initial transient phase has passed). This model was implemented in gPROMS Model Builder 4.0.0 (Process Systems Enterprise 1997); the graphs depicting the impact of fluctuations were created using ggplot2 (Wickham 2009).

In this example, the oscillation in concentration of B is represented using a sinusoidal function oscillating between 0 and two times inlet concentration of B with a frequency of 1 min, as shown in Fig. 3.8. This is a realistic scenario for a pulsating feed pump if the flow rate and pump frequency are low.

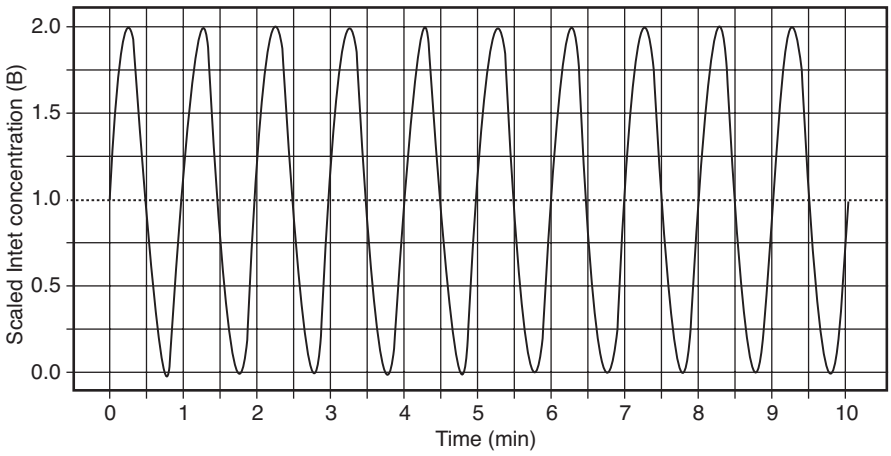


Fig. 3.8 Fluctuation in inlet concentration of B (scaled relative to steady-state concentration); sinusoid with frequency of 1 min

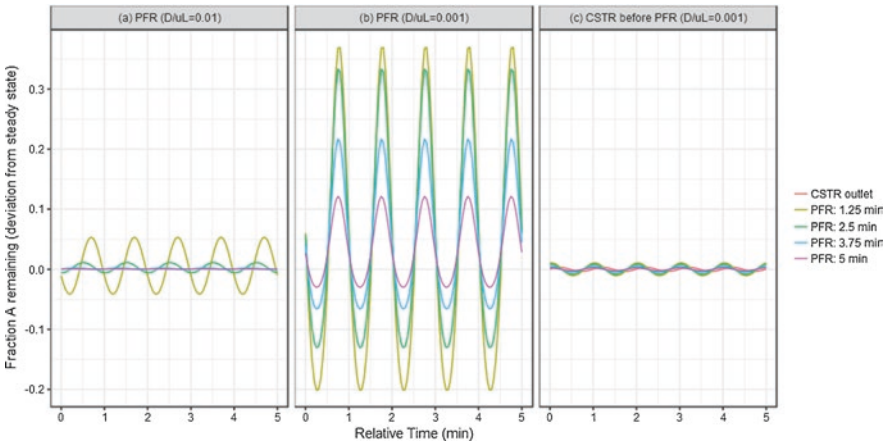


Fig. 3.9 The impact of dispersion number and reactor configuration, on dampening out process fluctuations

The impact of this fluctuation on the fraction *A* remaining at different residence times along the reactor is shown in Fig. 3.9 for three scenarios:

(a) PFR with dispersion number of 0.01 and residence time of 5 min.
(b) PFR with dispersion number of 0.001 (closer to plug flow) and residence time of 5 min.
(c) CSTR with a residence time of 2 min followed by a PFR (dispersion number of 0.001) with residence time of 5 min.

The steady-state conversion is different at different lengths (τ) along the reactor, and there is a time shift of fluctuations along the reactor. Therefore, Fig. 3.9 lists the fluctuations as deviations from steady-state conversion to make the y-axis comparable. Likewise, the time axis at different reactor lengths is appropriately shifted to make the x-axis comparable. Figure 3.9a shows that at $D/uL = 0.01$, there is not much of an impact of inlet fluctuations even halfway through the reactor (2.5 min residence time). With a PFR closer to plug flow ($D/uL = 0.001$, Fig. 3.9b), the impact of inlet fluctuations dampens out more slowly (notice the higher amplitudes at different residence times in Fig. 3.9b versus Fig. 3.9a). However, one way to dampen out inlet fluctuations when a PFR is closer to ideal plug flow is to include a CSTR with a small residence time prior to the PFR. Figure 3.9c shows a scenario where a 2 min residence time CSTR is introduced prior to the PFR. The introduction of the CSTR has a remarkable impact of dampening out fluctuations (Fig. 3.9c versus Fig. 3.9a, b) compared to a stand-alone PFR (with D/uL of 0.01 or 0.001). The reason for this is because the 2 min residence time CSTR considerably dampens out the concentration fluctuation in catalyst solution *B* which feeds into the PFR (shown in Fig. 3.10).

Understanding how axial dispersion dampens out fluctuations in feed mass flow rates helps us to select the appropriate time averaging for the mass flow measurement. In the above example, oscillations with 1 min frequency are dampened out in

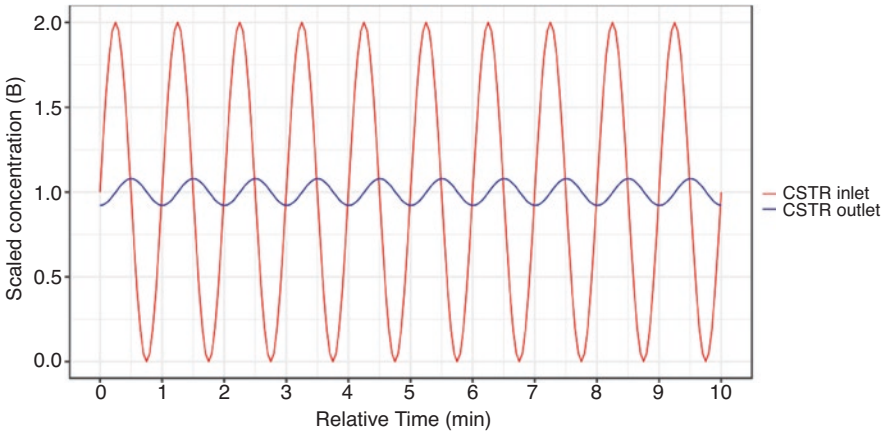


Fig. 3.10 The 2 min τ CSTR dampens out the concentration fluctuation in reagent *B*

a reactor system composed of a CSTR with 2 min τ in series with a PFR with 5 min τ. Therefore, it is appropriate to use 1 min time averaging on the measurement of mass flow rate of catalyst solution *B* pump. Subsequently, the standard deviation of pumping rate, which is used to set sigma values for pumping capabilities, can be based on the 1 min time averaged mass flow data from a flow meter, rather than instantaneous mass flow data. In other words, if two feeds mix in this CSTR, then the instantaneous mass flow rate can pulsate as long as the rolling average mass flow rate for 1 min is accurate. However, if the PFR is not preceded by a CSTR, then the appropriate time period for averaging the mass flow rate measurement is less.

3.6 Importance of High *L/d* When Scaling Up in the Laminar Flow Regime

If flow in a tubular reactor is in the laminar regime with Reynolds number (*Re*) in the range 0.1–2000, then *D/uL* increases when *Re* increases, assuming the reactor has fixed length to diameter ratio (*L/d)*. It may seem counterintuitive, but this relationship of *D/uL* and *Re* has been known and published in the literature and textbooks for many years, and it is shown graphically in Fig. 3.11 (Levenspiel 1962).

Axial dispersion was tested in seven coiled tube reactors ranging in size from about 700 ml to 8 L. The testing was done by making a step change from 100% THF to a 60%/40% (v/v) THF/toluene mixture flowing into the tube and monitoring the concentration of toluene in the reactor effluent by on-line Raman probe. (May et al. 2012) Data is shown in Table 3.1. A typical goal is to achieve $D/uL < 0.01$ in a PFR. Looking at the three smallest reactors (632–776 ml), it is clear that axial dispersion number increases as *L/d* decreases. The same is true for the medium-sized reactors (2.0–2.2 L) and the larger-sized reactors (7.1–8.7 L), especially in the

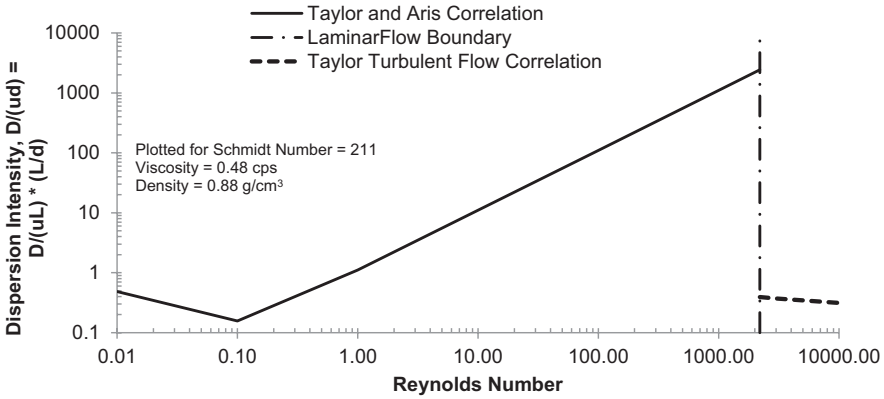


Fig. 3.11 Taylor and Aris correlations for dispersion intensity versus Reynolds number

Table 3.1 Axial dispersion versus tubular reactor size, *L/d*, and *Re*

	V (ml)	Inner *d* (mm)	*L/d* (length to inner *d* ratio)	*D/uL* (at τ = 90 min)	*D/uL* (at τ = 5 min)
632–776 mL	682	2.02	105,813	0.0003 (*Re* = 134)	0.0005 (*Re* = 2420)
	776	4.57	10,336	0.0013 (*Re* = 67)	0.0039 (*Re* = 1215)
	632	6.38	3106	0.0047 (*Re* = 40)	0.0049 (*Re* = 710)
2.0–2.2 L	2003	4.57	26,684	0.0016 (*Re* = 174)	0.0025 (*Re* = 3136)
	2185	9.55	3194	0.0091 (*Re* = 91)	0.0031 (*Re* = 1637)
7.1–8.7 L	7140	7.75	19,553	0.0008 (*Re* = 367)	0.0002 (*Re* = 6598)
	8726	15.7	2845	0.0385 (*Re* = 220)	0.0019 (*Re* = 3967)

laminar flow regime (τ = 90 min). Consider the three reactors with *L/d* about 3000 (3106, 3194, 2845). In these three examples, *L/d* about 3000 is sufficiently high for the 632 ml reactor with 6.38 mm inside diameter (*i.d.)*, since *D/uL* is less than 0.01 at both the high and the low flow rates. It is acceptable for the 2.2 L reactor with 9.55 mm *i.d.*, but *D/uL* is 0.009 in the lower flow rate scenario. However, dispersion results for *L/d* about 3000 are not acceptable for the 8.7 L reactor with 15.7 mm *i.d.*, because axial dispersion number is 0.0385 for the lower flow rate scenario. As expected, the data shows that higher *L/d* is required to maintain low axial dispersion when the reactor is scaled up to larger *i.d.*, and still in the laminar flow regime. On the other hand, the same 8.7 L reactor with 15.7 mm *i.d.* has very low axial dispersion (*D/uL* is 0.0019) when flow rate is increased enough so that flow is in the turbulent regime (*Re* = 3967). The key takeaway is that one must be especially cautious with large reactors, large *i.d.*, long residence times, operating in the laminar flow regime, because this can result in high axial dispersion. Another key takeaway is that you cannot scale up a tubular reactor by keeping L/d constant, if flow is laminar.

3.7 Calculated Peclet Number Does Not Correlate Well with Measured Axial Dispersion

The Peclet number does not correlate well with measured axial dispersion for tubing with high *L/d* and diameter greater than about 1 mm in the laminar flow regime. This is an important fact, because it describes the situation for most of the continuous flow tube reactors that have run in GMP production at Eli Lilly (please see the first seven entries in Table 1 in another chapter in this book) (Johnson et al. 2020). Due to notation, the vessel dispersion number appears to be identical to the inverse of the Peclet number. This is a common misunderstanding. The Peclet number is given by the equation:

$$Pe_L = uL / D_m$$

where Pe_L is the Peclet number, the ratio of convective to diffusive transport; u is the average flow velocity; L is the characteristic length; and D_m is the mass diffusion coefficient.

However, it should be noted that the dispersion coefficient (D_{eff}, also called D earlier in this chapter) used in the vessel dispersion number is several orders of magnitude larger than a mass diffusion coefficient (D_m). In order to reconcile this larger dispersion coefficient with traditionally measured mass diffusion, Taylor proposed the correction given by the following equation:

$$D_{\text{eff}} = D\left(1 + \frac{1}{192} Pe_d^2\right)$$

where Pe_d is the Peclet number using the tubing diameter as characteristic length for the vessel.

Use of this correlation would allow for prediction of a vessel dispersion number from a mass diffusion coefficient, reactor dimensions, and flow rate. In order to confirm the performance of Taylor's correction, it was compared to vessel dispersion numbers fit by least squares regression to experimental transition data.

Experimentally measured vessel dispersion numbers have been presented for tubular reactors of several scales. See Table 10 in the publication on thermal imidazole cyclizations in coiled tubes (May et al. 2012). In the discussion that follows, the experimentally measured axial dispersion numbers listed in Table 10 of the publication were compared to the Taylor-corrected Peclet number correlation and the noncorrected inverse Peclet number based on molecular diffusion. The Taylor-corrected inverse Peclet numbers agreed well with the experimental dispersion numbers provided the tubing *i.d.* was less than 1 mm. In cases where the reactor *i.d.* was >1 mm, the correlation matched experimental data poorly. Table 3.2 shows the cases with <1 mm *i.d.* tubing, where agreement was found. This would also be the case for most micro-reactors, because characteristic dimension of micro reactors is <1 mm. The full data set, highlighting lack of correlation in larger diameter tubing

Table 3.2 Vessel dispersion numbers fit to experimental data, compared to vessel dispersion numbers predicted by Taylor's correction for micro-reactors in highly laminar regimes

id	L	Vessel dispersion experimental	Taylor predicted
mm	m	$\tau = 90$ min	$\tau = 90$ min
0.56	7.14	0.0003	0.0003
0.56	9.97	0.0002	0.0003
0.56	18.4	0.0004	0.0003
0.56	83.6	0.0002	0.0003

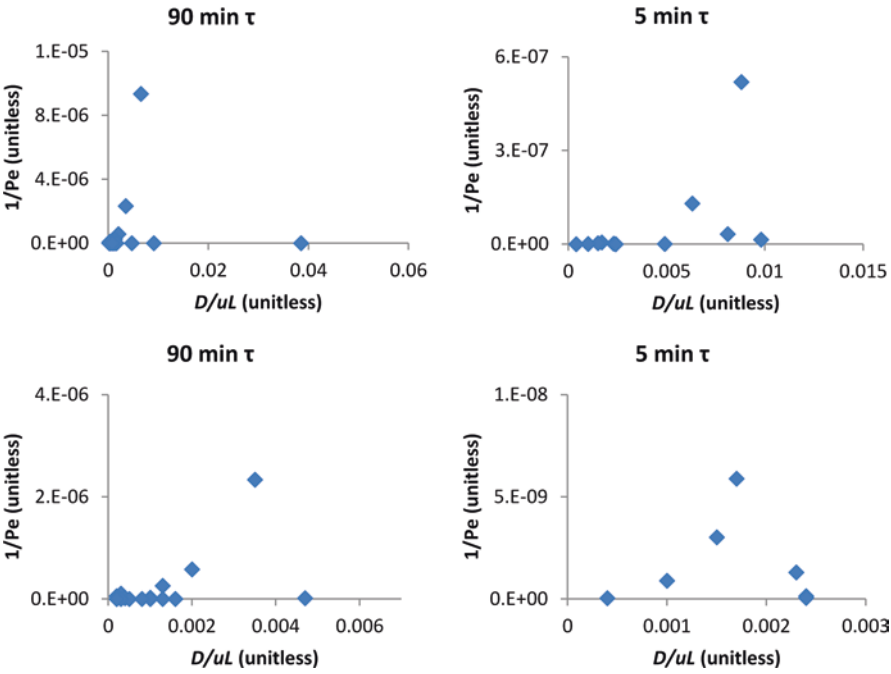


Fig. 3.12 Relationship between the inverse Peclet number and measured Axial dispersion. The top left shows data for 90 min τ and the bottom left shows a local view near the origin. The top right shows data for 5 min τ (where $Re < 1000$) and the bottom right shows a local view near the origin

and at shorter residence times, is displayed in Figs. 3.12 and 3.13. If the correlation between $1/Pe_L$ and D/uL was strong, then each of the plots in the figures would be a straight line with slope = 1. It is clear from the data in the figures that the correlation was poor.

Except for micro-reactors, the agreement between calculated prediction and experimental measurements was poor. A particular case will be highlighted here; a 7.1 L reactor, which was used for a thermal imidazole cyclization reaction in a cGMP API manufacturing campaign (May et al. 2012). The reactor was 7.75 mm *i.d.* and 151 m long. In this study, the experimental data for $V/Q = 90$ min was fit by least squares regression to give a vessel dispersion number of 0.0008, which

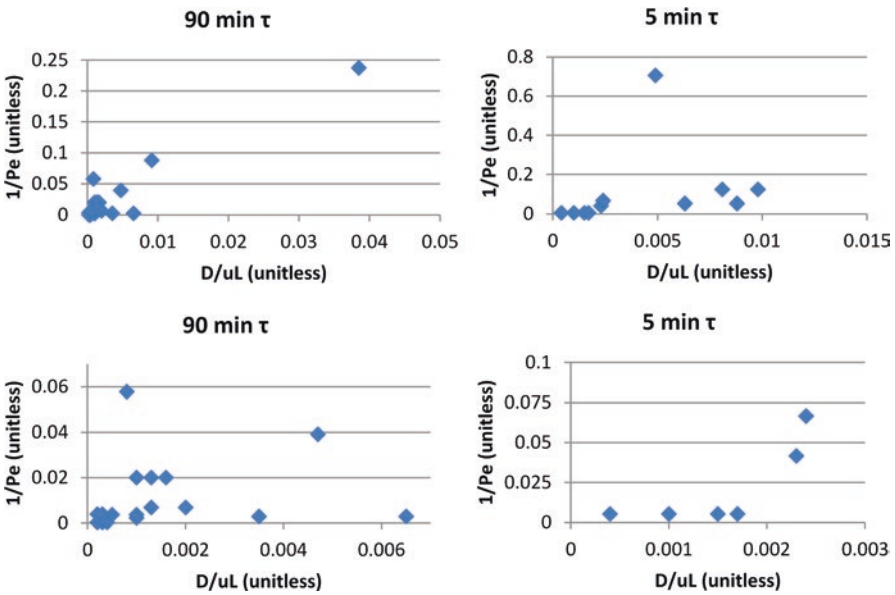


Fig. 3.13 Relationship between the Taylor corrected inverse Peclet number and measured Axial dispersion. The top left shows data for 90 min τ and the bottom left shows a local view near the origin. The top right shows data for 5 min τ (where $Re < 1000$) and the bottom right shows a local view near the origin

corresponds to a low degree of axial dispersion. When using a diffusion coefficient, the inverse Peclet number is $2*10^{-10}$. This small inverse Peclet number corresponds to near ideal plug flow. However, it is different than measured *D/uL* value by more than 5 orders of magnitude. Using Taylor's correction, the predicted vessel dispersion number is 0.06, which corresponds to a high amount of axial dispersion. In order to obtain a stream of 99.9% concentration at reactor start-up, these three vessel dispersion numbers would require 1.12, 1.00, and 2.03 volume turnovers, respectively. This is significant because each reactor volume turnover is almost 1.5 h. The experimental data and least squares fit are plotted along with the ideal plug flow predicted by the inverse Peclet number and the nonideal flow predicted by the Taylor correction in Fig. 3.14. It should be noted that the assumptions made in the Danckwerts' boundary conditions, which are used to analytically solve the dispersion model, collapse at dispersion numbers larger than 0.01. Therefore, the Taylor predicted F-curve shown is modeled as 14 ideal CSTRs-in-series to give a similar number of system turnovers for 99.9% of the transition as the predicted vessel dispersion number

Other methods, in addition to those presented here, exist for predicting axial dispersion (Edwards and Richardson 1970; Lemoullec et al. 2008). However, no single approach has fully captured all of the chemical, flow, and reactor effects that lead to the phenomenon of axial dispersion. In order to fully characterize a tubular reactor, its residence time distribution, and the impact of axial dispersion on

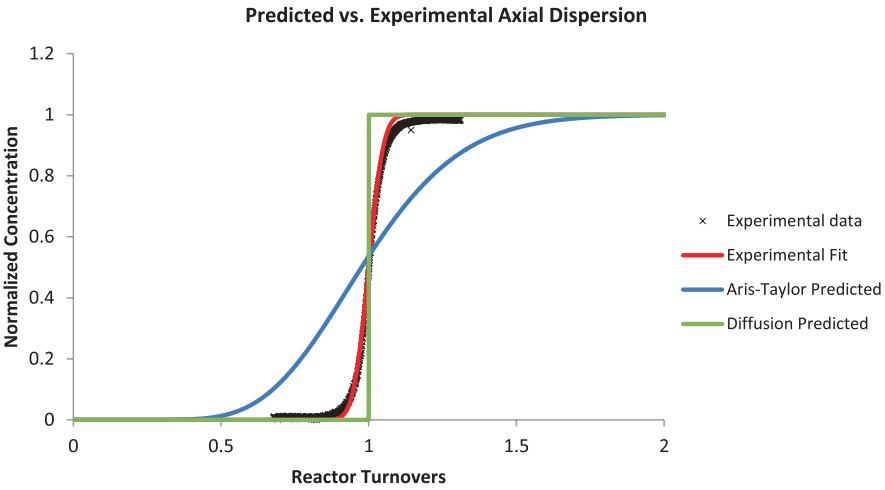


Fig. 3.14 Measured, predicted, and model fits to F-curves for a plug flow reactor, showing how poor the Peclet number predictions match reality for a 7.75 mm *i.d.* PFR in the laminar flow regime

reaction conversion, the best approach may be to fit vessel dispersion number to experimental data, measured after a step change or pulse change using representative process fluids and the actual substrate.

3.8 Quantifying *D/uL* with Solvents and Nonreactive Tracers Is Not Always Representative of *D/uL* for the Real Substrate in the Real Process

Start-up transition during the actual production campaign is one of the best opportunities to accurately quantify axial dispersion. The PFR is initially filled with solvent before the start of production. At time zero, reagents begin flowing into the PFR. Sampling or PAT analysis determines product concentration versus time at the reactor outlet for the first 1.5–2 reactor volume turnovers, thus generating the F-curve data points. The F-curve data points are fit numerically to the theoretical dispersion model to quantify the real dispersion number for the real product molecule. A consequence of this methodology is that axial dispersion quantification does not occur during equipment qualification. Quantifying dispersion number with solvents and nonreactive tracers during equipment qualification is still informative, but it will not necessarily be representative of the real substrate in the real process.

This fact was exemplified in a GMP continuous production campaign to manufacture prexasertib (Cole et al. 2017). A SNAr reaction was run in a 2.9 L coiled tube reactor at 70 °C temperature and 3 h mean residence time τ. The PFR was made with 91.5 m long 6.35 mm *i.d.* perfluoroalkoxy (PFA) tubing. First, nonreactive

solvents were used for F-curve measurements, switching from MeTHF to xylenes. The time to steady state was 3.3 h, and axial dispersion number was 0.0005. However, when the F-curve was subsequently measured for the actual pyrazole reaction product in DMSO during the start-up transition of the real process, the time to steady state was 4.8 h and the real axial dispersion number was 0.012. Mean residence time was about 3 h for both the solvent test and the actual reaction process, therefore the target 3 h reaction time was achieved. Data for the 2 F-curves is shown in Fig. 3.15. The nonideality can be seen in Fig. 3.16, which shows the same data mathematically converted to C-curve transitions. By definition, the F-curve is the integral of the C-curve. Therefore, the trends in Fig. 3.16 were generated by plotting the derivative of the F-curve data shown in Fig. 3.15. The C-curve for xylenes is normally distributed around the mean, but the C-curve for reaction product in

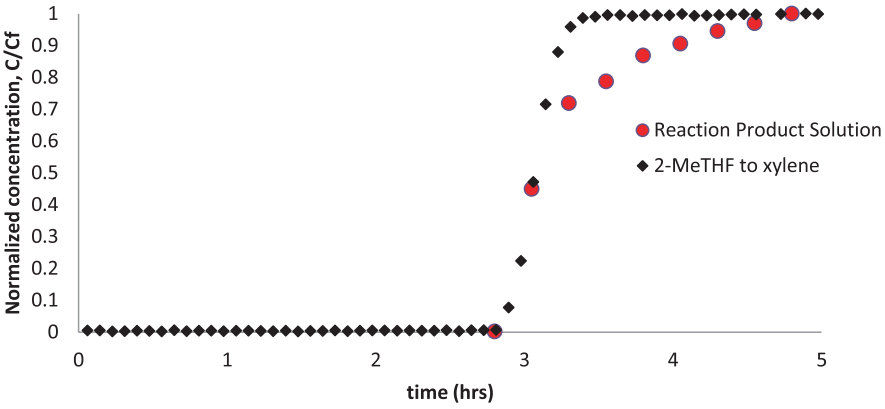


Fig. 3.15 F-curve transitions showing that RTD for reaction product is different than for solvent-only testing in this example

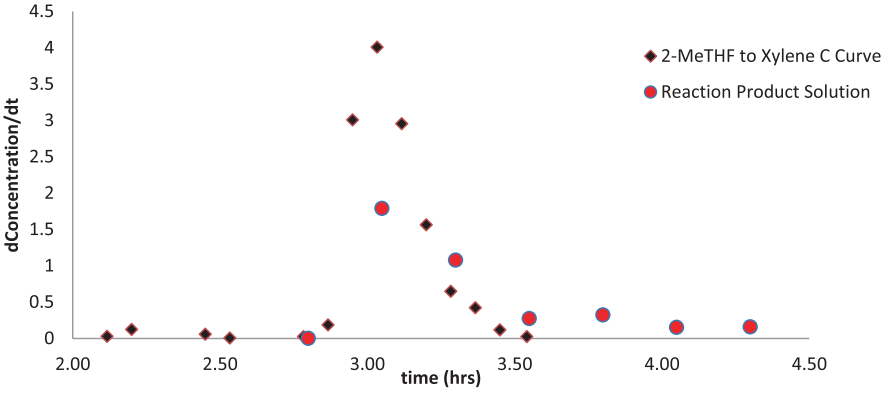


Fig. 3.16 C-curve transitions showing that RTD for reaction product is different than for solvent-only testing in this example

DMSO is not normally distributed. This is seen in the highly nonsymmetric C-curve in Fig. 3.16. The important message from this example is that measuring RTD with a nonreactive tracer is not always representative of the RTD of the actual product in the actual reaction mixture.

3.9 τ Versus V/Q and the Impact of Thermal Expansion

τ is defined as mean residence time. It depends on thermal expansion of the fluid in the reactor. A common mistake is to divide the volume of liquid in the reactor (V) by volumetric pumping rate at ambient temperature (Q) and declare that it equals reaction time. This is usually not correct. Reaction time must be corrected for thermal expansion of the fluid in the reactor. Comparing reaction time batch to reaction time flow necessitates quantification of τ, the real mean residence time, and not V/Q. Likewise, measuring reaction kinetics in a flow reactor necessitates quantification of τ. However, τ is not always known and it is not always easily measured. V/Q, on the other hand, should always be known, and it is a more practical value to list in the manufacturing instructions. Q is the flow rate set points for the pumps, which is measured by mass flow meters and known densities, and reactor volume V should be known. Therefore, a best practice is to report both τ and V/Q together. For example, consider a thermal deprotection that ran in a PFR at 270 °C and 9.4 min τ (May et al. 2012). The solvent was THF, which has significant thermal expansion at reaction temperature 270 °C. Reaction volume was 4.51 mL, and volumetric pumping rate at room temperature was 0.284 mL/min; therefore, V/Q was 15.9 min. Suppose the researchers really intended to achieve 15.9 min reaction time, so they set the pump flow rate to 0.284 mL/min. The problem is that the actual reaction time would only be 9.4 min, which is only 59% of the desired τ. For this reason, good practice is to describe the reaction as follows: "V/Q = 15.9 min and τ = 9.4 min because of thermal expansion." The severe thermal expansion can be seen in Fig. 3.17. In this

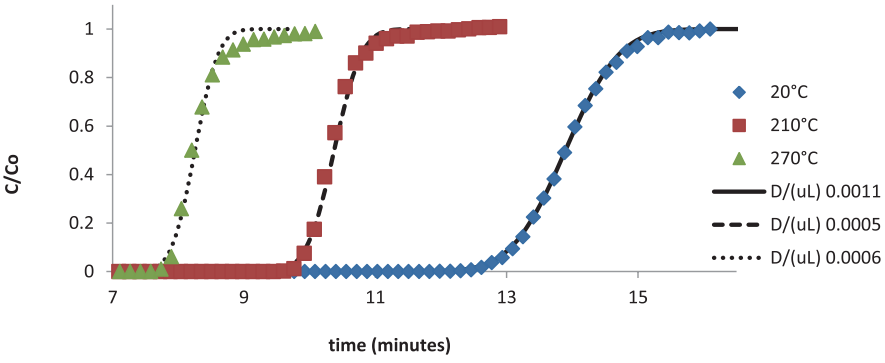


Fig. 3.17 F-curve transition curves showing the impact of thermal expansion on τ. (Reprinted with permission from May et al. (2012). Copyright (2012) American Chemical Society)

example, density at 270 °C is 0.53 g/mL and density at 210 °C is 0.69 g/mL, compared to 0.896 g/mL at 20 °C. For the same pumping rate, at 20 °C reactor temperature τ = 15.9 min, at 210 °C the thermal expansion of the liquid phase results in τ = 12.2 min, and at 270 °C τ = 9.4 min, with all else held constant except reactor temperature. The real τ for a continuous reaction like this with significant thermal expansion is best measured during start-up transition by measure the experimental F-curve

3.10 Overall RTD Contributors and Dampen Out Process Disturbances

While understanding the RTD of single unit operations is invaluable for equipment design, it is also necessary to understand the RTD of the complete process since it will determine how disturbances propagate, batch genealogy, and diversion strategy. Figure 3.18 illustrates a train consisting of a series of reactions and purification steps that operate continuously to produce a pharmaceutical active ingredient. Starting material 1 undergoes a protection reaction in PFR1, followed by a solvent exchange. The resulting protected intermediate is pumped to a surge can (dual surge can 1). As the amount of accumulated material in the surge vessel reaches the desired level, the dual surge can is discharged continuously to feed a series of two CSTRs (CSTR1 and CSTR2) and a coupling reaction (PFR2). The protected coupled material leaving PFR2 is quenched in CSTR3 and deprotected in CSTR4. The crude active ingredient is initially accumulated in a second surge vessel (dual surge can 2) until the desired level is reached. Then, the final purification step is fed. The resulting active ingredient is purified through a series of liquid–liquid extractions and the active pharmaceutical ingredient is accumulated in drums waiting for a final continuous crystallization step. In order to understand the potential dynamic response of the aforementioned train to disturbances, a mathematical model was built to describe the system overall RTD (gPROMs Model Builder). Table 3.3 lists the main RTD characteristics parameters of the pieces of equipment in the train listed above.

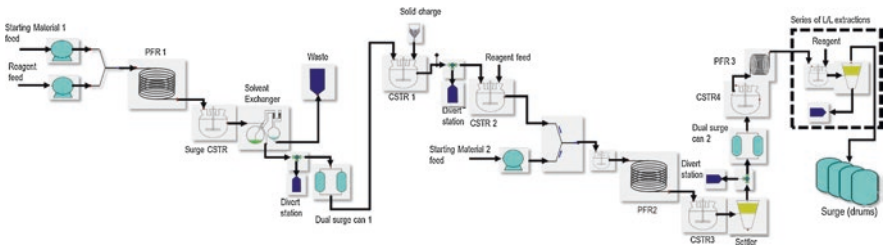


Fig. 3.18 Investigated continuous train

Table 3.3 Residence time distribution characteristic parameters for the investigated train

Unit operation	Dispersion number	Mean residence time (h)	Operation mode
PFR1	0.001	1.5	Continuous
Surge CSTR	∞	2.0[a]	Intermittent
Solvent exchanger	NA	3.0[a]	Intermittent
Dual surge cans 1	NA	25.0[b]	Fill-Empty
CSTR1	∞	3.0	Continuous
CSTR2	∞	0.5	Continuous
PFR2	0.001	1.5	Continuous
CSTR3	∞	0.5	Continuous
CSTR4	∞	0.5	Continuous
Dual surge cans 2	NA	7.0[b]	Fill-Empty
PFR3	0.001	0.3	Continuous
Liquid/liquid extractions in stirred tanks	∞	1.0	Continuous
Surge drums	NA	16.0[c]	Accumulation

[a]Equipment operates intermittently, and the residence time represents the time worth of material accumulated under standard operating conditions
[b]Dual surge cans operate in a fill–empty schedule and the residence time represents the time worth of material accumulated under standard operating conditions
[c]Surge drums operate in a fill-mode and the residence time represents the time worth of material accumulated the time the drum is replaced

Figure 3.19 exemplifies the propagation of a pulse-shaped disturbance (virtual tracer) with a duration of 1 h and a magnitude of 10 wt.% occurring at outlet of PFR1 (Fig. 3.19, inset). A few observations can be made:

- The average residence time of the train, determined by the material entering surge drums, is around 2 days. Long average residence times are a typical feature of drug substance continuous processes that have implications not only on PAT location and frequency, but also on start-up and shutdown strategies. Long residence times and low concentrations required not as frequent but highly sensitive analytical techniques. Due to long start-up and shutdown processes, unintended stoppages of the complete train are to be avoided. This is one of the main purposes of intermediate surge vessels that allow for decoupling of individual process sections. More details on the use of surge vessels can be found in the next section.
- While the severity of the simulated disturbance is high at its source (PFR1), its magnitude reduces three orders of magnitude at the point the tracer is accumulated in a drum. The dampening continues as the disturbance moves across the train, reaching the final drum with a maximum concentration of 0.025 wt.%.
- The shape of the RTD is primarily determined by the cycles of the surge cans due to the large amount of material accumulated in those vessels compared to the overall amount of material accumulated in the train.

Figure 3.20 shows the simulated composition of surge drums. The impact of the simulated disturbance can be seen in drums 3, 4, and 5, with only traces of the tracer observed in drums 2 and 6. This type of simulation is extremely useful to map the

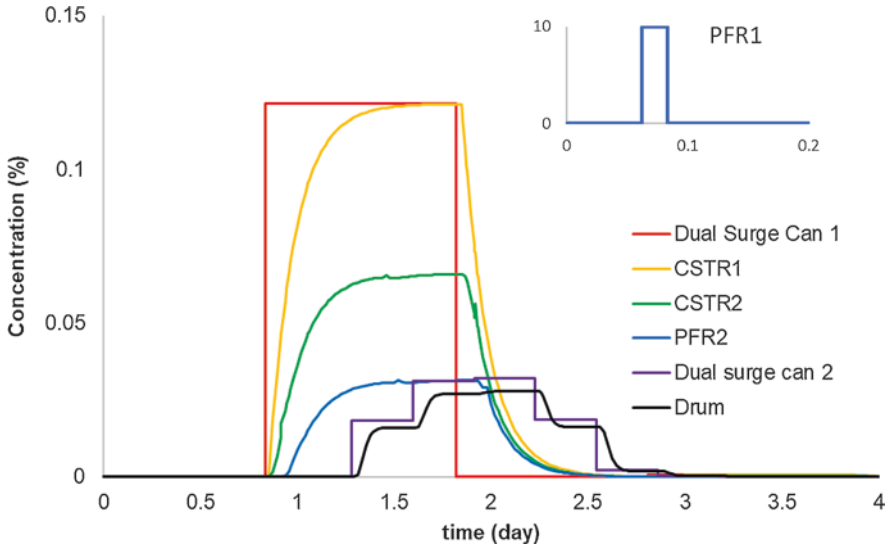


Fig. 3.19 Dynamic response to a pulse input at the outlet of PFR1 at t = 1.5 h, duration = 1 h, magnitude = 10 wt.%

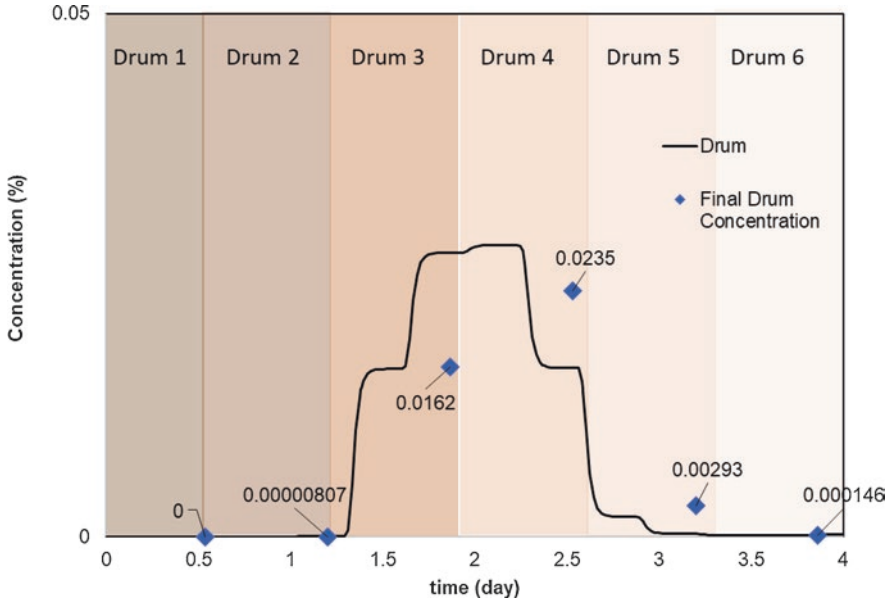


Fig. 3.20 Local composition of surge drums (line) and final drum concentration of tracer (symbol)

impact of magnitude/duration of disturbances as well as the need to divert potentially nonconforming material before it reaches the drum, but it can also be used to determine batch genealogy of complex processes/schedules. Intermediate surge

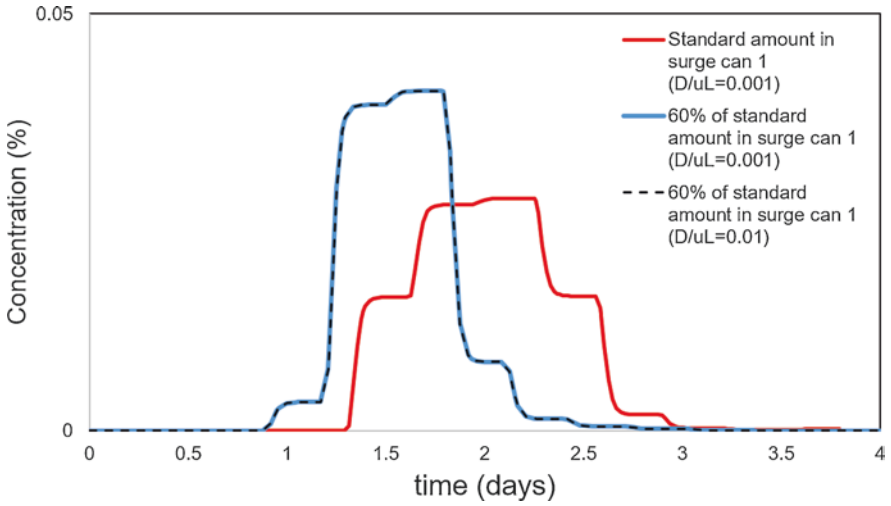


Fig. 3.21 Impact of dual surge can volume and PFR axial dispersion on overall RTD (drum input concentration)

vessels are designed to accommodate process needs. In case of stoppages or reduced flow rate upstream, the amount of material in the surge vessel will reduce over time. Stoppages or process slow-downs downstream will increase the amount of material accumulated in the surge vessel. Figure 3.21 shows the significant impact on the final RTD of reducing the volume accumulated in dual surge can 1. It also shows the negligible impact on overall RTD of changing all PFR dispersion numbers 1 order of magnitude. Axial dispersion number 0.01 versus 0.001 may be important for conversion versus distance in the reactor, but in this example it is not important for overall RTD and lot genealogy determinations.

3.11 Advantages and Disadvantages of Surge and Stock Vessels

In an efficient continuous plant, parallel stock vessels are used between reagent solution makeup tanks and the continuous reactor. These serve as reagent solution feed tanks. They are often used in parallel so that the off-line stock tank can be filled while the on-line stock tank feeds the continuous process, and they switch back and forth each time one becomes completely empty. Solid reagents are first dissolved into solvents in temperature-controlled, agitated vessels, and then the solutions are transferred to the simpler stock tanks that feed directly into the continuous reactors. This way, the agitated vessel is free to make more solutions, so no flow interruptions occur. The agitated vessels that make up feed solutions are generally more complex and more expensive because they have contained solids charge ports, agitation, and

heating/cooling capability. In contrast, the stock tanks may only need to hold homogeneous solutions at room temperature under slight positive pressure (e.g., 1 barg). The solutions are typically filtered while they are pushed or pumped from the solution makeup tank to the feed stock tank. It is operationally simplest if the stock tank contents are homogeneous solutions at room temperature with at least 1-week stability.

Furthermore, two or three parallel surge vessels between unit operations throughout the continuous train offer many advantages in a GMP manufacturing environment:

1. Decoupling the unit operations.
2. Simplifying operational logistics.
3. Simplifying start-up and shutdown transitions.
4. Simplifying automation and control.
5. Allowing brief stoppages of individual unit operations for troubleshooting or scheduled cleanouts of crystallizers or slurry reactors.
6. Reducing the need for in-line redundancy of things like on-line process analytical technology (PAT).
7. Providing distinct points where the quality of the material can be assessed and forward processed or diverted (three surge vessels in parallel facilitates off-line analytical).
8. Dampening out process disturbances and facilitating conservative calculations of acceptable magnitude and duration of disturbances.
9. Eliminating start-up and shutdown transition waste in some circumstances.
10. Simplifying investigation of the impact of procedural deviations.

Decoupling the unit operations means that one unit operation can be stopped for troubleshooting without needing to stop upstream or downstream reactors or separations unit operations. Operational logistics are simplified and number of operating staff required is reduced during troubleshooting events if the surge vessels are strategically used. The sections of the continuous processing train that are running well can be left alone while troubleshooting a single unit operation. Start-up and shutdown transitions are easier and simpler because the entire continuous train does not need to be started at the same time. The first reactor or the first section of the flow train is started and the product solution flows into surge vessels. After this section of the flow train is lined out at steady state and after a surge vessel is partially filled, the operations staff can move on to start up the next unit operation(s) in series in the flow train. This also facilitates planned start times of each section of the flow train, so that the same small group of experts can start each section but get some sleep between sections. If multiple synthetic route steps are started in series, then a sample from the surge vessel can be taken to the lab to run the use test for the next step before the downstream sections of the continuous train are started up. More frequent switching back and forth between parallel surge vessels minimizes the amount of material at risk, but increases the analytical frequency and decreases dampening of process disturbances. Higher frequency switching also increases the operator activity, unless the switching back and forth between surge vessels is automated.

Automating the back and forth switching of parallel surge vessels is strongly recommended because it also is better for minimizing heel volume. Minimizing the heels when switching back and forth greatly reduces overall RTD for the process train and reduces deviation boundaries. If the contents are homogeneous solutions at room temperature, then the surge vessels can be located a distance away from the reactors and separations unit operations in a "tank farm" area.

Should a particular pair of parallel surge vessels be operated close to full capacity or at low capacity? If it is more likely that the upstream sections of the train will have problems and need to be shut down periodically, then the surge vessels should be operating at relatively high capacity, so that downstream sections of the continuous train will be able to continue to run when upstream is shut down. On the other hand, if it is more likely that the downstream sections of the train will have problems and need to be shut down periodically, then the surge vessels should be operating at relatively low capacity, so that upstream sections of the continuous train will be able to continue to run while downstream is temporarily shut down. If neither, then 50% of full capacity is a good choice, so that there is equal decoupling in both upstream and downstream directions.

Furthermore, if it is a high priority to keep a particular section of the process train running without shutdown, then the fill level strategy will be slightly different. For example, if continuous reactions in series have an unstable intermediate, then it will be best to minimize the potential for shutdown of that section. Upstream and downstream sections of the continuous processing train may be more suitable for stable hold. If this is the case, then parallel surge vessels upstream from that section should be operated at higher fill levels, while parallel surge vessels downstream from that section should be operated at lower fill levels.

If one section of the flow train between surge vessels needs to "catch up" because it was temporarily shut down, then the other sections can be deliberately slowed down, for example, 80% of normal flow rates, to allow the volumes in surge vessels to be restored to original target operating levels. Therefore it is desirable to design operating ranges into the continuous process so that it is not a problem to temporarily increase mean residence time in each reactor or separations unit operation. The overall benefit is that the rest of the process train can continue uninterrupted if one unit operation goes down for maintenance, and subsequently the surge vessels can be rebalanced afterward. This is especially important for fully continuous multistep processes with a large number of continuous unit operations. It is also desirable to design the process with stable holds in as much of the flow train as possible.

Automation and control can be much simpler when sections of the flow train are decoupled via surge vessels. Otherwise, multiple levels of automation loops may be required. In this case, a stoppage of one unit operation would impact control of the adjacent continuous unit operations. If the unit operations are decoupled, then each section of the continuous system can be automated individually without the need for nested control loops, and it is easier to troubleshoot individual reactors or separations unit operations of the system without impacting the other reaction steps. This is especially beneficial the first time a new continuous process will run in

manufacturing, before "working out the bugs." Once a continuous process is well established and the entire flow train is functioning properly, some of the parallel surge vessels can be removed, the process can operate with less decoupling, and the nested control loops and more complex control programs applied. Decoupling with surge vessels is like starting a new continuous process with "training wheels" until all the reactions and separations unit operations are proven and reliable.

In addition, surge vessels allow for scheduled cleanouts of crystallizers or slurry reactors, which are prone to encrustation over time. For example, you may want to pause continuous crystallization 12 h for cleanout once every 200 h, and then continue running, without stopping the rest of the flow train.

The need for in-line redundancy of things like on-line PAT is reduced if you use surge vessels. For instance, what if there is no surge in the continuous train, the process relies on PAT, and the PAT stops working? Backup on-line PAT is one option to allow the process to continue to run without stopping. Alternatively, samples can be taken manually from a downstream surge vessel for off-line analysis in the event that the on-line PAT fails. This is a business decision because it puts more material at risk, but does not force a shutdown in the event that the PAT is temporarily not working.

Quality is an important driver for the use of surge vessels. These surge vessels simplify the process and offer distinct points where the quality of the material can be assessed. If surge vessels are used in parallel and switch back and forth, then they also make it more feasible to divert off-spec material in the event that it is needed, because the material can be diverted from the surge vessel before flowing downstream. On the one hand, it is less likely that any material would need to be diverted as a result of a small process disturbance, because the surge vessels provide a buffer against brief fluctuations. On the other hand, if an entire surge vessel full of material does need to be diverted to waste, then it may be a larger amount of material at risk versus not having surge capacity in the system.

There are many disadvantages to surge vessels as well. They increase process footprint, increase amount of material in the process at any one time and therefore increase the quantity of material at risk at any time, broaden RTD, and therefore expand deviation boundaries, and require a larger number of pumps for the overall process train, because another pumping system is typically needed after each surge. When a process becomes more mature, eliminating the surge vessels is a natural goal. However, it can be invaluable to incorporate them into a new unproven continuous process train in GMP manufacturing.

3.12 Single Surge Vessel Versus Two or Three in Parallel

Option 1 If you have two parallel surge vessels, then one is filling from the upstream process while the other is emptying to the downstream process. The switch is made when one of the cans empties. For example, suppose the process flow is

100 L per day, the surge vessels are each 100 L, and the strategy is to operate with equal upstream and downstream decoupling. The vessels will operate at a level so that they switch back and forth once every 12 h, and there is a total of 50 L material in both surge vessels combined. In this example, there is up to 12 h of available decoupling in both the upstream and downstream directions. If there is a change so that the vessels are switching back and forth more frequently than once every 12 h, this means that upstream throughput is falling behind or has stopped for a time. If they are switching back and forth less frequently than 12 h, this means that downstream throughput is falling behind or has stopped for a time. Throughput adjustments can be made to rebalance surge vessel operating level. Putting the surge vessels are on weigh scales or installing level sensors on the surge vessels facilitates real-time mass balances within the process flow train.

Option 2 If the target mean residence time for the material in the surge vessels is low, then it may be better to use a single agitated surge vessel that is filling the same time that it is emptying. Operationally, this is the simplest approach. For example, suppose that the process upstream and downstream has several CSTRs with mean residence time in the range of 2–4 h. Suppose that the target mean residence time in the surge vessel is 4 h. In this example, the broadening of RTD by the surge vessel is not as significant compared to the RTD of the rest of the process. Here it is acceptable to use a single surge vessel with a CSTR residence time distribution because it does not significantly change the deviation boundaries and lot genealogy. One advantage is that it is less complex than the dual surge vessels option. However, if throughput becomes imbalanced and the CSTR surge vessels fills to higher operating levels, then it can have significant impact on RTD.

Option 3 A third option is to have three surge vessels in parallel. One is filling, one is emptying, and the third is sampled and waiting for an off-line analytical result before it switches on line. The disadvantage is that this requires more vessels and more complexity in the switching between vessels.

3.13 Surge Vessels Serve to Dampen Out Process Disturbances

In a continuous process step including reaction, extraction, distillation, and crystallization, there was an increase in an impurity for about 1 reactor volume turnover. (Johnson et al. 2012) A process hold led to higher levels of an epimerization product. During a planned weekend shutdown, the formation of a trans-impurity in the reaction resulted in about 1.8% of this impurity in the reactor rather than the steady state <1%.

Figure 3.22 shows that the trans-impurity spiked during a shutdown period, but quickly returned to steady-state conditions one volume turnover after the flows restarted.

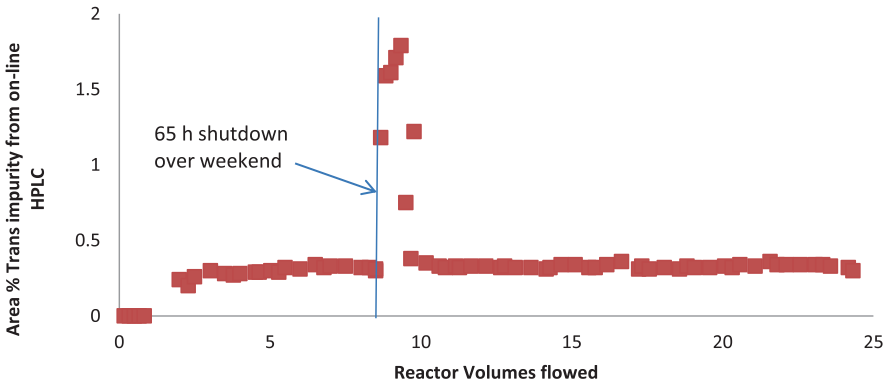


Fig. 3.22 Impurity concentration versus reactor volumes flowed for a PFR production run, showing a spike in impurity due to planned reactor shutdown. (Reprinted with permission from Johnson et al. (2012). Copyright (2012) American Chemical Society)

How did this spike in impurity track through the rest of the continuous train to the isolated solids?

Figure 3.23 shows the fully continuous process with surge vessels; 24 h surge vessels were installed after reaction, after extraction, and after solvent exchange distillation.

Figure 3.24 shows the process flow sheet in the gPROMS Model Builder® simulation.

Simulation of the scenario described above (reactor shutdown leading to impurity spike) reveals a significantly diffuse and diluted propagation of this impurity spike at the outlet of the second MSMPR (Scenario A, 24 h fill–empty surge tanks, impurity spike all within one surge vessel at surge point 1). This is shown in Fig. 3.25.

Additionally, it shows sensitivity to the timing of the upset. Splitting the upset across two surge vessels at surge point 1 gives a broader and less concentrated profile at the process outlet (Scenario B). Scenario B was closer to the actual experimental scenario. Two more hypothetical scenarios were simulated with fill–empty style surge cans, but reducing the periodicity to 12 h (Scenario C—spike not split, Scenario D—spike split in half between two consecutive 12 h surge vessels). The impact of smaller surge tanks decreases the length of time before material from the disturbance exits the process, and decreases the duration of the disturbance at the outlet. Also, the dampening effect is less (more concentrated impurity), and the effect of timing is less dramatic, as shown in Fig. 3.26.

Mode of operation can have a dramatic effect on the outlet distribution, as evidenced by Scenarios E, F, and G below. These scenarios represent "live" CSTR surge tanks of 1, 4, and 8 h residence times, respectively. General characteristics of this mode are significantly earlier breakthrough, but a much more diffuse profile for a given residence time, as shown in Fig. 3.27. Three CSTR surge vessels with 8 h τ

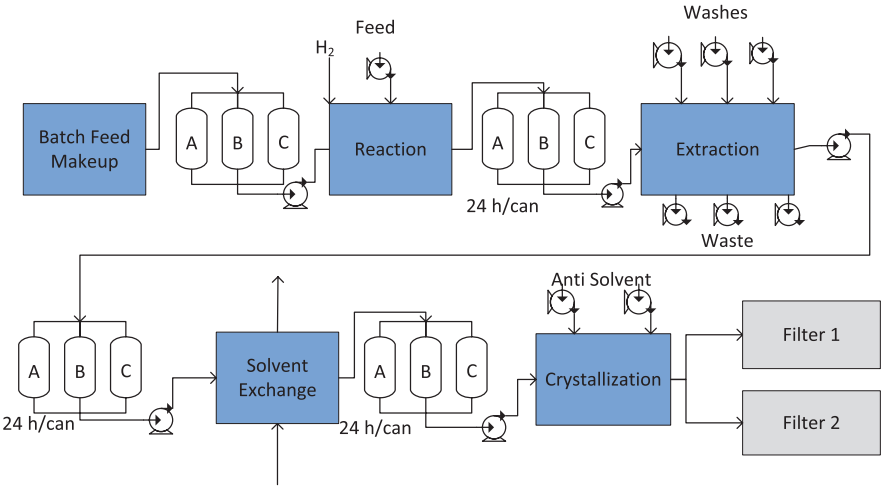


Fig. 3.23 Schematic of continuous process with four flow unit operations in series, decoupled by surge vessels in-between. A model of the process train was created using gPROMS ModelBuilder® 4.0.0.54901, primarily focused on residence time distribution for each unit operation. ModelBuilder® is a flow sheet based dynamic differential-algebraic equation solver utilizing both built-in and custom models. Each unit operation model is described briefly below:

Reaction: The reaction is modeled as a PFR with dispersion number (D/uL) = 0.0002. No reaction is considered explicitly; rather, the final product mixture is fed to the inlet, which is suitable for RTD modeling given that the PFR is almost ideal plug flow

Extraction: The mixtures are considered as two-phase systems, with the composite composition calculated as the unit is a CSTR with fixed volume. Organic–aqueous separation is considered instantaneous for the calculations. Each settler is considered as two separate CSTRs (organic/aqueous) with fixed volume

Solvent Exchange: The solvent exchange is treated as a semi-batch operation, with each batch being well-mixed and then separated to bottoms and distillate phases, because that is how the unit operation was physically run (Johnson et al. 2012)

Crystallization: The crystallization is treated simply as two single-phase CSTRs for the calculations. No solid phase is considered in these calculations because we are tracking a dissolved impurity, other than accounting for the volume occupied by the solids.

Surge Tanks: Surge tanks are treated as variable volume well-mixed vessels. In the fill–empty operating mode, three tanks reside between each unit operation and each tank is filled and then fully emptied in sequential order at a given periodicity. Zero heel was assumed when switching back and forth. This is the operating mode that was actually used for the real process. One vessel was filling, one vessel was feeding downstream, and the third vessel was waiting for off-line analytical results. The process was actually accumulating surge drums just upstream from crystallization, but it does not affect the calculations for dampening the process disturbance. For comparison, "live operating mode" for the surge tanks was also modeled. In the hypothetical live operating mode, only one tank resides between each unit operation, and it simultaneously receives feed and dispenses it. Volume is nearly constant, and initial volume divided by flow rate determines the residence time

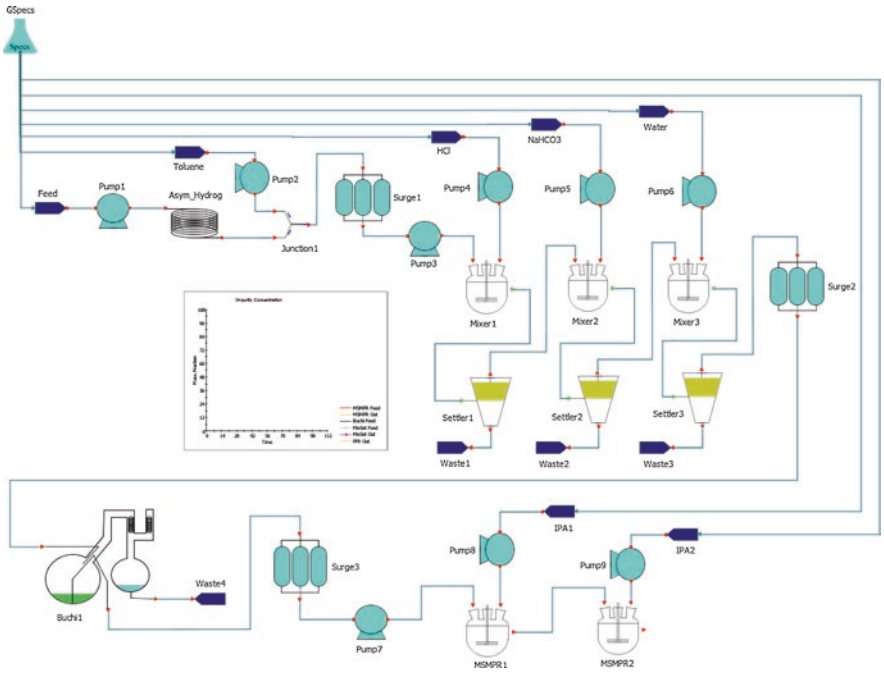


Fig. 3.24 Process flow sheet in the gPROMS Model Builder® simulation

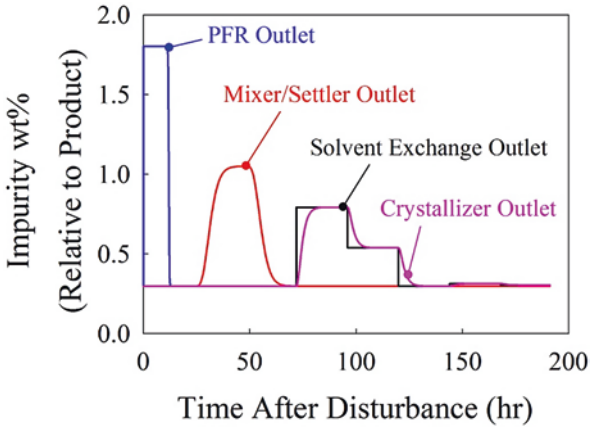


Fig. 3.25 Progression of the impurity spike through the system (Scenario A) shown as impurity content at the outlet of each unit operation

each add 24 h τ to the continuous train, but they add about 80 h to overall RTD. This is the disadvantage of CSTR surge vessels with long τ

An additional hypothetical scenario can be considered where the catalyst pump to the reactor shuts off for an hour, effectively sending an hour long pulse of unreacted starting material to the first surge tank. Even though the reactor sends an hour

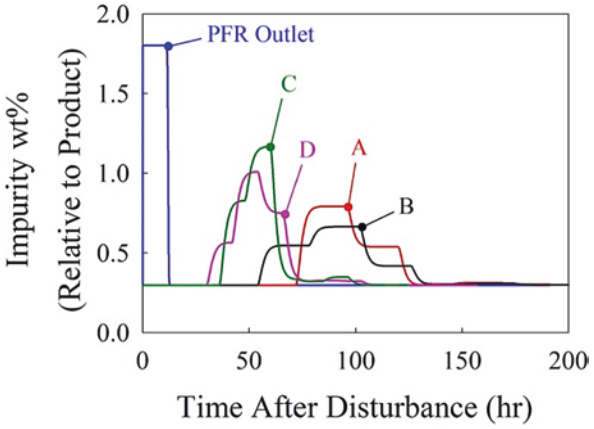


Fig. 3.26 Comparison of crystallizer outlet impurity wt% for four different "fill–empty surge" scenarios: all surge vessels as discrete vessels of 24 h flow capacity (A, B) and 12 h flow capacity (C, D). Additionally, the effect of disturbance timing is probed: the disturbance is contained within a single vessel immediately downstream from the reactor (A, C) or is split across two vessels (B, D) at surge point 1

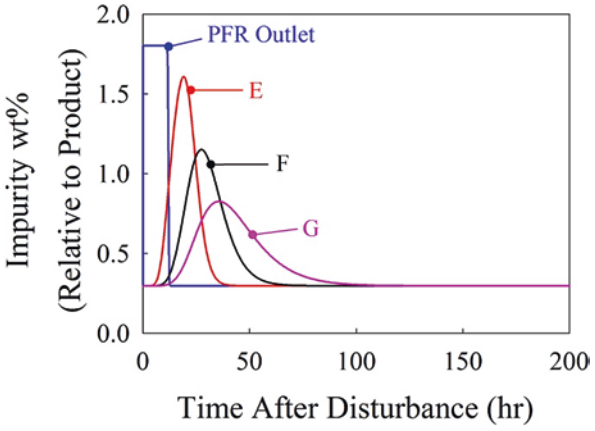


Fig. 3.27 Comparison of crystallizer outlet impurity wt% for three different "live surge" scenarios: all surge vessels as 15 min (E), 4 h (F), and 8 h (G) residence time CSTRs

long pulse of nearly 100% starting material to the first surge tank, the 24 h fill–empty surge strategy only tops out around 3% starting material exiting the crystallizers, as shown in Fig. 3.28. In this actual process, 3% unreacted starting material would have been rejected to the filtrate in the crystallization. Operating with very small CSTR surge vessels results in a much quicker propagation, but still only reaches around 8% at a maximum.

An additional scenario can be considered where the catalyst pump to the reactor shuts off for 10 min, effectively sending a 10 min long pulse of unreacted starting material to the first surge tank. Base case is 24 h fill–empty surge cans between unit

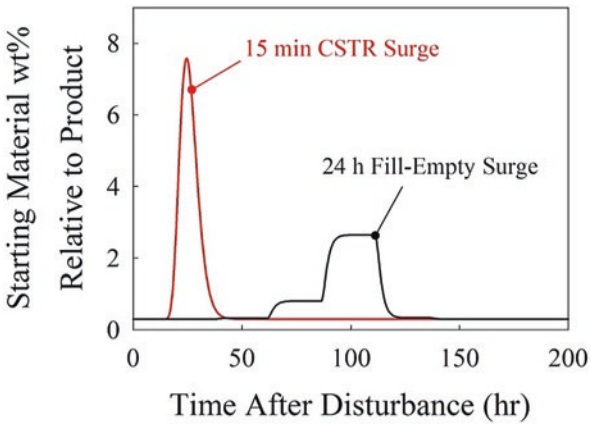


Fig. 3.28 Impact of operating mode on the propagation of a 1 h long stoppage of the catalyst pump for the reaction

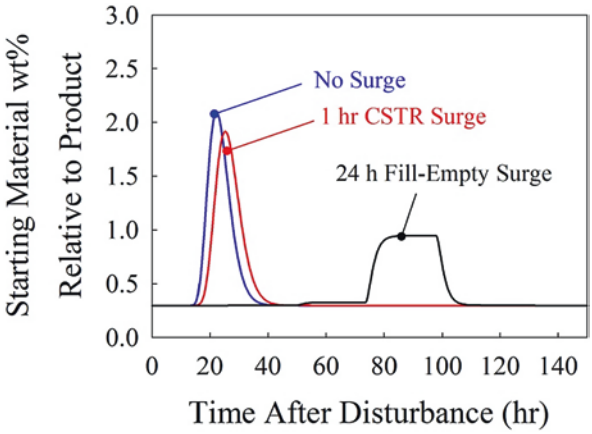


Fig. 3.29 Impact of operating mode on the propagation of a 10 min long stoppage of the catalyst pump for the reaction

operations, because that was the actual process. Even though the reactor sends a 10 min long pulse of nearly 100% starting material to the first surge tank, the process only reaches around 1% starting material exiting the crystallizers, as shown in Fig. 3.29. Operating with no surge vessels results in a much quicker propagation, but still only reaches around 2% at a maximum. If 1 h residence time CSTR surge vessels are utilized, there is a slight impact to disturbance dispersion and timing, but the dispersion in the system is already sufficient that the surge vessels are almost negligible to the residence time distribution.

The usage of dynamic flow sheet modeling is valuable for defining a strategy for surge capacity and for contaminated product diversion. The modeling can help to predetermine the limiting magnitude and duration of disturbance that is acceptable to progress downstream, versus the magnitude and duration of disturbance that should be diverted at the outlet from the continuous reactor.

In the same continuous process example, the throughput was not the same for each continuous unit operation, as shown in Table 3.4 (Johnson et al. 2012).

Table 3.4 Total continuous processing time for each unit operation in series

Continuous unit operation	Total run time for the campaign (hours)
Reaction	282
Extraction	282
Distillation	279
Crystallization/filtration	341

This demonstrates the importance of decoupling the crystallization/filtration from the rest of the continuous train by using surge vessels. Crystallization feed solutions accumulated in an inventory of additional 24 h drums. Surge vessels should be placed where a natural change in feed rate is expected, and in some circumstances, it is wise to plan for surge inventory, especially upstream of continuous crystallizations which can be problematic and suffer from periodic shutdowns.

Acknowledgments Jeff Niemeier, Brian Haeberle, Rick Spears contributed to residence time distribution characterization in PFRs by measuring axial dispersion, and quantifying importance of high L/d in the laminar flow regime. Declan Hurley contributed to lot genealogy and deviation boundaries as a function of RTD. Steve Jeffery and Robert Moylan contributed to sections on advantages and disadvantages of surge and stock vessels. James Cashman and Aoife Corrigan contributed to the start-up transition, dampening, and deviation boundaries considerations. We thank Bret Huff for initiating, leading, and sponsoring the continuous reaction design and development work at Eli Lilly and Company.

References

Aris R. On the dispersion of a solute in a fluid flowing through a tube. In: Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, vol. 235(1200): The Royal Society; 1956. p. 67–77.

Cole KP, Groh JM, Johnson MD, Burcham CL, Campbell BM, Diseroad WD, et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow cGMP conditions. Science. 2017;356:1144–50.

Danckwerts PV. Continuous flow of materials through processing units. Ind Chemist. 1954;30:102.

Edwards MF, Richardson JF. The correlation of axial dispersion data. Can J Chem Eng. 1970;48(4):466–7.

Fogler HS. Elements of chemical reaction engineering. 3rd ed: Prentice Hall PTR; 1999.

Johnson MD, May SA, Calvin JR, Remacle J, Stout JR, Diseroad WD, Zaborenko N, Haeberle BD, Sun WM, Miller MT, Brennan J. Development and scale-up of a continuous, high-pressure, asymmetric hydrogenation reaction, workup, and isolation. Org Process Res Dev. 2012;16(5):1017–38.

Johnson MD, May SA, Kopach ME, Groh JM, Cole KP, Braden TM, Shankarraman V, Merritt JM. Chapter 16: Design and selection of continuous reactors for pharmaceutical manufacturing. In: am Ende DJ, am Ende MT, editors. Chemical engineering in the pharmaceutical industry. 2nd ed: Drug Substance/API; 2019.

Johnson MD, et al. Intermittent flow and practical considerations for continuous drug substance manufacturing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020.

Lemoullec Y, Potier O, Gentric C, Leclerc JP. A general correlation to predict axial dispersion coefficients in aerated channel reactors. Water Res. 2008;42(6):1767–77.
Levenspiel OS. Chemical reaction engineering. An introduction to the design of chemical reactors. New York, NY: Wiley; 1962.
Levenspiel OS. The chemical reactor omnibook: Oregon State University Book Stores; 1979.
May SA, Johnson MD, Braden TM, Calvin JR, Haeberle BD, Jines AR, Miller RD, Plocharczyk EF, Rener GA, Richey RN, Schmid CR. Rapid development and scale-up of a 1 H-4-substituted imidazole intermediate enabled by chemistry in continuous plug flow reactors. Org Process Res Dev. 2012;16(5):982–1002.
Process Systems Enterprise. gPROMS. www.psenterprise.com/gproms. 1997–2016.
Taylor G. Dispersion of soluble matter in solvent flowing slowly through a tube. In: Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, vol. 219(1137): The Royal Society; 1953. p. 186–203.
Taylor G. Dispersion of matter in turbulent flow through a pipe. Proc R Soc Lond A. 1954a;223:446–68.
Taylor G. The dispersion of matter in turbulent flow through a pipe. In: Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences, vol. 223(1155): The Royal Society; 1954b. p. 446–68.
Weber WJ Jr, DiGiano FA. Process dynamics in environmental systems. New York, NY: Wiley; 1996.
Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer; 2009.

Chapter 4
Intermittent Flow and Practical Considerations for Continuous Drug Substance Manufacturing

Martin D. Johnson, Scott A. May, Jennifer McClary Groh, Timothy Braden, and Richard D. Spencer

Abstract Intermittent flow enables slurry flow out of continuous stirred tank crystallizers without solids plugging or clogging. It enables semi-continuous filtration, washing, and re-dissolving downstream from continuous crystallization. It allows solvent exchange distillation with strip to dryness in rotary evaporators to be a legitimate manufacturing unit operation for small-volume continuous processes. Intermittent flow back pressure regulation and vapor–liquid separation downstream from continuous high-pressure hydrogenation reactors tolerates a small amount of solids precipitate flowing out of the reactors without clogging or plugging, and they promote efficient pressure purge stripping of excess gas reagent. Intermittent flow stirred tank reactors are a practical alternative to plug flow reactors (PFRs) for heterogeneous reactions. Eleven examples of continuous reactions are given that have been run at manufacturing scale in PFRs. Mean residence time ranges from 0.7 to 24 h in the 11 examples; therefore, it is not necessary for a reaction to be extremely fast in order to be a viable candidate for flow chemistry. This chapter gives many general guidelines on how to design and operate a continuous process, avoiding many of the common operational, equipment, analytical, and process chemistry pitfalls. The continuous process checklist serves to help prevent common oversights.

Keywords Continuous processing · Drug substance · PFR · CSTR · Intermittent flow · Recycle · Crystallization · Filtration · Distillation · Reaction

M. D. Johnson (✉) · S. A. May · J. M. Groh · T. Braden · R. D. Spencer
Eli Lilly and Company, Indianapolis, IN, USA
e-mail: johnson_martin_d@lilly.com; may_scott_a@lilly.com; groh_jennifer_mcclary@lilly.com; braden_timothy@lilly.com; spencer_richard_d@lilly.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_4

Abbreviations

API	Active pharmaceutical ingredient
CSD	Crystal size distribution
CSTR	Continuous stirred tank reactor
DARA	Direct asymmetric reductive amination
DCS	Distributed control system
FBRM	Focused Beam Reflectance Measurement
i.d.	Inside diameter
MSMPR	Mixed suspension mixed product removal
PAT	Process analytical technology
PFA	Perfluoroalkoxy
PFR	Plug flow reactor
PIDs	Process and instrumentation diagrams
PMI	Process Mass Intensity
RTD	Residence time distribution
τ	Mean residence time

4.1 Intermittent Flow

Intermittent flow enables solids processing, heterogeneous reactions, different modes of reagent addition, crystallization, filtration, more efficient solvent exchanges, back pressure regulation without restricting orifices, and complete conversion for reactions. Other researchers have used intermittent flow. In the work of Adams this is known as semi-continuous operation (Adams and Pascall 2012). Adams describes the forced cyclic process with no steady states, and demonstrates that it is possible to achieve multiple separations steps and high reaction conversion using fewer vessels than would be required in truly continuous operation.

4.1.1 Slurry Flow Out of Continuous Stirred Tank Crystallizers

Intermittent flow of slurries is especially important when average volumetric throughput is less than about 200 mL/min and internal tubing diameter is more than about 4 mm. If Reynolds number is less than about 2000, then flow is in the laminar regime, which is not sufficient for keeping the solids suspended and keeping the solids from clogging. Solids gradually settle in the piping and fittings and eventually accumulate enough to clog the process tubes. Settling and accumulation of solids especially occurs at locations of fittings, elbows, tees, valves, or any other slight constriction or expansion in the flow path. On the other hand, flowing slurry quickly and intermittently out of a stirred tank for brief time periods enables high

enough linear velocities to achieve turbulence. Intermittent flow is generally achieved by pressure differences and sequenced automated block valves, but it can also be achieved by automated intermittent mechanical pumping.

Continuous crystallization was utilized for kinetic rejection of a chiral impurity in the manufacture of a key intermediate in the synthesis of LY500307 (Johnson et al. 2012). A fully continuous process step was used to generate 144 kg of penultimate in laboratory fume hoods and a laboratory bunker. The continuous process train included high-pressure asymmetric hydrogenation, liquid–liquid extraction, solvent exchange distillation, crystallization, and filtration. Crude reaction product solution had ee in the range 92–94% throughout the continuous campaign. Crystallization was required to upgrade the advanced intermediate to greater than 99% ee. Unfortunately, crude ee was on the unfavorable side of the eutectic; therefore, thermodynamic equilibrium in the crystallizer would fail to provide the required ee upgrade. However, kinetics favored crystallization of the desired chirality product, because the undesired enantiomer crystallized out of solution more slowly. Continuous crystallization in stirred tanks in series is superior to batch for kinetic impurity rejection because the crystallizers operate in the kinetic regime, by definition, with steady-state supersaturation. Stirred tank crystallizers are often called mixed-suspension mixed-product removal (MSMPR) crystallizers. The product was in solution prior to crystallization, with toluene being the main solvent. Isopropyl alcohol anti-solvent and cooling were used to generate the supersaturation that drove the crystallization. The desired compound dissolved in toluene was continuously pumped into the first of two MSMPRs in series. Anti-solvent was continuously pumped into both MSMPRs. A sketch of the setup is shown in Fig. 4.1.

Overall average slurry flow rate was 61 mL/min from MSMPR1 to MSMPR2, and it was 119 mL/min from MSMPR2 to the dual alternating filters. Without intermittent operation, the slurry flow would have been in the laminar regime and resulted in solids clogging over time. However, because slurry flowed intermittently with about 1 m/s linear velocity, the process ran a total of more than 400 h without

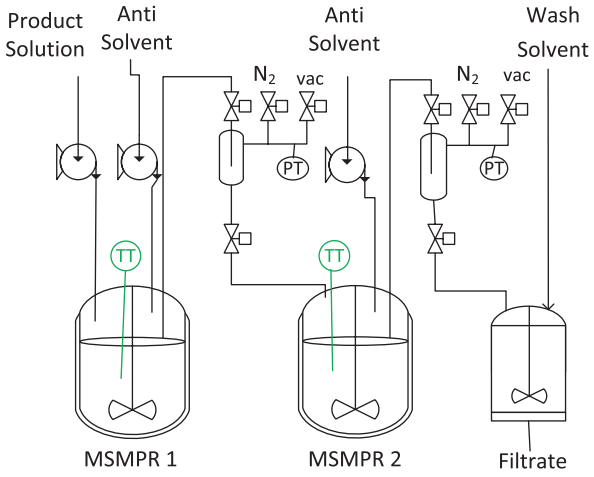


Fig. 4.1 Sketch of 2 MSMPRs in series with intermittent flow slurry transfer zones

fouling or clogging in the transfer tubes. The process was stopped each weekend. The longest individual operation time without stopping was 92 h. At the end of the 92 h there was no sign of clogging or fouling in any of the process tubing. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 5 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 30 min. Slurry slug volume between the MSMPRs was 0.3 L, and slurry slug volume between MSMPR2 and filter was 3.6 L. Please see Table 12 in the publication for more details on the intermittent slurry flow out of each MSMPR (Johnson et al. 2012). The filter cake was automatically intermittently washed after each transfer of slurry to the filter to keep the undesired enantiomer from precipitating in the filter cake. Two four-valve transfer zones designed for intermittently pumping slurries out of MSMPR crystallization vessels are shown in Fig. 4.1. Each four-valve transfer zone operated with a repeating sequence controlled by the automation system. Consider the transfer zone between MSMPR 1 and MSMPR 2 in the figure. The vacuum valve opened until pressure in the zone reached about 300 mm Hg, then it closed. The inlet valve opened to pull slurry at about 1 m/s linear velocity from MSMPR 1 into the transfer zone, then it closed. Slurry pulled out of MSMPR1 to decrease slurry level in the vessel to the dip tube, which was positioned at a height corresponding to 5.3 L remaining in the vessel. The nitrogen valve opened until the zone was pressured up to about 2 bar, then it closed. Finally, the outlet valve opened to push slurry into MSMPR 2 at about 1 m/s linear velocity, then it closed. Total time for this sequence was about 40 s. The sequence repeated automatically once every 5 min, controlled by the deltaV distributed control system (DCS). A similar automated sequence was used to transfer slurry to the filters once every 30 min. Similar to MSMPR1, slurry pulled out of MSMPR2 to empty the vessel to dip tube level, which was 8.9 L. Therefore, level sensors were not needed for automated level control in either of the crystallization vessels. A picture of the 8 L transfer zone for transferring slurry from MSMPR2 to the filter is shown in Fig. 4.2.

It is important that the automated valves have approximately the same internal diameter as the perfluoroalkoxy (PFA) tubing, minimizing any constrictions or expansions. In this example, the process tubing and the valve inside diameter were about 1 cm. Constrictions or expansions are typical locations for solids to build up and eventually clog. Automated ball valves with internal diameter as close as possible to the internal diameter of the PFA tubing work well. Figure 4.2 shows tubing and valves with internal diameter about 1 cm. It is also important to avoid elbows or any other types of fittings with sharp bends, because these are common locations for solids fouling. It is best if the PFA tubing flows straight into a two-way ball valve and straight out the other side with no bends. Valve orientation is also important because of the impact of gravity. If the valve is mounted vertically at the bottom outlet of the transfer zone, then solids can settle on the ball valve during the time that the transfer zone is filling and the slurry is accumulating in the zone. As seen in Fig. 4.2, the outlet valve from the bottom of the transfer zone was mounted horizontally and at a higher elevation than the bottom outlet. PFA tubing at the transfer zone outlet bends smoothly 270° from the bottom outlet up, around, and into the side of the outlet valve. This way, if solids settle by gravity in the transfer zone, then they

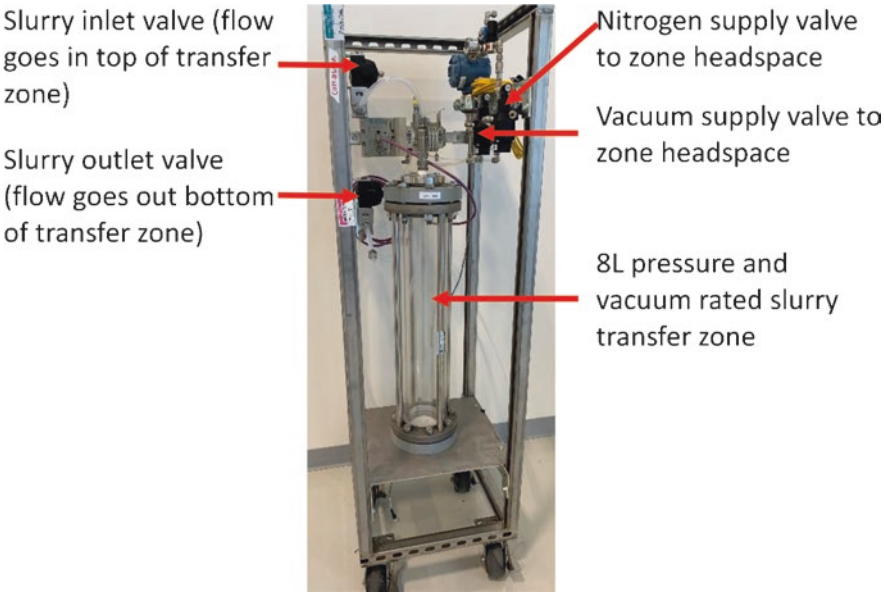


Fig. 4.2 Picture of 8 L four-valve transfer zone for transferring slurry from MSMPR2 to a filter

settle down into the smoothly curving part of the PFA tubing, so that when flows commence the solids will have less opportunity to hang up in the tubing.

Continuous crystallization was used for impurity rejection in the middle of a multistep continuous process for merestinib (Reizman et al. 2019). A picture of the 50 L glass MSMPRs with overhead stirring and custom baffle cages inserted is shown in Fig. 4.3.

One of the main benefits of running this process continuously was that the product was a highly potent compound. Small-volume continuous allowed the process to run in laboratory fume hoods at commercial manufacturing scale. Similar to the previous example, the crystallization was driven by cooling and anti-solvent. However, this was primarily a cooling crystallization, and the feed was kept heated at 50 °C to maintain solubility into the first MSMPR-in-series. Two feeds continuously flowed into the first MSMPR, the pharmaceutical compound dissolved in THF, and cyclohexane anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 15 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 15 min as well. Overall average slurry throughput was 179 mL/min, which would have been in the laminar flow regime if flows out of each MSMPR were truly continuous. However, 2.7 L slurry transferred quickly by turbulent flow once every 15 min, with linear velocity about 1 m/s, through 1 cm inside diameter PFA tubing. The end result was that the GMP manufacturing campaign ran for 17 days with no stopping and no clogging or fouling in the transfer tubes. The campaign produced more than 180 kg of API. Please see the publication for more details (Reizman et al. 2019). Like in the previous example, slurry flowed

Fig. 4.3 Picture of the 50 L MSMPRs

from MSMPR2 to the dual filters using an 8 L four-valve transfer zone, as described and shown in Figs. 4.1 and 4.2. However, in this example, slurry flowed from MSMPR1 to MSMPR2 by a different mechanism. Vacuum was temporarily applied to MSMPR2, which pulled slurry from MSMPR1. Between the MSMPRs was only a simple 1 cm *i.d.* curved PFA tube that arched up and over from one vessel to the next. During the transfer time, an automated block valve closed the vent from MSMPR2, a second automated block valve opened the vacuum supply to MSMPR2, and a third automated block valve opened an extra nitrogen supply to the headspace of MSMPR1 so that it remained at atmospheric pressure during the transfer and did not suck back from the vent bubbler. In this manner, slurry level in MSMPR1 was reduced to dip tube level once every 15 min. This method is generally preferred for cooling crystallizations, because the time for slurry to travel from MSMPR1 to MSMPR2 is only about 2 s; thus, it does not have time to cool and precipitate in between the vessels. Figure 4.4 shows a picture of the skid with automated block valves that accomplished the transfer from MSMPR1 to MSMPR2. The skid was designed and constructed to control slurry flows for up to four MSMPRs-in-series, which is why the extra valves, transmitters, and vacuum pots are seen in the picture.

A detailed description of the automated sequence used for this type of intermittent transfer, and a design sketch of the apparatus, is given on pages S9 through S11 in the supporting information section of a publication (Kopach et al. 2016).

Continuous crystallization was used for impurity rejection and isolation of a cytotoxic API in a multistep continuous process for tasisulam (White et al. 2012). One of the main benefits of continuous processing for this product was that the entire cytotoxic segment of the synthetic route was run in laboratory fume hoods in

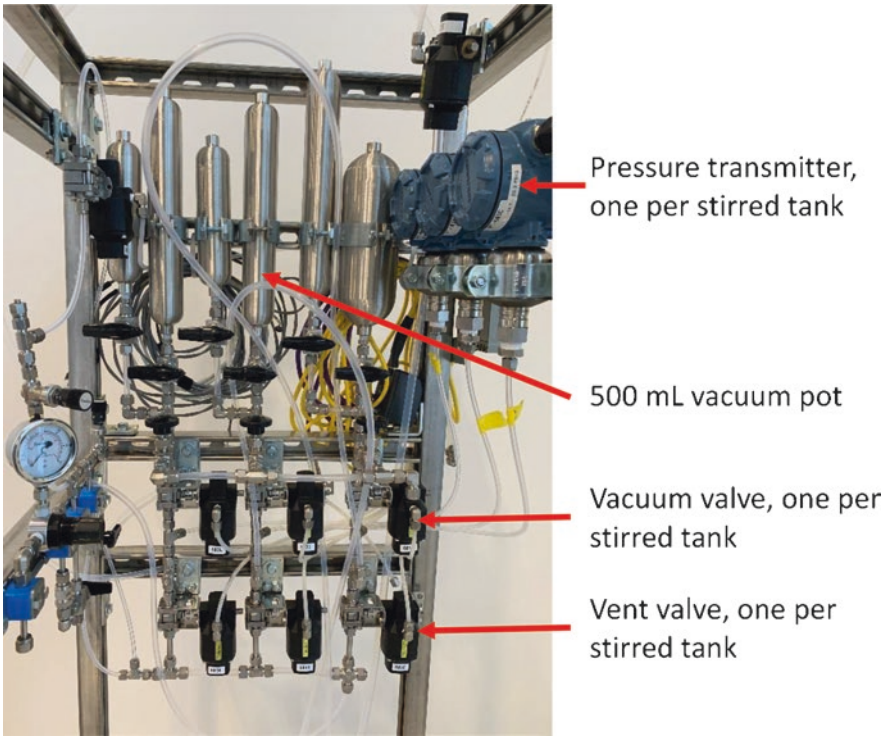


Fig. 4.4 Automated skid with sequenced block valves that transferred slurry from MSMPR1 to MSMPR2

disposable, inexpensive flow equipment. Two MSMPRs-in-series were used for the continuous crystallization. Two feeds were continuously pumped into the first MSMPR. One was the desired compound dissolved in 60/40 isopropyl acetate and isopropanol. The second was heptane anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 3 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 30 min. Before scaling up to production, the crystallization was operated at research scale. Overall average slurry flow at research scale was 0.6 mL/min. PFA tubing with 3 mm inside diameter was used for slurry flow. Therefore, the slurry flow at research scale would have had about 0.001 m/s linear velocity. Solids would have precipitated in the tubing and clogged if flow was truly continuous. However, because of intermittent flow, linear velocity in the 3 mm *i.d.* tubing was actually about 0.1 m/s in the research scale continuous crystallization experiments, which prevented solids clogging.

Continuous crystallization was used for impurity rejection in the middle of a multistep continuous process for prexasertib (Cole et al. 2017a). Crystallization was important to impurity control strategy, necessary to reject pyrazine and regio isomers downstream from a continuous S_NAr reaction. Purity of the crystallized solids was in excess of 99.8 area% throughout the campaign. There were many benefits of

continuous processing for this product, as described in the publication. One of the main benefits was containment of the entire process in laboratory fume hoods, which was important because the product was cytotoxic. Two feeds continuously flowed into the first MSMPR, the pharmaceutical compound dissolved in DMSO, and methanol anti-solvent. Slurry flowed intermittently from MSMPR1 to MSMPR2 once every 3 min, and it flowed intermittently from MSMPR2 to the dual alternating filters once every 6 min. The next section describes the importance of the dual alternating filters for eliminating solids handling.

4.1.2 Filtration Downstream from Continuous Crystallization

In the prexasertib example, continuous crystallization flowed into semi-continuous filtration. Semi-continuous filtration eliminated manual handling of the genotoxic and cytotoxic intermediate. Therefore, the impurity rejection benefits of crystallization were realized, without the manual handling of solids that would typically be required in batch. The automated dual filters switched back and forth once every hour. The on-line filter received slurry slugs from the crystallizer, while the off-line filter was automatically rinsed with methanol, washed with MTBE, and dried with nitrogen. Then, the product was automatically dissolved off the filter in formic acid. Formic acid was the solvent and reagent for the downstream BOC deprotection continuous reaction. A simplified sketch of the automated intermittent flow dual filters in shown in Fig. 4.5.

A detailed step-by-step description of the automated dual filter process is given in the supplementary information for the publication (Cole et al. 2017a). A picture of the dual filter skid is shown in Fig. 4.6.

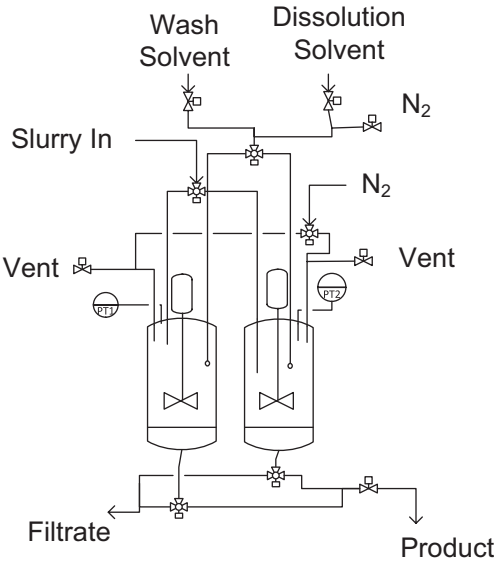


Fig. 4.5 Sketch of the automated intermittent flow dual filters

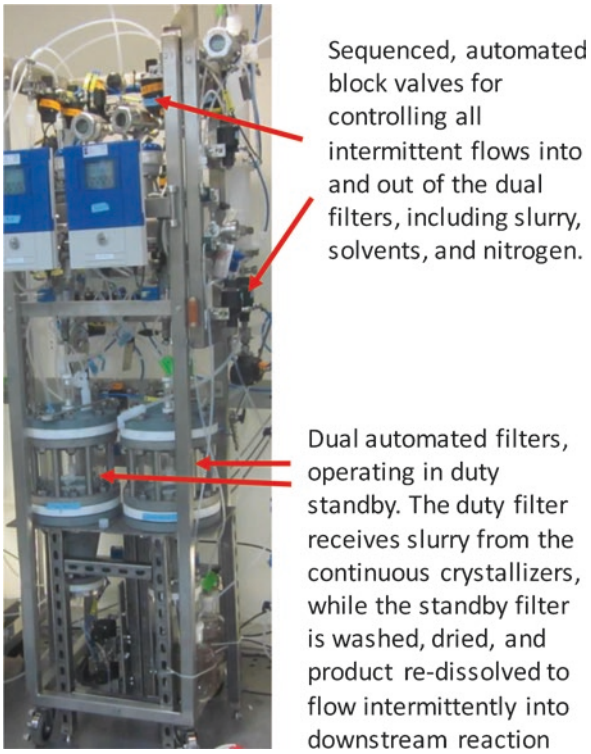


Fig. 4.6 Picture of the automated intermittent flow dual filters

4.1.3 Solvent Exchange Distillation with Strip to Dryness in Rotary Evaporators

Intermittent flow also enables efficient solvent exchanges in the middle of an otherwise fully continuous process train. Solvent exchange from high boiling point solvent to low boiling point solvent is much more efficient with stripping to dryness and then adding back the lower boiling solvent. This is not possible in a typical batch vessel in a manufacturing plant because of the need to maintain a minimum stir-able volume, but strip to dryness solvent exchange is possible using intermittent flow 20–50 L laboratory rotary evaporators, which can give commercial scale throughput for small-volume continuous processes. The automation charges product solution, then strips off the first solvent, then empties the distillate, then adds back the second solvent dissolving the product, and then empties. The DCS automated control system repeats the sequence, typically about once per hour, for the entire duration of the continuous campaign. It is not truly continuous, it is really automated repeating batch, but it is practically continuous, and it is an effective way to incorporate solvent exchange into the middle of an otherwise continuous process train in laboratory fume hoods.

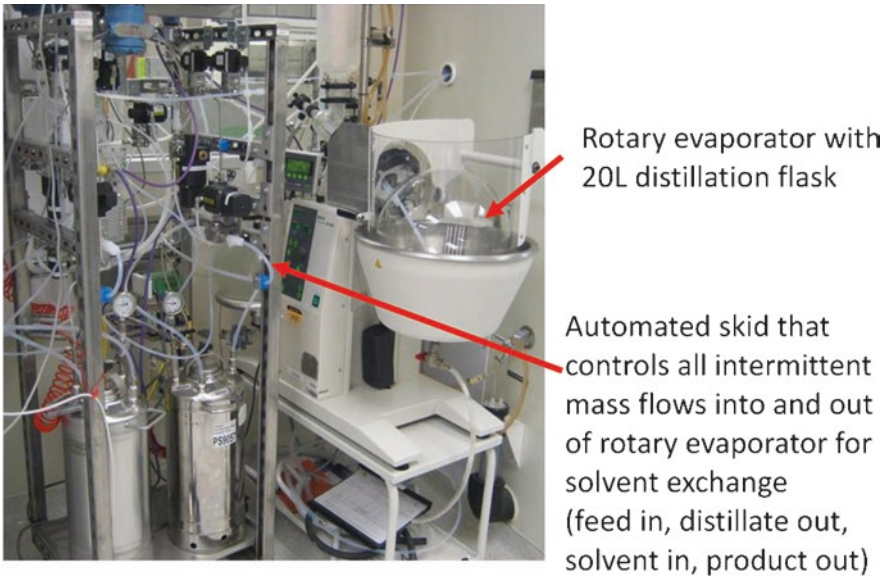


Fig. 4.7 Picture of automated intermittent solvent exchange distillation equipment

Automated intermittent solvent exchange with strip to dryness was used in the continuous process steps for tasisulam (White et al. 2012). This was a particularly challenging solvent exchange. Desired product was dissolved in a mixture of toluene, methyl THF, and residual water after a continuous Schotten-Baumann reaction and liquid–liquid extraction. The goal was to solvent exchange into a precise ratio of 60/40 isopropyl acetate and isopropanol. Residual solvent target was <0.1% water and <0.5% toluene. Furthermore, achieving the precise ratio of isopropyl acetate to isopropanol was important for the subsequent continuous crystallization. These are challenging requirements for either a batch solvent exchange distillation or a truly continuous solvent exchange distillation. However, the automated intermittent rotary evaporator with strip to dryness achieved the solvent exchange goals with ease. The solvent exchange system that was used to process 20 kg cytotoxic API at 5 kg/day throughput is shown in Fig. 4.7.

A simplified sketch of the apparatus is shown in Fig. 4.8.

Solids were present temporarily in the rotary evaporator during the solvent strip, but the solids were redissolved prior to product solution flowing out of the evaporator.

Automated intermittent solvent exchange with strip to dryness was also used in the continuous process steps for prexasertib (Cole et al. 2017a). In this example, a formate salt isolation of a cytotoxic intermediate was eliminated by using stripping to remove formic acid to less than 0.8 equivalents relative to the API. This was not feasible by either batch distillation or truly continuous distillation. Batch distillation could not accomplish the formic acid removal because the extended times required for distillation resulted in impurity formation. Truly continuous distillation was not feasible because of the solid precipitate that forms when formic acid is stripped away. The intermittent flow rotary evaporator was able to achieve formic acid removal

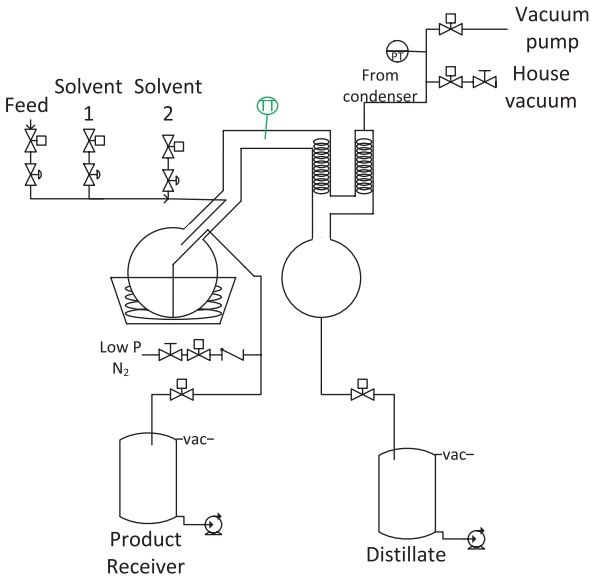


Fig. 4.8 Sketch of automated intermittent solvent exchange distillation equipment

without significant impurity formation because of the ability to strip to a dry thin film. The automated sequence charged formic acid solution of the API, then charged lactic acid, then stripped to dryness, then charged water, then stripped to dryness again, emptied the distillate, then charged water and THF, dissolved the API, and emptied the product solution. A detailed step-by-step description of the automated intermittent evaporator sequence is given in the supplementary information for the publication (Cole et al. 2017a). Figure 4.9 shows the front side of the automated skid that controlled all the flows into and out of the rotary evaporator and controlled the automated pressure profile. Figure 4.10 shows the back side of the unit. From the back side, the Coriolis mass flowmeters are seen, which were used by the DCS system to achieve accurate and precise mass charges of each intermittent flow.

4.1.4 Back Pressure Regulators and Vapor–Liquid Separators Downstream from Continuous High-Pressure Hydrogenation Reactors

Continuous reaction was used for an Ru-catalyzed direct asymmetric reductive amination (DARA) reaction for producing an API intermediate. The reaction operated with 68 bar hydrogen pressure. Continuous processing offered safety, throughput, and capital cost advantages compared to batch high-pressure hydrogenation. Intermittent flow was used at the exit of the continuous reactor to depressurize reactor contents and forward reaction product material to a surge vessel. First, the process was developed at research scale with about 0.2 kg/day throughput, then it was demonstrated at pilot scale with about 0.7 kg/day throughput, and finally it was

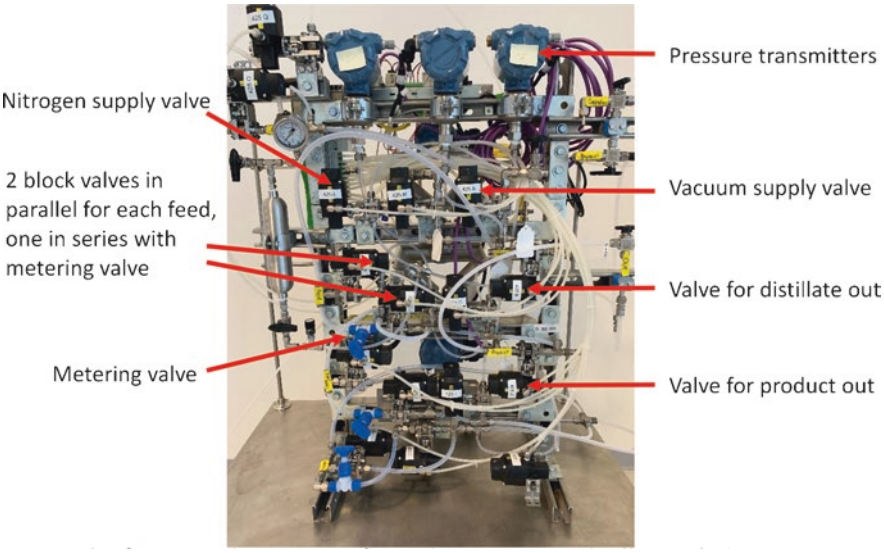


Fig. 4.9 Front side of automated intermittent flow solvent exchange distillation skid

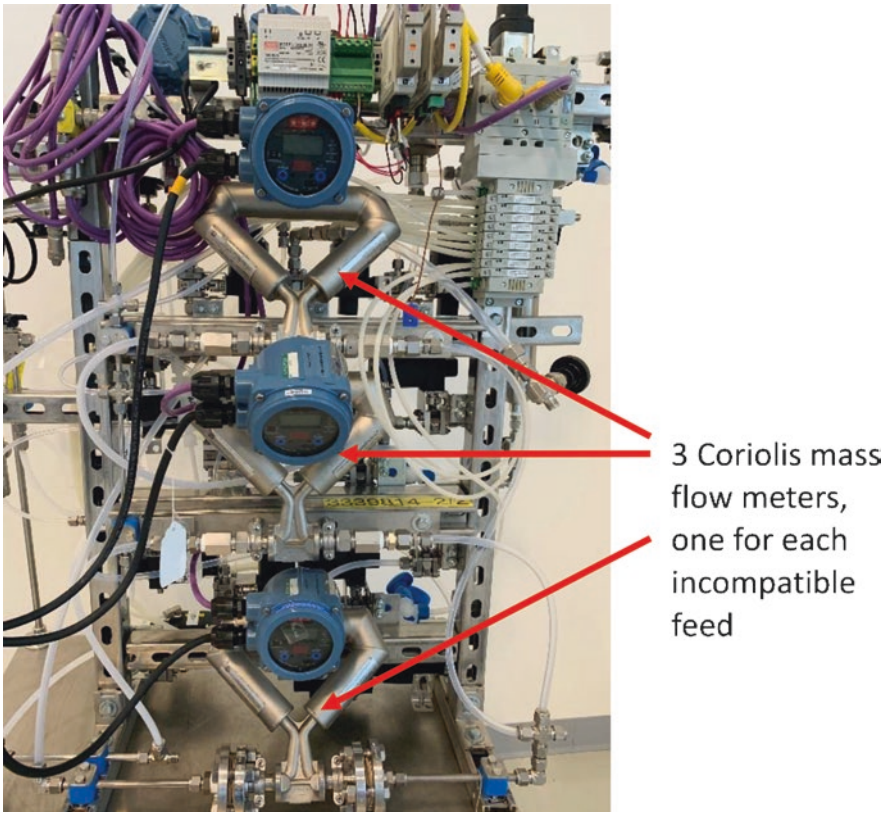


Fig. 4.10 Back side of automated intermittent flow solvent exchange distillation skid

scaled up to manufacturing for a 3000 kg GMP validation campaign with throughput about 100 kg/day (Johnson et al. 2016; Changi et al. 2017). Intermittent flow at the reactor exit was used at all three scales. The main advantage of intermittent flow at the PFR outlet was that it tolerated some amount of solids precipitate in the back pressure regulation and gas/liquid separation section, without solids clogging or fouling. A detailed description of the design, automation, and operation of the intermittent flow back pressure regulator is given on pages S14 to S25 of the supporting information to the publication (Johnson et al. 2016). Excess gas reagent and product solution temporarily pooled in the first of a series of pressurized vessels. This vessel operated at pressure equal to reactor outlet pressure. Intermittently, once every 14.5 min, the reaction product solution was transferred through expansion vessels in series by the opening and closing of sequenced automated block valves between the vessels. The pressure trends in the reactor as a result of this mode of operation are shown in the supporting information section of the publication. This type of back pressure regulation is an alternative to a standard restricting orifice back pressure regulator that would maintain truly continuous flow at the exit of the reactor. For example, a pressurized diaphragm dome style back pressure regulator, a spring-loaded back pressure regulator, or a small orifice automated metering valve could be used as an alternative. The traditional commercially available style back pressure regulator would also maintain constant pressure within the reactor, rather than the deliberate oscillating pressure swings. However, these have restricting orifices that can clog and foul with solid particles. In contrast, the intermittent flow approach using the expansion chambers in series and the automated block valves does not have any restricting orifices. When an automated block valve is opened to allow intermittent flow from one chamber to the next, the flow path has large diameter and material transports at extremely high linear velocities. At production scale, the valves have at least a 1 cm inside diameter flow path. The end result is that the intermittent flow back pressure regulation system can run for longer times and tolerate small amounts of solids without clogging or fouling. Figure 4.11 shows a research-scale intermittent flow back pressure regulation system, and Fig. 4.12 shows a pilot-scale intermittent flow back pressure regulation system.

The pressure pots and automated block valves are stainless steel. They are pressure rated to more than 100 bar, although they are typically used at pressures 70 bar or less.

The same type of intermittent flow back pressure regulator was used for the hydroformylation example given in another chapter in this book (Johnson et al. 2020a); 50/50 H_2/CO was the reagent gas mixture, and the reactor operated at 68 bar pressure. Continuous reaction had safety and capital cost advantages compared to batch. In that example, catalyst particles precipitated because the catalyst/ligand was not soluble in the mixed aldehyde product. In addition, some of the methylmethacrylate polymerized, forming solids in the reactor. The slurry exiting the reactor contained about 1 wt.% solids, much of which was sticky polymer. Nevertheless, the reactor operated continuously for the entire 314 h continuous campaign without any signs of fouling or clogging at the outlet of the reactor through the back pressure regulation section, depressurization, and gas/liquid separation.

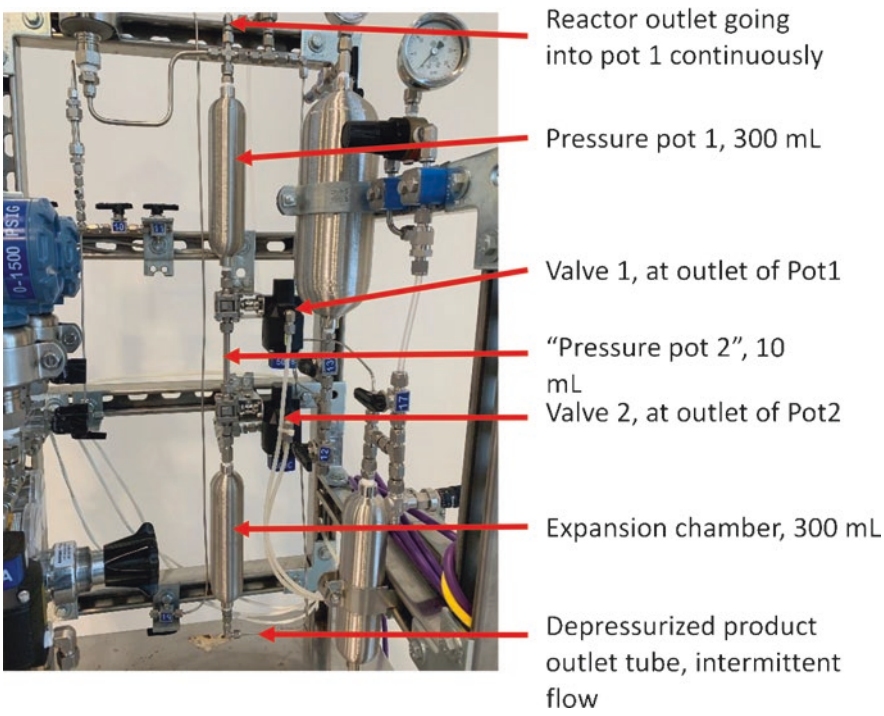


Fig. 4.11 Research-scale intermittent flow back pressure regulation system

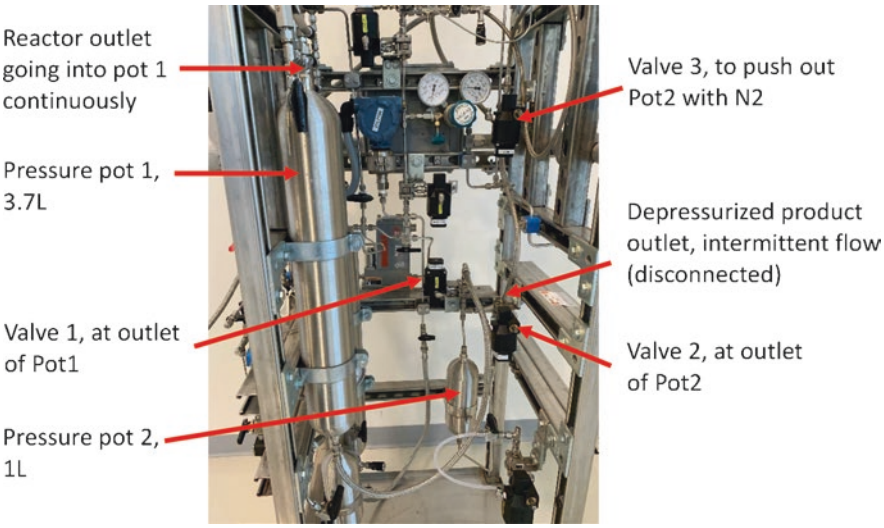


Fig. 4.12 Pilot-scale intermittent flow back pressure regulation system

Intermittent flow was used at the exit of a packed catalyst bed reactor because of catalyst fines that exit the reactor during the first several hours of start-up with each new catalyst bed. A hydrogenolysis reaction was run in the continuous packed bed reactor because it minimized formation of a desF impurity compared to batch. For details about the chemistry and the reactor design please see the publication (Zaborenko et al. 2015). If a traditional flow-restricting orifice back pressure regulation device had been used, it would have fouled because of the small catalyst particles. Filtration of the tiny catalyst particles was not an option because of the buildup of high pressure drop across the filter, restricting flow. The intermittent flow back pressure regulation at the exit of the packed catalyst bed reactor was a more reliable approach because it operated for 92 h without any signs of fouling. Material of construction was hastelloy C276 rather than stainless steel because of the HCl used in the hydrogenolysis reaction.

Intermittent flow was used downstream from a high-pressure reductive amination reactor in order to facilitate gas/liquid separation and hydrogen stripping from the liquid product solution. The manufacturing plant produced 2000 kg of product in a GMP registration stability campaign. The continuous reactor ran with 50 bar hydrogen pressure. The reasons for continuous processing, the reaction process, and automated equipment are described in detail in the publication (May et al. 2016). The main safety advantage of running this high-pressure hydrogenation reaction continuously was that the hydrogen supply, reactor, back pressure regulator, gas liquid separator, and gas stripping were all located outside the building. Reaction product solution flowing back into the building was practically free of hydrogen. This is something that is feasible continuous but is not feasible batch, because batch autoclaves must be opened for charging materials. They must be inside the building in a controlled bunker. In contrast, the continuous reactor is always sealed; therefore, it can be located outside. Detailed designs, description of automated sequences, and pressure trends are given on pages S25 through S37 in the supporting information of a different publication (Johnson et al. 2016). This gas–liquid reaction was relatively clean, meaning that solids precipitates were not expected, only vapor and liquid phases. Therefore, a dome diaphragm style restricting orifice back pressure regulator was used to maintain constant back pressure of 50 bar at the exit of the reactor. However, immediately downstream from the back pressure regulator, intermittent flow was used to separate excess hydrogen gas from the liquid product solution without misting out the vent. The intermittent flow chambers in series stripped hydrogen from the solution with nitrogen and forward material intermittently to a surge vessel downstream once every 19.5 min. Intermittent flow chambers were efficient at gas stripping because of the ability to do pressure purges.

4.1.5 Stirred Tank Reactors for Heterogeneous Reactions

A separate chapter in this book described an intermittent stirred tank reactor used for a Suzuki-Miyaura cross-coupling reaction in the merestinib continuous process (Johnson et al. 2020a; Cole et al. 2016, 2019). As described in that example, the

intermittent operation of the fill–empty reactor enabled full conversion in just 20 min, while it would have required much longer residence times in a CSTR or CSTRs-in-series. Furthermore, intermittent flow allowed the catalyst to be added last at a controlled flow rate, which improved catalyst activity. A 6 L stirred tank was used for a 100 hour continuous run to make 22 kg of product. The reactor turned over automatically about 45 times per day, therefore reactor volume was about 50–100 times smaller than what would have been required in traditional batch processing. It was not a truly continuous reaction, rather it was automated repeating batch, but it was practically continuous, and it integrated well with an otherwise fully continuous process in laboratory fume hoods.

An intermittent flow stirred tank reactor was used for a thermal cyclization reaction in which the product precipitated in the reactor (White et al. 2014). This would have clogged a plug flow tubular reactor. A truly continuous stirred tank reactor was not used because it would have required a much longer reaction time to reach the same conversion. The reactor operated under elevated pressure, therefore filling and emptying the reactor was more feasible with intermittent flow than truly continuous flow. The benefit of intermittent flow compared to batch was that the reaction ran at extremely high temperatures and pressures, beyond the capabilities of most batch reactors. The reactor turned over once every 8 min, therefore it was more than 100 times smaller than a batch reactor operating once per day. Temperature was 265 °C and pressure was 41 bar; 300 g of advanced intermediate was produced with 700 automated turnovers of a 25 mL autoclave.

An intermittent flow stirred tank reactor with catalyst recycle was used for an enantio-selective Aza-Henry reaction (Tsukanov et al. 2016). The reagents were homogeneous solutions, but the product was insoluble in the reactor. The process was more efficient and the reaction time was shorter when the catalyst and the excess reagent were recycled. A safety benefit of intermittent flow compared to batch for this reaction was that the intermittent flow reactor was smaller, reducing the amount of nitro alkane in the system. The reactor turned over once every 40 min, and 25 g product was made with 16 automated turnovers of a 250 mL reactor.

An intermittent flow stirred tank reactor was used for a nitro group reduction (Cole et al. 2017b). A trickle bed hydrogenation reactor with fixed catalyst packed into a column could have been used as a truly continuous plug flow reactor for this type of reaction. However, a trickle bed reactor requires a larger catalyst particle size than what is used batch, because small particle size catalysts create excessive pressure drop as the reaction solution and hydrogen flow through the column. Therefore, a benefit of the intermittent flow stirred tank compared to trickle bed for this chemistry was that the same catalyst that was developed for the batch reaction could be used for the intermittent flow reaction. Another benefit of the intermittent flow reaction compared to traditional batch hydrogenation was that the reactor was about 2 orders of magnitude smaller for the same overall process throughput. This is a significant safety advantage because it reduces the amount of hydrogen in the system. Moreover, the 1 L hydrogenation reactor was approved for operation in a laboratory fume hood, while a 100 L batch hydrogenation reactor surely would have needed a specially designed hydrogenation bunker, which is not available to most research labs. The small particle size heterogeneous catalyst was sequestered in the reactor for multiple turnovers. The homogeneous solution reagents and hydrogen

flowed into the reactor, and products flowed out of the reactor, but the solid catalyst particles were maintained within the vessel by filtering in situ. Overall catalyst used was less than the batch process, but the molar equivalent of catalyst in the reactor was higher compared batch because the catalyst was sequestered for about 25–50 reactor volume turnovers. Average time per reactor intermittent flow in and out of the reactor was 17 min; 1.9 kg advanced intermediate was produced with 107 automated turnovers of a 1 L autoclave reactor. Furthermore, part of the campaign was run with less feed solvent which resulted in a slurry feed. The intermittent flow stirred tank hydrogenation reactor enabled the slurry feed, which would not have been possible with a truly continuous packed catalyst bed reactor.

In general, intermittent flow stirred tank reactors (automated repeating batch) are similar to truly continuous reactors in many ways. Fast heat-up and fast cool-down is achieved in heat exchangers, while material flows in and flows out of the reactor. Intermittent reactors have higher heat transfer A/V compared to standard batch. Heat-up time, cool-down time, and time at reaction temperature is more scalable than standard batch. Intermittent reactors remain at approximately constant temperature and pressure at all times. They achieve a large number of turnovers per day. They are small compared to batch for the same throughput. If reaction time is 30 min, then the intermittent stirred tank is about 100 times smaller reactor volume compared to batch (assuming 48 h start to start cycle time for batch campaigning). Intermittent reactors are small enough so that they fit in lab hoods for 5–10 kg/day processes. The reagent feed tanks gradually empty and the product tanks gradually fill over time. These should be switched about once per day or once per shift.

In addition, there are many benefits of intermittent flow stirred tank reactors compared to PFRs. They can handle heterogeneous reactions with solids in flow, and two-phase liquid–liquid with long reaction times (τ longer than 10 min). They facilitate a much wider range of reagent addition strategies, for example, all-at-once addition, controlled addition of one or more feeds, any order of addition of multiple feeds, or co-addition, depending on which gives higher yield and/or minimizes key impurities (PFR is mainly all-at-once stoichiometric addition at the reactor inlet). They can run closer to end of reaction conditions, for example, remove and add back 10% of reactor volume each cycle, if it benefits impurity profile.

Furthermore, there are obvious benefits of intermittent flow stirred tank reactors compared to true CSTRs. They require much less reaction time for the same conversion, and thus smaller reactor volumes, for positive order reactions. They can operate with all-at-once reagent addition, controlled addition of one or more reagents, or co-addition, depending on which gives higher yield and/or minimizes key impurities (CSTR is only co-addition).

4.2 Recycle

There are many potential benefits of incorporating recycle into a continuous process. Continuous reaction with recycle of unreacted reagents after a downstream separation step can improve overall yield and selectivity of some chemical transformations. Late forming impurities can be avoided, while still maintaining high yield,

by running at incomplete conversion and recycling reagents back to the reactor inlet (White et al. 2014). For equilibrium limited reactions, recycle can be used to achieve higher overall conversion by removing a product or by-product and recycling reagents to drive the reaction forward. If there is a selectivity advantage of partial conversion, then recycle keeps the reaction at the conversion where the ratio of desired to undesired product is highest, yet maintains high yield by recycle of reagents. If there is a benefit of high stoichiometric ratio of reagents in a continuous reactor, then recycle can be used to achieve high relative stoichiometry in the reactor by separating downstream and recycling the reagent used in excess. This can be accomplished without using a large excess of reagent overall. Recycle can be used to achieve lower overall catalyst loading but higher instantaneous catalyst loading by recycling a homogeneous catalyst (Tsukanov et al. 2016). Finally, recycle is well known to achieve lower process mass intensity by recycling solvent, and achieve higher overall yield in by recycling product from filtrate. Process Mass Intensity (PMI) is defined as mass waste generated divided by mass product. Recycle could be one of the most powerful aspects of continuous processing compared to batch. Chemical engineering texts on material and energy balances often teach recycle incorporated into process flowsheets for continuous processes (Felder and Rousseau 1986). It is the key to achieving higher yield at the same purity, or the same yield with higher purity, and minimizing waste, compared to batch.

4.3 Common Misconception About Needing Fast Reactions in PFRs

Kinetics do not necessarily need to be fast in order for a reaction to be a viable candidate for continuous processing. It is common practice to speed up reaction rates by operating at elevated temperatures in continuous reactors; however, there may be a trade-off between accelerating reactions and thermal stability of reagents and products. Reaction selectivity and impurity profile often suffer at higher temperatures.

Table 4.1 lists continuous reactions that have been scaled up by Eli Lilly and Company to pilot and plant scale, most of them in GMP production. Compared to what is seen in the majority of literature on continuous chemistry, the examples listed in the table show long PFR reaction times (0.7–24 h) and large PFR reactor volumes (3–360 L). The table lists the impurity issues that become more significant when the reaction is run at higher temperatures. As seen in the last column in the table, isomeric impurities, dimers, chiral impurities, degradation of product, and other impurities may result, if reaction temperature is increased in order to reduce τ.

The imidazole cyclization with 90 min τ (May et al. 2012) and the hydrazine addition with 90 min τ (Cole et al. 2017a) were both run at moderate temperature to minimize deprotection of product (Table 4.1). The thermal EE deprotection with 40 min τ (Reizman et al. 2019; Cole et al. 2019; Frederick et al. 2015) was run at moderate temperature to avoid hydrolysis of an amide bond. The S_NAr reaction was

Table 4.1 Continuous reactions with long τ that have been scaled up in PFRs by Eli Lilly and Company to pilot- and plant-scale production

Reaction in PFR	τ	Material produced	PFR vol.	T	P	Issues at higher temperatures (and shorter τ)
Imidazole cyclization	1.5 h	29 kg GMP	7 L	140 °C	69 bar	Deprotection of product
Thermal deprotection of ee group	0.7 h	183 kg GMP	7 L	170 °C	25 bar	Hydrolysis of an amide bond
Hydrazine addition	1.5 h	26 kg GMP	1.5 L	130 °C	20 bar	Deprotection of product
S_NAr	3 h	31 kg GMP	3 L	70 °C	1 bar	Dimer
Acid deprotection of boc group	4 h	24 kg GMP	12 L	25 °C	1 bar	t-butyl amide
Hydroformylation, Rh catalyst	24 h	178 kg	32 L	55 °C	70 bar	Lower isomer selectivity
Asymmetric hydrogenation, Rh catalyst	12 h	144 kg	73 L	70 °C	70 bar	Lower ee
Reductive amination, Ir catalyst	12 h	2000 kg GMP	380 L	25 °C	55 bar	Cis isomer
Asymmetric reductive amination, Ru catalyst	5 h	3000 kg GMP	200 L	130 °C	60 bar	Dimer

run at moderate temperature and 3 h τ (Cole et al. 2017a), and a direct asymmetric reductive amination was run at moderate temperatures and 5 h τ, to minimize dimer impurities (Changi et al. 2017). The BOC deprotection was run at mild temperature and 4 h τ to minimize a t-butyl amide impurity (Cole et al. 2017a). A hydroformylation was run at moderate temperature and 24 h τ to minimize formation of the undesired linear aldehyde isomer (Johnson et al. 2020a). An asymmetric hydrogenation was run at moderate temperature and 12 h τ to minimize an undesired enantiomer (Johnson et al. 2012). A reductive amination was run at moderate temperature and 12 h τ to minimize the cis isomer (May et al. 2016). These all lead to longer reaction times, larger required reactor volumes, and less need for fast mass transfer and heat transfer rates.

The last four entries in the table are homogeneously catalyzed reactions where there are significant cost, quality, and environmental reasons for minimizing the amount of catalyst/ligand, which can represent a significant part of the overall manufacturing cost. The loading of Rh, Ir, Ru is minimized for environmental stewardship, and also for reducing levels of these metals in the product for toxicology reasons. It is vitally important to sufficiently remove metals from pharmaceutical products, and this can require additional separation steps. The best option is to put less metals in the process in the first place by using high substrate-to-catalyst ratios, which results in long (>12 h) reaction times and large reactor volumes, for example, the 360 L PFR in GMP manufacturing for homogeneously catalyzed hydrogenation (May et al. 2016).

In addition, a practical reason for not accelerating the reactions by going to higher temperatures in some circumstances is process fit. When these reactions are integrated into a fully continuous process, it is best if the reactions can tolerate flow stoppages or throughput changes. This is usually more feasible at the milder temperatures and longer target τ. This way, if the process must stop for an extended time period and then restart, risk is lower for holding material in the reactor.

Whether the reason for longer τ is to minimize impurities, improve process fit, or reduce catalyst and ligand loading, the downside is the need for larger volume PFR reactors. However, PFRs are generally low cost and the trade-off is favorable to install larger reactors.

4.4 Continuous Process Checklist

Prior to running a continuous process demonstration, the following questions should be answered.

Flows:

- Achieving and maintaining accurate and precise mass flow rates for each of your continuous feeds is one of the most critical aspects of flow chemistry. Do not start a continuous process unless you know that your mass flow rates will be accurate, consistent, and quantifiable. How do you know that your mass flow rates will be correct? Do catch and weighs. This means that you collect the liquid in a tared container for a measured amount of time, weigh the mass of liquid collected, then use the information to calculate mass flow rate. This can also be done if feed solutions and product solution are continuously collected on data logging balances, by calculating change in mass versus time. Design your system in a way that you will have redundant measures of mass or volumetric flow rates, so that you get double-checks on mass flowmeters. For example, the primary measurement is performed using a Coriolis mass flowmeter (because this is more useful for feedback control), and the secondary measurement is change in feed vessel mass or level versus time. Calculate real-time mass balances to make sure they remain 95–105%, for example, mass balance for the previous hour, by looking at change in mass of all inputs, outputs, accumulation, and generation over a given time period.
- What if one or more of the pumps gives oscillating flow rates for example, every few minutes? Does the process dampen this out sufficiently at reactor inlets to tolerate the fluctuations? You may need to include a small mixing pot or stirred vessel at the inlet to a PFR.
- If this is a reaction with gas reagent (e.g., H_2), then how do you know the real reaction gas flow rate? Verify the mass flowmeter reading by monitoring gas supply cylinder pressure versus time.
- How do you know that your check valves are going to prevent backflow? They won't. Check valves fail. Use them only as a backup line of defense.

- Have you done pump testing? For example, if you plan to use a peristaltic pump for a solution with THF solvent at 60 °C, then can you set up a pump-around recirculation loop with a small amount of the feed in an inerted box with secondary containment, and pump the feed for a long duration to make sure the peristaltic pump tubing will hold up to the hot solvent? It is better to test failure limits like this with small amounts of material in safe environments.
- How will you know if the instantaneous liquid and gas flow rates are steady? If there are oscillations in the liquid or gas flow rates, then how will you know how much they oscillated?
- Are your tubing and piping line sizes correct? If this is a technology transfer from one site to another, are any line sizes different than where the process was previously run? Oversized tubing can result in carryover and dead zones, while undersized tubing or fittings can result in high pressure drop and plugging.
- Are there any process or vent lines for which the inner diameters are too small? Look through every inch of your process. Typical examples: You want size of tubing and fittings on suction side of pumps to be larger than discharge side of pumps, preferably 3/8″ or larger for 100 L/day scale. You want process tubing and fittings to be at least ¼″ even at smallest research scale if there is any potential for solids moving through the tubing. You want the aqueous overflow from gravity decanters to be at least ½″ until after the siphon break, even at the smallest research scale, because of the impact of water surface tension.

Feed solutions:

- What is the composition of all feed solutions prepared batch? This includes weights and volumes of all components, and also the density and molarity of each of the feed solutions.
- Can you prepare feeds that are all homogeneous solutions at room temperature? This greatly simplifies the complexity of the continuous process and improves reliability. If not, then try to change solvents and concentrations so that you have homogeneous solution feeds at room temperature. A small amount of effort here to obtain solubility can save hours of effort later resolving fouling/clogging issues. The alternatives are to feed slurries or hot solutions, which are both more complex and difficult than homogeneous solutions at room temperature.
- Is there any reason why composition in feed tanks would change over time? For example, could you be losing solvent or volatile reagent to evaporation out the vent? This could cause stoichiometry to drift over time and the process to go out of spec. Make sure that you do not have a live nitrogen sweep in the headspace of the feed tank.
- Is the feed solution in the feed tank well mixed? For example, if charging multiple cans or drums of feed solution to the feed vessel, are they well blended so that concentration is uniform throughout?
- Is the feed solution stable? What is the stability of starting reagent solutions over time? The goal is at least 1 week stability at feed tank temperature.
- Is the feed solution filtered? What if in-line filters are clogging during the continuous run? It is best to filter all homogeneous solution feeds before you start

pumping them especially if you are using pumps with small clearances like gear pumps. In-line filters should still be used at the inlet of gear pumps even if the feed was already filtered before it was in the feed tank. Use dual in-line filters with valves so that you can switch to the second filter and clean the first when it starts to clog or blind, and switch back and forth.

- If in-line filters are installed, how do you know the pressure drop across them so that you can swap them out before they foul? Do you have a pressure indicator between the filter and the pump inlet?
- Do you have enough tanks, vessels, and cans for feed solutions, surge, and product solutions? Can you avoid filling and emptying a vessel at the same time? Can you avoid frequent manual refills? It is best if refilling is not more frequent than once every 12 h if you are doing it manually.
- Will you have enough materials? Are you making up enough solution volume for all feeds including excess of some feeds in case you need to change stoichiometric ratio? Which feed do you want to run out first, that is, what is the limiting reagent? Preparation of a small excess of inexpensive feeds is recommended. You don't want to run out of another feed before your limiting reagent is used up completely at the end of the campaign.
- Is a solvent feed tank ready to go, and tubing, valves, and Tee connected, in case you want to do a solvent pushout in the middle of the campaign? Is there a switching valve in line and ready for the easy feed swap?
- If you are ordering feed solutions from a vendor rather than mixing the solutions yourself, and if the solution needs to stay inerted, then does the vendor transport vessel have the needed fittings, valves, and connections so that you can transfer the solution out of the shipping container and into your plant feed vessel in an inert fashion?

Process Parameters and Data Collection:

- How will you measure to determine the real reaction temperature? Do you have redundant temperature measurements? Is temperature measured on the process side or shell/jacket side? How will you know the location of the hot spot in the reactor? Actual reaction temperature can be difficult to know and measure. Temperature measurements at the outlet of a reactor or unit operation can be inaccurate because of ambient cooling, even if they are only inches away from the jacket. Therefore, do not measure temperature at the outlet and assume that it is the same as the temperature inside the reactor. For PFRs, you may be able to use the shell side temperature and call it reaction temperature, but you must know heat transfer coefficients, reaction energetics, and kinetics so that you can calculate hot spots. What in-process temperatures do you plan to monitor? What is the number and location of in-process thermocouples (at the mixing Tees and also anticipated reactor hot spots?)
- How do you know that the actual reaction temperature at the hot spot will be the same when you scale up in a PFR?
- What will be the difference in jacket temperature (research vs. production scale) to keep reaction temperature the same when you scale up in a CSTR? For exam-

ple, if the reactor is a CSTR, then the difference between jacket temperature and reactor internal temperature will be larger when scaling up, because heat transfer surface area per to volume ratio (A/V) decreases when you scale up.

- How will you measure to determine the real reaction pressure? Do you have redundant pressure measurements? What is the pressure at the reactor inlet and the reactor outlet? Do you anticipate pressure gradually drifting over time due to fouling? Will you know if pressure drift and fouling is happening?
- How will you measure to determine the real reaction τ? Can you measure start-up transition curve and get τ from the midpoint of the F-curve? Can you calculate τ from reactor volume, % liquid filled, liquid flow rate, and thermal expansion of the liquid? What will be the impact of off-gassing in the reactor on τ? How do you know if the PFR will be completely liquid filled? How do you know if the pressure is high enough to prevent a gas phase? If the process exceeds critical temperature, then does it also exceed critical pressure? Actual reaction τ can be difficult to know and measure. τ is best measured during startup F-curve transitions from solvent to steady state.
- What is the difference between τ and V/Q for the continuous reactors? For example, for thermal deprotection in THF at 150 °C, we should say, "V/Q = 120 minutes and τ = 100 minutes because of thermal expansion." V/Q is more practical to specify for the plant operating ticket, because Q is the flow rate set point for the pumps, and the reactor volume should be known. However, τ is more important because it is the real reactor residence time. For example, if we want to compare reaction time in batch to reaction time flow, or if we want to measure reaction kinetics in the flow reactor, then we must use τ, the real mean residence time.
- What is the acceptable operating range for all of the important process parameters, for example τ, temperature, solvent volumes? Can you do anything to widen the acceptable ranges, for example, run at lower temperature if it allows you to have wider operating windows for τ because it makes product more stable to end of reaction conditions?
- How do you know the real reactor volume? This may be different than manufacturer specified, and it may be difficult to measure for a PFR because of the difficulty to get it 100% liquid filled and then completely empty.
- Does the liquid flow in the bottom and out the top of the coiled tube reactor? This is especially important for larger than 4 mm inside diameter tubing.
- How will you measure to determine the real reaction liquid flow rate? Do you have a secondary check in addition to a mass flowmeter for each stream, for example, change in mass or liquid level in feed tanks over time? Can you do catch and weighs? Do you know density of feed solutions?
- How do you know all the feeds were inerted and the reactor was inerted? Is the system leak free? Prove it with leak tests and pressure/vent or pressure/vacuum purges.
- How do you know that you will be able to calculate mass balance? Do you have enough vessels on weigh scales, vessels with level transmitters, or mass flowmeters to account for all inputs and outputs? How will you know if accumulation is happening? Is the data being collected, so that you will be able to close mass

balances for each lot? Are you manually weighing feed drums going into the process and product and waste drums coming out before and after they are filled or emptied, along with start and stop times for each drum?

- How do you know that experimental data for all important operating parameters will be collected and preserved? What is logged to the distributed control system (DCS) data historian (if available) and what is not? How will you retrieve the data? What do you need to write down hourly on checklists because it is not logged automatically? You can make a table of readings for the operator to write down at a specified frequency. How will that information be preserved?
- Is your numerical model predictive? What perturbations or step changes can be done during the continuous run to test if the model is predictive?
- What information are you planning to document after the scale-up continuous run? Are you set up to collect all of this information?
- Is there video of the entire process? If not can you get a video of the entire process during the demonstration?
- Do you expect anything to change gradually over time, like fixed catalyst bed activity? If so, do you have a planned frequency of change-out?
- Does anything need preconditioning or seasoning, like a new packed catalyst bed or metal walls of a reactor to be used for a catalytic reaction?

Scale-up

- What is the appropriate scale-up factor for a production campaign? This depends on intended duration of flow campaign. In this campaign, is it important to demonstrate longer time periods, for example, if using a packed catalyst bed reactor or running a continuous crystallization with long-term encrustation potential, or is it more important to generate product as quickly as possible?
- Are your process flowsheets and spreadsheet calculations representative for scale-up? Do you have a scale-up spreadsheet with calculations of all flow rates, volumes, masses? Or are you using a flowsheeting program?
- Have you done the engineering calculations on heat transfer rates, sizing heat exchangers, mass transfer rates, sizing mixers, line sizes, vessel sizes, agitation systems, filtration times, and settling times?
- If running continuous crystallization, how will you ensure that % supersaturation is constant with scale-up? What is the shear sensitivity of the crystals? What is the impact of agitation on particle size distribution, and impact of shear on secondary nucleation? How does it impact filtration rates and filter cake wash as a function of the attrition, and might attrition be greater at larger scale? Also, if running continuous crystallization, have the scaled-up vessels been properly designed for complete solids suspension mixing? Multiple flat baffles, vessel height to diameter ratio of 1.0 or less, and down-pumping impellers are recommended for solids suspension mixing.
- Are you planning to do a long continuous run 24 h per day with multiple unit operations simultaneously? If so, then it is best to do 12 h continuous demo runs beforehand, through small sections of the continuous train at a time, during the day shifts. This allows you to make sure everything is working correctly before

trying to start up the entire flow train for the long run. Scheduling people for 24 h per day coverage is a significant burden on resources especially if it is an R&D facility that does not normally work shifts around the clock. Make sure the 12 h daytime runs work out the bugs in the continuous train before beginning the 24 h coverage and committing larger amounts of materials.

- Do you need to run scale-up demonstrations in the lab before scaling up to the manufacturing facility? Here are 4 reasons why you might:

  1. Heat and mass transfer rates. Mass transfer changes because you cannot maintain all the same mixing parameters with scale-up. Heat transfer changes because heat transfer area/volume changes and fluid mechanics change with scale. Thus, heat transfer distances and film layers change at the boundaries. If scaling up a mixer-settler or centrifugal extractor, then what is the stage efficiency? How close does each stage operate to equilibrium? What is the relationship between τ, agitation rate, and scale on stage efficiency?
  2. Fouling/encrustation. Solids accumulating on surfaces depends on shear rate and fluid mechanics such as Reynolds number and linear velocity, which are differently scaled up, for example, in a continuous crystallization. Fouling also depends on blend time, which impacts local supersaturation levels, and mixing rate versus crystallization rate. Mixing rates and blend times change with scale-up. Another example is fouling and plugging in heat exchangers, which can be much worse scaled-up because the ΔT between jacket and process is usually higher in the scaled-up unit. ΔT must be higher when scaling up to overcome lower A/V, unless the heat exchanger maintains the same characteristic dimension as for numbering up micro-reactors.
  3. Technology development. If a reactor or unit operation is new and not tested at scale because it did not previously exist, then it may be best to test the scaled-up prototype before starting a GMP manufacturing campaign, where it is more difficult to make changes. This can often be done with short duration experiment for proof of concept and proof of sufficient heat and mass transfer rates, for example, 3–5 volume turnovers.
  4. Material production.

  Chemistry:

- Are you sure that the chemistry is going to work as expected (purity, yield) with the actual starting material lots and solvents? Have you run these exact planned conditions continuously previously? Can you run the reaction batch to confirm by use test, before starting the continuous process? Furthermore, after reagent solutions are made up in feed tanks for the continuous campaign in the plant, before starting flows, it is best to get a sample of the actual feed materials and do a batch (or small-scale flow) reaction off-line in the lab to confirm that the feed solutions are good to go.
- If the reaction uses a catalyst, then have you tested multiple lots of catalyst/ligand/support to make sure reaction results are consistent across a number of catalyst lots?

- How tight are the stoichiometry requirements if the continuous reaction mixes two reagent feeds together? What reagent is in excess, and how much excess is acceptable? It is best if there is an acceptable range on stoichiometry and the excess reagent target is set such that if one of the mass flows is off by 10% the reaction results will still be acceptable.
- Is a seasoning run warranted? Consider a brief reaction run to condition the reactor surfaces, especially if the reactor is a PFR.
- Is the flow system clean? What was used in the reactor previously, and how could it interfere with reaction? At research scale, it is usually best to install a new tube reactor if it is a new chemistry in a PFR. At production scale, it is best if the PFRs are disposable and dedicated to a specific chemistry. PFR tubes are typically inexpensive, and cost is negligible compared to the cost of repeating or delaying an experiment or production campaign.
- Was chemical reaction safety analysis completed? This includes ARC and DSC and calculations for heat removal. Were calculations done for worst-case scenario heat release for exothermic reactions? For example, if running a Grignard formation in CSTRs, did you make sure the reaction is dilute enough so that it is not possible for all the solvent to boil off in an exothermic event (latent heat of vaporization for the total solvent mass is greater than maximum heat of reaction)?
- Did you consider appropriate material compatibility, flammability, reactivity, reaction calorimetry, and thermal stability test data? Do you have appropriate materials of construction? How do you know that there will be no corrosion problems? Was coupon corrosion testing done for this reactor material of construction? All wetted parts of the reaction system including gaskets, o-rings, pressure reliefs, and gauges should be considered.
- What is the chemical stability of starting reagent solutions over time?
- Do you have data on stable hold points?
- Do you have solubility at room temperature where you need it or want it? For example, it is desirable to have solubility at room temperature for liquids flowing in/out of extractions, solvent exchanges, adsorption columns, in addition to reagent feed solutions.
- Is the product stable to end-of-reaction conditions in the reactor? If not then use a PFR, tightly control τ, and have solvent supply vessel ready to swap and valve and pump out the reactor contents in the event of a process stoppage.

Heat and Mass Transfer:

- How do you know that you will have sufficient heat transfer rates? Make sure heat transfer calculations have been done and you know heat-up and cool-down times and distances flowing through the tubes. Make sure heat exchangers are not undersized or oversized. What will be the hot spots for reaction exotherms predicted from the model of reaction rate, heat of reaction, and reactor heat transfer? Is it possible to insert thermocouples to measure the hot spots? Do you know overall heat transfer coefficients for CSTRs and can you remove the heat generated with the jacket alone, or is an internal cooling coil needed as well?

- Is sufficient heat exchange provided for process liquids flowing into and out of reactors or separators?
- If individual circulators are used instead of house cooling system, then do the circulators have enough cooling capacity to remove heat from reaction exotherms and maintain desired process temperature? What are the consequences if the circulators or cooling devices fail during the campaign? Are the safety interlocks in place for shutdown on high-temperature alarms?
- How do you know that mass transfer rates will be sufficient? Will mixing be sufficient when two streams Tee together? If a static mixer is utilized, then do you know the minimum recommended linear velocity through the static mixer for sufficiently fast mixing, given the viscosities and volumetric flow ratio of the two feeds? Are mixing zones or stirred tanks at the inlet to PFRs sufficiently large to dampen out flow fluctuations?
- If the process has separations like liquid–liquid extraction, then what is the relationship between τ, agitation rate, and scale on stage efficiency?

Residence Time Distribution (RTD):

- Is RTD known for each unit operation? CSTRs with τ more than 10 min and blend time less than 1 min typically can be modeled as ideal CSTRs, as long as inlet and outlet tube/pipe positions are sufficiently separated to prevent short-circuiting. If a PFR is used, F-curves should be measured during start-up transitions from solvent to reagents/products, to quantify RTD of substrate. Is the start-up transition time to steady state understood, and is the information used in the decision of when to start collecting product?
- How do you know what RTD will be overall for a continuous process train with multiple unit operations together in series? A numerical modeling software package may be needed.
- Will the planned data collection provide sufficient information to calculate lot genealogy? Documentation will include the time of switching reagents, switching parallel surge vessels, heels in feed and surge vessels, overall τ, and overall RTD.
- Does the numerical model of overall RTD facilitate the calculation of deviation boundaries?
- Is there a plan to quantify the potential difference between RTD of the actual product versus RTD of nonreactive tracers or solvents? These may not be the same for many reasons, including interactions with tube walls or adsorption/desorption from column packing materials.
- Are CSTRs strategically placed where process fluctuations or flow oscillations are expected? CSTRs serve to dampen out process fluctuations. If the reaction has fast kinetics and impurity profile at end of reaction conditions is favorable, then a CSTR is more forgiving to process fluctuations than a PFR. If reaction kinetics is not fast, then CSTR for initial stoichiometric mixing followed by PFR for full conversion may be a good option.
- The importance of knowing RTD is explained in another chapter in this book (Johnson et al. 2020b).

Stop/Restart:

- Is the process designed for ease of stop/restart if adjacent unit operations are having problems and need to be stopped temporarily, or in case upstream material was diverted to waste temporarily?
- Because of resource limitations, it may be preferable to run 16 h/day rather than 24 h/day in a production campaign. This is typically preferred in R&D, because it is easier on people, and it is advantageous to demonstrate that the process has sufficient stop/restart capabilities before going to manufacturing.
- If the process is stopped and restarted, then what automation interlocks might get tripped? For example, if there is an automation interlock based on low pressure, and if pumps are stopped, then does the low pressure alarm trip and do something like close an automated valve? If so, then make sure to reopen the valve before restarting.

Start-up and Shutdown

- Has the research-scale continuous process been started up with the exact same procedure and timing intended for production? This is important so that you know start-up transition will be as expected. For example, solid precipitates can form during start-up transitions and potentially clog lines, depending on timing feeds.
- When will you start collecting product, when will you switch collection tanks, and when will you stop collecting product? These should all be planned in advance. It may be advisable to switch collection vessels more frequently during start-up transition while you are deciding when to start forward processing, and then less frequently once the process reaches steady state. This approach can help minimize start-up transition waste.
- How long will it take to reach steady state? For drug substance, it can take days for a fully continuous multistep process train to reach steady state. Recycle increases the time to steady state. Also, some processes are always in transition, for example, packed catalyst bed reaction with drifting catalyst activity.
- Is it possible to achieve zero start-up transition waste for some of the sections to the continuous processing train? Will a surge vessel downstream from the reactor dampen out the impact of start-up such that the entire transition can be forward processed?
- What are the manufacturing complexity trade-offs? For example, you could start-up a countercurrent extraction train full of solvents rather than empty. This would cause start-up dilution and transition waste, but it would be much simpler than gradually filling each vessel in semi-batch fashion.
- Has a detailed plan for start-up and shutdown transition been created, and are the feed and product cans, tubing and valves set up so that you can follow the plan? For example, you may want to start the process flowing solvents only, then switch to reagent solutions at time zero. This means that you need additional feed vessels and you need switching valves installed. Also, the timing of switching each of the reagents from solvent may not be the same, depending on the time it takes

to flow from the pumps to the mixing Tee for each. The opposite order occurs at shutdown transition. You need a source of solvent ready to switch on-line to the pump inlet for solvent pushout at the end.

- Is the volume of all feed lines measured, and are the times that it takes for each feed to reach the mixing *T* at the reactor inlet known? If not simultaneous, then which reagent or catalyst feed solution do you want to reach the mixer at the inlet to the reactor first? For example, you may want the limiting reagent to reach the reactor inlet last.
- Do you know which reagent you want to run out first, and do you have excess of the other reagents?
- If you are running a PFR or continuous extractor, then do you plan to start with the process full of solvent and then switch over to reagents? Do you plan to do the opposite during shutdown for solvent pushout, meaning that you will switch from reagents to solvents and keep pumping at the same flow rates until the product is pushed out? If so, then can the downstream processing steps handle the start-up and shutdown dilution? For example, a downstream distillation may be able to normalize the concentration by stripping more solvent during start-up transition. If not, then do you need to divert reaction product solution that is not full strength? If you are planning to divert material that is not full strength, then are you using a PFR designed for low axial dispersion to minimize product loss?
- CSTRs can be started up with the reactor in semi-batch mode. Then, flows start when reaction reaches full conversion in the vessel. This method can eliminate start-up transition waste for the reactor. However, if this is the plan, then have the experiments been done to prove that semi-batch start-up mode meets product quality needs? Subsequently, semi-batch shutdown with gradual emptying of the CSTR can eliminate shutdown transition waste from the reactor. Does data exist to support this mode of operation. Also, do you have capabilities to pump down the CSTR at controlled rate during shutdown transition? If flow goes out a dip tube, then will the dip tube be gradually lowered, and if so, is it designed with the capability to gradually lower it without getting into the impellers?
- If you are running a continuous distillation, then plan to delay flowing out of the distillate bottoms until the evaporator reaches steady-state concentrations. The total mass pumped into the evaporator before starting outlet flows should be calculated ahead of time.
- If you are running a continuous crystallization in stirred tanks, one option is to start the process with a batch crystallization in the first stirred tank, then start flows when the vessel is filled and at its crystallization endpoint. Another option is to start with the first stirred tank filled with final crystallization solvent composition and fully seeded with solids closely resembling steady-state particle size distribution. How many stops and restarts do you expect for cleaning out encrustation? Do you have enough seed crystals on hand depending on your desired strategy and expected restart frequency?
- Are manual bypass valves installed where needed for start-up and shutdown transitions? For example, if you expect a small amount of solids precipitation or solids eluting from packed beds during start-up transition but not at steady state,

then do you have two three-way valves installed so that you can bypass on-line HPLC samplers, IR probe flow cells, or other sensitive or easily fouled process components until the process reaches steady state? Process analytical technology (PAT) is useful during start-up transition because it helps to determine time to steady state and help decide when to start collecting product, but if it leads to solids fouling then it may be better to temporarily bypass it.

- If you are running continuous crystallization in MSMPRs, how many system volume turnovers are required to reach steady-state concentration, and how many turnovers are required to reach steady-state crystal size distribution (CSD)? This could be determined by sampling once per turnover and analyzing with something like a Malvern, or it could be measured with an in situ Focused Beam Reflectance Measurement (FBRM). What is the target percentage supersaturation in all vessels at steady state, what was it in the research and development runs, and how will you measure supersaturation in the production run? One option is an in situ probe like IR calibrated for concentration. Another option is sampling, immediately filtering the slurry sample, diluting the sample quantitatively so that product remains in solution, and quantifying potency, then comparing to equilibrium concentration at these conditions. What is the best way to start up the process? Is the first MSMPR started batch-wise before initiating flows? Is it seeded? If start-up is done batch-wise to establish slurry in the MSMPRs before starting flows, then what is the purity of the solids by HPLC or NMR at steady state compared to batch startup? What is the polymorph form of the crystals, CSD, and crystal shape for batch start-up versus steady state?
- How will you quench excess hazardous reagents like LDA or BuLi downstream from the reactor, or activated Mg at the end of processing? Have a plan to do this in a safe and controlled manner that does not generate high concentrations of hazardous gas by-product.
- What is the plan for waste disposal?

Diverting:

- Where are the best points for diverting? This is typically immediately downstream from a PFR, where RTD is narrow, because it minimizes the amount of material that must be diverted before reestablishing the process within specifications. Is RTD known for the system? Does a plan exist for when and why to divert? For example, calculating a rolling average for impurity concentration over a time period equal to downstream fill–empty surge vessels is a reasonable approach to calculating acceptable magnitude and duration of disturbances. Do you have a decision tree for diverting decisions, and will you be set up and ready to divert with valves and catch vessels if needed?
- What are the points in the system of broad residence time distribution? For example, CSTRs, MSMPRs, and surge vessels have broad RTD. It is best to divert before the point of broad RTD, because broad RTD usually results in more material being diverted before that part of the process has returned to steady state or within range. It is better to divert at points of narrow RTD immediately after the reactors.

- Where is the divert valve positioned relative to on-line LC sample point. It is best to insert a delay loop between the on-line LC sample point and the divert valve. For example, a simple highly plug flow coiled tube with high *L/d* ratio can be used for a delay coil. The delay coil should have τ greater than LC method plus result turnaround time (e.g., 30 min). This allows you to "see into the future" with your on-line LC rather than getting results too late to respond. Of course, the on-line LC is not actually seeing into the future, but it is measuring what the composition will be at the divert valve 30 minutes after the sample is taken.

Analytical:

- How do you know that analytical will be correct and accurate? Make sure that the analytical method is up and running and that you have verified it is measuring your product and key impurities correctly before starting flows. Will the process require analytical results in real time? Make sure that on-line and off-line analyses confirm each other. Periodically take samples for off-line analysis at the same time and from the same location as the on-line PAT. Will on-line HPLC be used? Do you have a manual sampling point at the same location as your on-line HPLC, so that you can confirm on-line results with an off-line sample?
- Do you have a backup analytical plan so that the process can continue to run if the PAT is down temporarily? For example, the decision may be to take manual samples for downstream surge vessels and analyze them off-line until the PAT is back on line. More material is at risk while the PAT is down, because if a downstream surge vessels must be diverted it is probably a larger amount of material to divert. However, all material is still analyzed before it is pumped into the next step.
- What will be the first off-line analytical sample of the continuous run during start-up transition? When will it be expected to be pulled from the process? How soon will you need it analyzed in order to respond with a process adjustment if needed? How quickly can you get it analyzed and results reported from the lab? Is the analyst ready and expecting the sample? How do you know whether or not the sample represented steady state? Is this known flat-lining of PAT, or by proving the same result with multiple off-line samples pulled at some frequency?
- Is the sampling plan realistic, that is, is it possible to keep up with the number and frequency of samples that you plan to take? Where will samples be placed once pulled? How will the samples be labeled? For any continuous process, it is best to include the date and time in the sample name, because then every sample is necessarily unique and you can look back at process parameters at that time? What off-line sample analytical will be needed near real time throughout the campaign? Do you have a unique way to label and a unique place to put those urgent samples? What is required analytical frequency at start-up versus at steady state?
- How do you know that the samples will be representative? The sample position must be designed so that it receives representative flows, that is, it is not in a dead leg or dead zone. How big are the waste cuts and how big is the dead leg in the sample zone? How should I quantify the volume or mass of each waste cut to

make sure it was sufficient? Justify why you believe that the samples are representative.
- What if the sample port clogs with solids? Have a plan, and install the proper valves so that the sample port can be unclogged without disrupting the process.
- It is possible for PAT to interfere with the process in a negative way? It is best if the PAT does not obstruct the process flows, for example, restrict flows or introduce valves in line with the main process flows. If using on-line HPLC, how do you know that you will not get diluent, quench, or derivatizing reagent back into the process? You should always have more than one automated block valve separating the process from these PAT-related solutions, in case one valve fails.
- If you are using a sample port, then does the process dip tube go to the bottom of the overflow tee (via bored through fitting) for manual sampling points and for on-line dilution cart automated sampling? The overflow Tee is a good way to get samples without the potential to interrupt process flows, but you must install the dip tube such that the sample zone is not a low dead leg.
- Did you install manual bypass valves so that you can bypass the probe or auto sampler? This may be important during start-up transition if you have solids issues, for example, the outlet of a packed catalyst bed reactor where solids fines can elute from the reactor for a time when a new bed is put on line. Furthermore, this may be important at any time in the long run if you need to swap out parts or do maintenance or troubleshooting on the PAT.
- Is there a plan for how to use PAT for troubleshooting? What could go wrong in the process and what would be the PAT response to look for? If conversion or selectivity of a continuous reactor changes, a typical question is this: What reagent feed was changed 1 τ before the composition at the reactor exit changed? For example, if the reactor has a 6 h mean residence time and conversion drops at 12:00, then you will look to see if one of the reagent or catalyst feed lots was switched at 06:00.
- If the process has batch collection vessels after a continuous reactor, then have you selected the size of collection vessels that best meets your needs for analytical testing and forward processing decision? For example, you may want 24 h collection vessels, so that you only need to do analytical for forward processing once per day. The decision is based on a balance between minimizing the required analytical frequency and minimizing the amount of product at risk at a given time (putting too many eggs in one basket).

Operations:

- Do you have written start-up, shutdown, and operating procedures? Is there a plant "ticket" written to give instructions for transitions as well as steady-state operation?
- How can you work out the bugs first in a solvent only run? Are the planned solvent run conditions identical to the process conditions that you plan to use in the real chemistry campaign?
- Do you have a plan for daily checks for long continuous runs? Check all things daily that you know could be problematic. For example, check liquid level in

circulators. If the liquid level gets too low because of evaporation or other loss of heat transfer fluid, then the circulator could shut down. Be proactive and keep them topped up. Think about all the heating and cooling utilities. Also, check the level in feed vessels that are filled infrequently (e.g., once per week). Check the amount of solids buildup in continuous crystallizers. Check pressure drop across in-line filters and replace if needed. Are vent bubblers still bubbling, indicating that vent system is remaining inert with nitrogen and that vent piping and reactor headspaces are sufficiently sealed? Check to see if manual nitrogen rotameters are still operating at desired settings. Check transparent tubing at the outlet from pressure relief valves, or vessels at the outlet from relief valves, to make sure that valves have not relieved and discharged material from the process. Do secondary calculations of mass flow rates by tank outages to verify that mass flowmeters are still accurate. Check to make sure there is no buildup of material in vent lines. Are filtration rates slowing, indicating that filter pads need to be cleaned?

- Do you have an operator checklist and data table prepared? What readings do you want recorded once per hour? For example, check liquid levels in vessels, temperatures and pressures not recorded by the DCS. Check DCS trends for pressures, temperatures, and levels that are recorded, including pressure at the suction side of all pumps. Check interface levels in mixer-setters. Check slurry transfer tubes for continuous crystallizers and filters. Look for signs of pump cavitation. Confirm that no feed vessels are about to run out.
- How frequently do you plan to switch product collection vessels? Even if the product collection vessel has capacity for several days, make sure to deliberately switch your product collection vessel at least twice per day during a long multi-day run. That way, if there is a process upset you have less material at risk in any one product collection vessel. Write your instructions with flexibility so you have the ability to switch to a new product collection vessel at any time.
- How will you prevent reactant accumulation, or total mass accumulation? Make sure vessels are not gradually filling if they are supposed to be at constant level. How will you prevent overfilling vessels? This especially applies to product collection vessels.
- Do you plan to run more than one continuous unit operation simultaneously? If so then it is better to run each of them individually first to work out the bugs before putting it all together in a real production campaign.
- Is there a plan for monitoring vent knockout vessels? It is best to monitor them automatically with alarmed level detectors or weigh scales, but they can be monitored manually if frequency is proceduralized and monitoring frequency is sufficiently high that they cannot overfill.
- What will happen if you temporarily lose power during the run? Would this be a safety hazard? What are the fail-safes? What items would be difficult on the restart after power is restored?
- Is everybody ready (engineering, chemistry, operations, automation, analytical), and do all know the start date and the planned people coverage? This is generally well known for real manufacturing campaigns, but it has the tendency to be less well communicated in laboratory development runs.

- Do you have enough people? For example, suppose that you intend to run a three-step fully continuous process in manufacturing, each step with reaction, separation, and purification unit operations. The start-up transition will require the most people, and the most technical expertise. You might need two process engineers, two process chemists, one automation engineer, two analytical chemists for on-line PAT, six operators around the clock during start-up transition to get the entire train up and running. After the entire process reaches steady state, the staff can decrease. The main point is that start-up transitions are resource intensive, and require the technical experts, and a mix of R&D people who developed the process and manufacturing plant personnel collaborating on the plant floor together.
- Make sure all vessels are inerted before flammable solvents flow into the vessel, especially if they flow into the vessel through nonconducting tubing?
- Grounding and bonding is necessary when flowing from one vessel to another to prevent static charge buildup and sparking. Avoiding the use of nonconductive heat transfer fluid with nonconductive piping in the heat transfer fluid pump around loops.
- Conduct process hazard reviews (e.g., What-if or HAZOP reviews, preferably involving multiple people). Hold pre-startup safety reviews to ensure equipment and controls were installed as designed, and to communicate critical safety issues to operations.
- Select a safe location for the flow experiment or manufacturing considering the number of people affected by a worst-case scenario.
- Program DCS calculations of real-time mass and energy balances for safety critical operations.
- Prevent plugging of vent lines, and prevent process materials flowing out vents. Prevent closed block valves in vent lines. Prevent backflow into feed lines or feed vessels, and prevent backflow of liquid from vent lines such as backflow of caustic from scrubbers. Prevent undesired phase changes in process lines like boiling or freezing.

Equipment:

- Have you operated the temperature control units for extended times prior to the real chemistry run? If it will be a multiweek continuous process run, then prove beforehand that all heat transfer units will maintain jacket set points for at least the length of the planned duration. For example, cooling circulators can gradually accumulate internal ice and lose cooling capacity over time. Make sure to get constant temperature baths adjusted to your desired operating temperatures the day before, if heat-up time is significant. For example, heat-up of 200 L water or oil baths used for submerging PFRs can take time depending on the power of your circulators.
- What equipment will be used for overall reactor system back pressure regulation? Is there a potential for solids during transitions? If so, then it may be best to use an expansion chambers in series back pressure regulator rather than diaphragm or spring-loaded regulators (Johnson et al. 2016).

- Is all of the equipment ready? Has it all been tested in a solvent run?
- Do you have redundancy where you may need it? For example, if you intend to run a long campaign with multiple continuous steps operating simultaneously, do you have redundancy of key temperature and pressure measurements, and on-line PAT instruments?
- Do you have equipment backups for pumps, peristaltic pump tubing, agitators, glass vessels and heads, automated open/close block valves, and temperature control units, all of which could fail or break in the middle of the campaign? Have a plan to stop flows to swap out equipment with minimal down time.
- Will electrical and utilities be sufficient to support all the portable flow equipment that must be hooked up? In particular, think about the electrical needs for heaters and chillers that can draw high amps and make sure there are enough high amp circuits, and that you do not have too many of these running off the same circuit.
- How do you know that the equipment will run without fail for the duration of the planned campaign? Off-line long-term reliability testing on key equipment before the real continuous chemistry run is recommended.
- Do you have a sufficient supply of fittings that are GMP approved and kept in a regulated storage facility such that you have backups and can replace process lines and fittings in case of leaks or fouling?
- Get started on specifying and ordering pressure relief valves at least 6 months in advance, and order backups of all pressure reliefs. Include pressure relief devices and/or auto-shutoff pressures on the discharge of all positive displacement pumps. Provide pressure relief devices immediately downstream from back pressure regulators (lower pressure set point to protect downstream equipment). Are all pressure reliefs compatible with their process stream?
- Do you have representative and complete process and instrumentation diagrams (PIDs)? Do the PIDs match the actual equipment in the plant or lab after it is set up, including all manual block valves? Have plant personnel walked the process lines referencing the PID?
- Does the equipment have sufficient venting so that vessels do not pressure up when filling (with vent paths free of block valves). Use separate vent headers when multiple vessels have incompatible headspaces. Provide vent knockout tanks to catch process materials in the event that vessels overfill and also catch bubbler liquid in case of suck back. Be especially careful with separation of liquids and vapors when a reactor is depressurizing and the gas/liquid mixture is flowing into a vessel and gas is exiting to a vent at the same time (e.g., use hydro cyclones to prevent misting out the vent).
- Is secondary containment installed where leaks, breaks, or overfills may occur? Examples are beneath pumps, bubblers, sample valves, and CSTRs.
- Include chemical sensors to detect leaks and hazardous concentrations, for example, hydrogen detectors in the tops of hoods where H_2 is used.
- Have all gaskets been checked for compatibility with the solvents in my process? For example, Viton® o-rings and gaskets can look like Kalrez®; therefore, if Kalrez® is needed for solvents like THF then check them all carefully. If the

wrong o-ring or gasket material is installed, and if it goes unnoticed during construction and assembly, then you should at least identify the problem during the solvent run before the real chemistry run.

- Does the equipment have alarms, interlocks and/or auto-shutoffs based on temperatures, pressures, and fill levels? Does equipment have emergency stops for power and automation fail safe?
- If this is a tech transfer, what is the difference between feed, surge, and product cans between the two locations? At the new location are the surge vessels and product receivers sufficiently sized so that the tank switching or tank emptying can be done only 1 or 2 times per day for each, if the intention is to do it manually?

Planned Duration of Scale-up Demonstration

- How long should the scale-up demonstration be run before the manufacturing campaign? Run at least long enough for the entire process to reach steady state. This might only require a few hours. Run longer for special circumstances. Several examples are given below. Many of these will not need to operate 24 h per day, if they return to steady state quickly after stopping and re-starting, but the total cumulative run time without cleanout may need to be several days. If possible, design continuous reactions and separations unit operations that can easily be stopped, held for a time (e.g., 10 h), and restarted without any negative consequence. It is better for manufacturing. Therefore, in most demonstration campaigns, it is more valuable to prove that you can run the continuous process during the days and stop flows overnight, rather than operate 24 h per day, and it is much easier on people.

    1. If you are using a packed catalyst bed or CSTR with sequestered catalyst and activity decreases over time, then you need to run for the expected catalyst life, which could be several days.
    2. If you are running a process with recycle, then a long time could be required for impurities to build up to steady-state concentrations in the recycle loop, perhaps 100 h.
    3. If you are running a packed bed adsorption, then it could take more than 300 h for the mass transfer zone to take shape along the bed length and another >300 h for the mass transfer zone to move one MTZ length down the bed to see if it is spreading over time.
    4. If the process has fouling, plugging, crusting of solids over time, like continuous crystallization, cryogenic PFRs, or WFE distillation.
    5. If it is possible for material to accumulate in a rector or a unit operation, for example, precipitate of metal impurities in continuous Grignard formation reactor, if the Mg is not 100% pure.
    6. If you are running a continuous filtration and dissolve-off process, and other insoluble solids and tiny debris accumulate on the filter pads over time.
    7. If you are running a continuous extraction and you are interested in the insoluble materials that accumulate in a rag layer between the liquid phases over time.
    8. If you need to test PAT for drift over time due to crusting, fouling, instrument drift, or tube swelling.

Around the Clock Operation:

- Do you need to run 24 h per day or can you stop flows overnight and restart in the morning? Most of the reactions and separations unit operations can be operated during the day shifts in R&D. You only need to keep the process running 24 h per day if there is a negative consequence of stopping and restarting flows. For example, if it is not a stable hold point. Here are a few examples:

  1. Asymmetric hydrogenation, where PFR τ is 12 h and the product is not stable at end of reaction conditions (Johnson et al. 2012).
  2. Continuous cooling crystallization. You may not want to hold the slurries in the MSMPRs hot overnight because of stability and potential solvent loss. If you cool down the MSMPRs, then it will change particle size and CSD. Once you start the process backup, it might take 24 h to reach steady-state CSD again in all the MSMPRs. If this is the case, it is better to run 24 h per day and not stop flows until the end.
  3. Continuous crystallization with kinetic rejection of impurities (Johnson et al. 2012). In these examples, it is better to run 24 h per day and not stop flows until the very end of the campaign. When flows stop the stirred tanks will reach thermodynamic equilibrium, which is obviously not desirable for demonstrating a kinetic impurity rejection.

Automation:

- Note that experts in automation engineering are required for running fully continuous processes.
- What level of automation will be used? What will be automated by the plant-wide DCS and what unit operations will be automated by small local PLCs? Will a plant-wide control system be used, or individual unit operation controlled with surge tank decoupling?
- Are the parameters tuned for all feedback control loops?
- Are unit operations linked in series, and if so, is sufficient time allotted for automation testing before a campaign?
- Are the skids programmed the same for the manufacturing campaign as they were at research scale or pilot scale in R&D? This is especially important for sequenced operations, like intermittent flow evaporators and filters, as described previously in this chapter. If this is a technology transfer from one site to the next, and especially if the automation platform is different at the new location, did the personnel who ran the process at the previous site, or in development, consult with the automation engineers at the new site to make sure startup and shutdown transitions, ramp rates, and ratio control are programmed consistently with the previous process?
- How can the system be started up if stopped in the middle? Which stop/restart sequences will be automated, and which will be manual.
- Is there an overall control screen where multiple variables can be viewed at once?
- Allot time prior to the production campaign to make sure there are no bad channels on the cards, that solenoids are functioning, that transmitters are properly spanned, etc.

Plugging/Fouling

- What is the potential for plugging and fouling during the continuous production campaign and how will it be measured? What changes can be made to decrease the chances of fouling and plugging? For example, if you are running a continuous crystallization, then operating at lower supersaturation, using a larger number of vessels in series so that less product needs to come out of solution in each vessel, installing slurry transfer tubing subsurface, and locating the inlet away from agitator shaft, walls, or baffles may all minimize solids fouling.
- Is tube size large enough to minimize plugging if you suspect you could have solids? Can you use high-velocity intermittent flow from vessel to vessel for slurries, as described previously in this chapter?
- Can you monitor for plugging and fouling visually? This is facilitated by using transparent glass vessels and PFA tubing for slurry transfers if the reactor, crystallizer, extractor, or evaporator runs at low pressure. If nontransparent metal tubing or vessels are used for high-pressure operations, then can you monitor for plugging with pressure transmitters? Can you provide mitigations for plugging and fouling (e.g., double block and bleed valves) so that you will not need to use a wrench on a clogged line under pressure?
- Have you installed pressure gauges and transmitters for measurement of pressure versus time at pump outlets (provides indication of gradual fouling)?
- What are the most likely unit operations for solids fouling? For example, continuous crystallization, filtration, or cryogenic reaction have fouling potential, so these should be a main focus during preparations.
- Does the process have any heat traced lines because the flowing stream would not be a homogeneous solution at room temperature? If so, is every inch of the process line sufficiently heated? Small "cold" spots can plug over time or when there is a flow stoppage.
- What are the smallest flow restrictions for process materials? For example, flow restrictions may be needle valves, small fittings, restricting orifices for mass flow control or back pressure control. These will plug first.
- What are the parts of the system where you may not have solubility at lowest possible process temperatures? What are the locations of potential supersaturation?
- Does the process have a continuous liquid/vapor separator downstream from a back pressure regulation system? If so, then plugging in vent lines from misting process solutions with dissolved substrate may be a concern. In this circumstance, design the process to vent slowly, design cyclones for V/L separation, use large vent lines, and avoid vent restrictions.
- If you are running continuous crystallization, then what is the expected rate of fouling with solids over time? How does fouling and crusting change with scale-up? Where are the most likely locations for solids fouling, for example, on the agitator shaft, on the baffles, at the location of feed entry into the vessel, in the tubing between vessels? How does the level of supersaturation impact fouling rates? How long do you plan to run the process at steady state before scheduled

periodic shutdowns for solvent cleanouts? Is the plan to operate the continuous crystallization until the operations staff makes a judgment call on when it is time for a cleanout? Do the fouling solids on the walls have the same purity as the flowing crystals? What is the chemical stability of the solids on the walls, in case some break off and flow downstream? This can be tested by getting a sample of solids stuck to a wall or baffle and analyzing them for purity. Where are the inlet points, where product solution and anti-solvent flow into the stirred vessel? It is best if they are on opposite sides of the vessel. Are the inlet tubes located at a sufficient distance away from impellers, shaft, baffles, and wall? Do they mix in a static mixing device in a pump-around loop? What is the mixing rate at these addition points?

- How do you plan to unclog a clogged line? For example, use back-and-forth pressure pulses or back-and-forth pumping as a first attempt to free the clog, before using a wrench to open up process equipment.

GMP related:

- Do you intend to have any start-up and shutdown waste? Or, do you intend to start-up and shut down in a manner that all products meet specifications and thus there will not be any transition waste?
- What will be the cleaning procedures at the end of campaign and after a process upset? How will you clean a long PFR that is difficult to inspect internally? It may be better to use inexpensive tubing and dedicate the PFR to a specific chemistry, disposing it when it is no longer needed for that specific product.
- How will you calculate lot genealogy? Do you have a plan to collect sufficient information on τ and RTD to calculate lot-to-lot carryover?
- Do you have a plan for deviation management (deviation boundaries—operating space, design space)?
- If this will be a validation campaign, then what is the required number of start-ups and shutdowns, required batch size and required steady-state operation time?
- What is the required analytical sampling frequency? Is the analytical sample for controlling process parameters taken from a point of narrow RTD, and is the analytical sample for forward processing decision taken from a point of broad RTD?
- What are the process parameters that you will use for feedback control (automated PID control loops based on T, P, flow, level, versus analytical PAT)?
- Do you intend to use on-line PAT for real-time information, or real-time forward processing, or real-time feedback control, or real-time release, or some combination of these? Do you need redundancy in case the PAT instrument stops working?
- What is an acceptable degree of variability and fluctuations of process parameters?
- How will in-spec process parameter adjustments be made? In other words, if no material is out-of-spec yet, but adjustments within the acceptable processing range are being made to process parameters to adjust closer to target set points, then how do you know that each operator would make the same adjustment? Use decisions trees or automation.

- What are the points in the continuous unit operations for diverting flow in the event of a deviation?
- Do you have any deliberately non-steady-state unit operations, such as fixed catalyst bed reaction where catalyst activity changes over time (e.g., 300 h) or fixed bed adsorption where solid phase becomes exhausted over time (e.g., 600 h). What if the continuous unit operation has deactivation over time?
- Where are you building in redundancy for processing equipment?
- What if the continuous unit operation has known fouling over time? What is the plan for periodic off-line solvent cleaning and flushing to remove solids buildup?
- Will this process deliver flexible batch size or fixed batch size?
- How will you document material movements from one part of the continuous train to another?
- Does the process include recycle? If so, how will you monitor for impurities building up in recycle loops?
- Will the GMP processing instructions and master batch record be handled differently than batch?
- What will be the material hold times?
- How can you use the numerical model to determine when to divert flow, or if diverted material is unacceptable?

Safety:

- See the continuous drug substance processing safety chapter in this same book (Johnson and Niemeier 2020).

References

Adams TA, Pascall A. Semicontinuous thermal separation systems. Chem Eng Technol. 2012;35(7):1153–70.

Changi SM, Yokozawa T, Yamamoto T, Nakajima H, Embry MC, Vaid R, Luciani CV, Wong SW, Johnson MD, Moher ED. Mechanistic investigation of a Ru-catalyzed direct asymmetric reductive amination reaction for a batch or continuous process scale-up: an industrial perspective. React Chem Eng. 2017;2:720–39.

Cole KP, Campbell BM, Forst MB, McClary Groh J, Hess M, Johnson MD, Miller RD, Mitchell D, Polster CS, Reizman BJ, Rosemeyer M. An automated intermittent flow approach to continuous Suzuki coupling. Org Process Res Dev. 2016;20(4):820–30.

Cole KP, Groh JM, Johnson MD, Burcham CL, Campbell BM, Diseroad WD, et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow cGMP conditions. Science. 2017a;356:1144–50.

Cole KP, Johnson MD, Laurila ME, Stout JR. An automated repeating batch with catalyst recycle approach to nitro group hydrogenolysis. React Chem Eng. 2017b;2:288–94.

Cole KP, et al. Small-volume continuous manufacturing of merestinib. Part 1. Process development and demonstration. Org Process Res Dev. 2019;23:858–69.

Felder RM, Rousseau RW. Elementary principles of chemical processes. New York: Wiley; 1986.

Frederick MO, Calvin JR, Cope RF, LeTourneau ME, Lorenz KT, Johnson MD, et al. Development of an NH4Cl-catalyzed ethoxy ethyl deprotection in flow for the synthesis of merestinib. Org Process Res Dev. 2015;19(10):1411–7.

Johnson MD, Niemeier J. Safety guidelines for continuous processing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020. (this volume).

Johnson MD, May SA, Calvin JR, Remacle J, Stout JR, Diseroad WD, Zaborenko N, Haeberle BD, Sun WM, Miller MT, Brennan J. Development and scale-up of a continuous, high-pressure, asymmetric hydrogenation reaction, workup, and isolation. Org Process Res Dev. 2012;16(5):1017–38.

Johnson MD, May SA, Haeberle BD, Lambertus GR, Pulley SR, Stout JR. Design and comparison of tubular and pipes-in-series continuous reactors for direct asymmetric reductive amination. Org Process Res Dev. 2016;20(7):1305–20.

Johnson MD, et al. Continuous reactors for pharmaceutical manufacturing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020a. (this volume).

Johnson MD, et al. Understanding residence time, residence time distribution, and impact of surge vessels. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020b. (this volume).

Kopach ME, Cole KP, Pollock PM, Johnson MD, Braden TM, Webster LP, Groh JM. Flow Grignard and lithiation: screening tools and development of continuous processes for a benzyl alcohol starting material. Org Process Res Dev. 2016;20:1581–92.

May SA, Johnson MD, Braden TM, Calvin JR, Haeberle BD, Jines AR, et al. Rapid development and scale-up of a 1 H -4-substituted imidazole intermediate enabled by chemistry in continuous plug flow reactors. Org Process Res Dev. 2012;16(5):982–1002.

May SA, Johnson MD, Buser JY, Campbell AN, Frank SA, Haeberle BD, et al. Development and manufacturing GMP scale up of a continuous Ir-catalyzed homogeneous reductive amination reaction. Org Process Res Dev. 2016;20(11):1870–98.

Reizman BJ, et al. Small-volume continuous manufacturing of merestinib. Part 2. Technology transfer and cGMP manufacturing. Org Process Res Dev. 2019;23(5):870–81.

Tsukanov SV, Johnson MD, May SA, Rosemeyer M, Watkins MA, et al. Development of an intermittent-flow enantioselective Aza-Henry reaction using an arylnitromethane and homogeneous Bronsted acid-base catalyst with recycle. Org Process Res Dev. 2016;20:215–26.

White TD, Berglund KD, Groh JM, Johnson MD, Miller RD, Yates MH. Development of a continuous Schotten–Baumann route to an acyl sulfonamide. Org Process Res Dev. 2012;16(5):939–57.

White TD, Alt CA, Cole KP, Groh JM, Johnson MD, Miller RD. How to convert a walk-in hood into a manufacturing facility: demonstration of a continuous, high-temperature cyclization to process solids in flow. Org Process Res Dev. 2014;18:1482–91.

Zaborenko N, Linder RJ, Braden TM, Campbell BM, Hansen MM, Johnson MD. Development of pilot-scale continuous production of an LY2886721 starting material by packed-bed hydrogenolysis. Org Process Res Dev. 2015;19(9):1231–43.

Chapter 5
Continuous Crystallization: Equipment and Operation

Yiqing C. Liu and Zoltan K. Nagy

Abstract Crystallization is an economical separation and purification unit operation commonly used in the pharmaceutical industry as the last drug substance manufacturing step to obtain crystalline form active pharmaceutical ingredient (API). The quality attributes of crystallization products such as purity, crystal size, crystal shape, and polymorphic form heavily influence downstream processes and may even affect the final product performance. Crystallization is one of last missing links in end-to-end continuous pharmaceutical manufacturing development because of its complex two-phase and stochastic nature. Recent advances and remaining challenges in the development of laboratorial- and industrial-scale continuous crystallization equipment, techniques, and control strategies are discussed in this chapter accompanied by a brief overview of crystallization theories as well as process analytical technology (PAT) applications in crystallization processes. The integration of continuous crystallization is also discussed in this chapter.

Keywords Crystallization · Continuous · Process analytical technology (PAT) · Mixed-suspension-mixed-product removal (MSMPR) · Oscillatory baffle reactor (OBR) · Plug flow crystallizer (PFC) · Oscillatory baffled crystallizer (OBC) · Filtration · Process intensification

Abbreviations

AFC	Antifouling control
API	Active pharmaceutical ingredient
ATR-FTIR	Attenuated total reflection-Fourier transform infrared

Y. C. Liu
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Z. K. Nagy (✉)
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Department of Chemical Engineering, Loughborough University, Loughborough, UK
e-mail: zknagy@purdue.edu

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series x1,
https://doi.org/10.1007/978-3-030-41524-2_5

ATR-NIR	Attenuated total reflection-near-infrared
ATR-UV/Vis	Attenuated total reflection-ultraviolet/visible
CFC	Continuous filtration carousel
CFD	Computational fluid dynamics
CIP	Cleaning in place
COBC	Continuous oscillatory baffled crystallizer, also known as OFBC
CSD	Crystal size distribution
DBC	Dynamic baffle crystallizer, also known as OBR
FBRM	Focused beam reflectance measurement
GC	Gas chromatography
HPLC	High-performance liquid chromatography
MOM	Method of moment
MS	Mass spectrometry
MSMPR	Mixed-suspension-mixed-product removal
MSZW	Metastable zone width
NMR	Nuclear magnetic resonance
OABA	Ortho-aminobenzoic acid
OBC	Oscillatory baffled crystallizer
OBR	Oscillatory baffle reactor, also known as DBC
OFBC	Oscillatory flow baffled crystallizer, also known as COBC
PAT	Process analytical technology
PBM	Population balance model
PFC	Plug flow crystallizer
PSD	Particle size distribution
PVM	Particle vision measurement
RTD	Residence time distribution
SASR	Solvent-to-antisolvent ratio
SEM	Scanning electron microscopy
STC	Stirred tank crystallizer
UPLC	Ultra-performance liquid chromatography
XRD	X-ray diffraction

Nomenclature

A	[m^2]	Heat transfer area
B	[#·s^{-1}·m^{-3}]	Secondary nucleation rate
b	-	Secondary nucleation order
C	[kg/m^3]	olution concentration
C_{in}	[kg/m^3]	Inlet solution concentration
C_p	[J·K^{-1}·kg^{-1}]	Specific heat
D	[m]	Reactor diameter
d_c	[m]	Crystal size
ΔC	[kg/m^3]	Absolute supersaturation
ΔH_c	[J/kg]	Heat of crystallization

D_{im}	[m]	Impeller diameter
D_p	[m]	Mass mean size of particles in suspension
d_s	[m]	Seed size
E_b	[J/mol]	Secondary nucleation temperature dependence
E_g	[J/mol]	Crystal growth temperature dependence
f	[#/m^4]	Population distribution of particles
f_{in}	[#/m^4]	Population distribution of particles at the inlet
G	[m/s]	Growth rate
g	-	Crystal growth order
g_c	[m/s^2]	Gravitational acceleration constant
h	[W·K^{-1}·m^{-2}]	Heat transfer coefficient
J	[#·s^{-1}·m^{-3}]	Primary nucleation rate
k_b	Empirical	Secondary nucleation kinetic constant
k_g	Empirical	Crystal growth kinetic constant
k_j	Empirical	Primary nucleation kinetic constant
k_v	-	Shape factor
L	[m]	Crystal characteristic length
L_0	[m]	Nuclei size
M	[kg/m^3]	Solid (magma) density
m	Empirical	Secondary nucleation solid concentration dependence
m_c	[kg]	Crystal mass
m_s	[kg]	Seed mass
N_{js}	[rad/s]	Minimum rotational speed required for suspension
Q_{agg}	[#·s^{-1}·m^{-4}]	Agglomeration mechanism
Q_{break}	[#·s^{-1}·m^{-4}]	Breakage mechanism
$Q_{\cdot i}$	[m^3/s]	Volumetric flow rate out of the ith MSMPR
$Q_{\cdot in}$	[m^3/s]	Inlet volumetric flow rate
Q_{nuc}	[#·s^{-1}·m^{-4}]	Nucleation mechanism
$Q_{\cdot out}$	[m^3/s]	Outlet volumetric flow rate
R	[J·mol^{-1}·K^{-1}]	Gas constant
Re_o	-	Oscillatory Reynolds number
S	-	Supersaturation ratio
s	-	Relative supersaturation
St	-	Strouhal number
T	[K]	Temperature
t	[s]	Time
Tw	[K]	Heat transfer wall temperature
V	[m^3]	Slurry volume
X	-	Weight percentage of solid to liquid content
x_o	[m]	Oscillation amplitude
z	[m]	Length along a PFR
Z_c	Empirical	Zwietering constant
δ	[m^{-1}]	Dirac delta function at L_0
μ	[Pa·s]	Viscosity
μj	-	jth moment of population distribution
ν	[m^2/s]	Kinematic viscosity

ρ_c	[kg/m^3]	Crystal density
ρ_{sol}	[kg/m^3]	Solution density
ω	[Hz]	Oscillation frequency

5.1 Introduction

Crystallization is a process where dissolved solutes transform from their solution phase to crystalline solid phase. It is a key unit operation present in the vast majority of pharmaceutical manufacturing processes to separate and isolate solid drug substance from its mother liquor. In oral solid dosage form production, it is often employed as the final purification step to obtain pure active pharmaceutical ingredient (API), serving the transitional role between drug substance and drug product manufacturing as shown in Fig. 5.1. Tailoring the crystal quality attributes such as crystal size distribution (CSD), shape, polymorphic form (multiple crystal structures of the same molecule, Fig. 5.2), and purity can have a significant impact on downstream processes as well as the quality of the final drug product (Lee et al. 2015; Zoltan K. Nagy, Fujiwara, et al. 2008; U.S. Food and Drug Adminiatration 2004; Wood et al. 2019; Yang et al. 2017). A crystalline particle exhibits a regularly arranged molecular structure which is often more stable than amorphous particles which do not have ordered molecular structures (Davey and Garside 2000). Typically carried out in batch, there has been, as of late, increasing interest in continuous crystallization development to emulate the progress made in developing continuous reaction chemistry and drug product operations to enable end-to-end continuous pharmaceutical manufacturing (Chatterjee 2012; U.S. Food and Drug Administration 2004, 2019). The intent of this chapter is to give a brief overview of crystallization theories and to introduce readers to laboratorial and industrial continuous crystallization equipment, techniques, and control strategies.

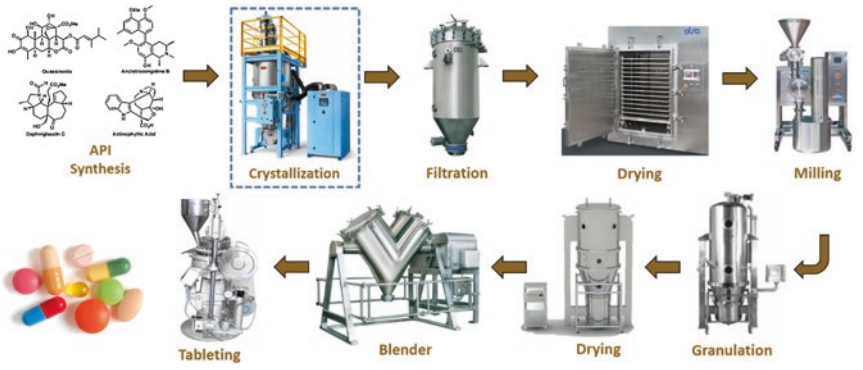


Fig. 5.1 Transitional role of crystallization in a typical manufacturing process of oral solid dosage form drugs

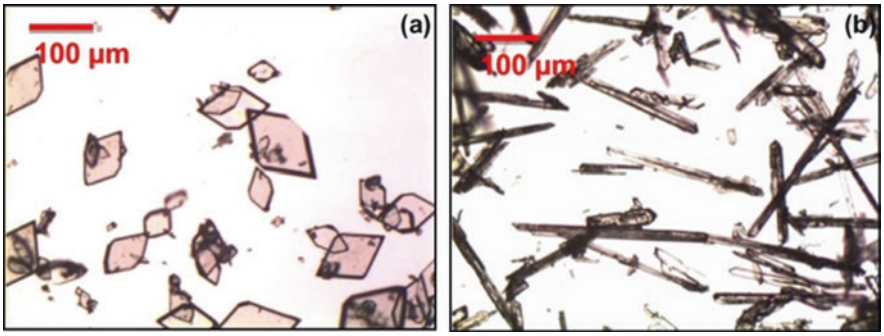


Fig. 5.2 An example of polymorphism: (**a**) prismatic form I of ortho-aminobenzoic acid (OABA) and (**b**) needlelike form II of OABA. (Permission obtained from Simone et al. © 2017 Elsevier Ltd)

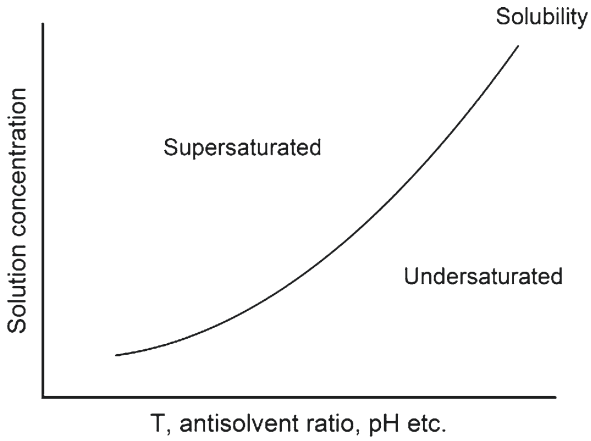


Fig. 5.3 Typical solubility curve and under–/supersaturated area

5.1.1 A Brief Review of Crystallization Theories

A typical solid solution-phase diagram is depicted in Fig. 5.3 where the equilibrium solution concentration (solubility curve) and the corresponding supersaturated and undersaturated areas are shown. During crystallization, the system starts undersaturated and must move across the solubility curve to become supersaturated. Supersaturated solutions are metastable: the system tends to retrieve back to equilibrium (solubility), but certain amount of supersaturation must be built up to initiate rapid crystal formation. A barely supersaturated system may remain "stable" for days, while a system situated deeper into the supersaturation area may become unstable in a matter of minutes resulting in crystallization. It is easy to understand that crystallization is a rate process driven by supersaturation whereas solubility is a thermodynamic property. Supersaturation may be generated by cooling, evaporation, antisolvent addition, reaction, pH manipulation, freezing, or a combination of

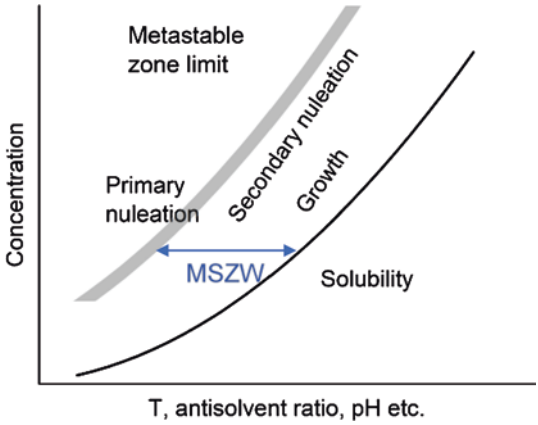


Fig. 5.4 The schematics of different zones for nucleation, growth, and dissolution in a typical phase diagram

the aforementioned methods, and the degree of supersaturation can be expressed as absolute supersaturation ΔC (defined as $C - C_{sat}$), supersaturation ratio S (defined as C/C_{sat}), or relative supersaturation s (defined as $\Delta C/C_{sat}$).

Crystallization is generally considered to be a two-step process where sufficient supersaturation must be accumulated resulting in births of crystals which then grow in size by incorporating solute molecules from the supersaturated solution. The birth of a crystal is called nucleation, and it consists of two mechanisms: primary nucleation, fast and spontaneous formation of nuclei without the presence of crystal particles, and secondary nucleation, birthing of nuclei triggered by existing crystal particles. As shown in Fig. 5.4, primary nucleation occurs near the thermodynamic metastable zone limit. For crystallization process design purposes, the supersaturation required for primary nucleation is often considered to be the metastable zone limit, while true thermodynamic metastable zone limit may be elsewhere which is difficult to measure due to the delayed detection of nucleation. The distance between solubility and the metastable zone limit in the phase diagram is called the metastable zone width (MSZW) which is an important guide for crystallization process design. It directly correlates to crystallization kinetics. MSZW is a kinetic property that can be affected by many factors including supersaturation generation rate (i.e., cooling rate, antisolvent addition rate, etc.), mixing dynamics, solvent properties, impurities, and solution history. Figure 5.5 illustrates the effect of cooling rate on the MSZW of aluminum potassium sulfate crystallization in water. Once primary nucleation takes place, the generated parent crystal particles will trigger secondary nucleation at a lower supersaturation as shown in Fig. 5.4. Because primary nucleation takes place at the brink of metastability, nuclei are generated in a stochastic, uncontrolled, and often undesirable manner. In contrast, secondary nucleation can be controlled by manipulating the parent particle's properties such as the size, polymorphic form, and load (amount). Therefore, in industrial batch crystallization processes, pre-generated particles are strategically added as "seeds" to bypass primary nucleation, regulate secondary nucleation, and promote growth. Sometimes nucleation may be preferred or required, which is often the case in continuous

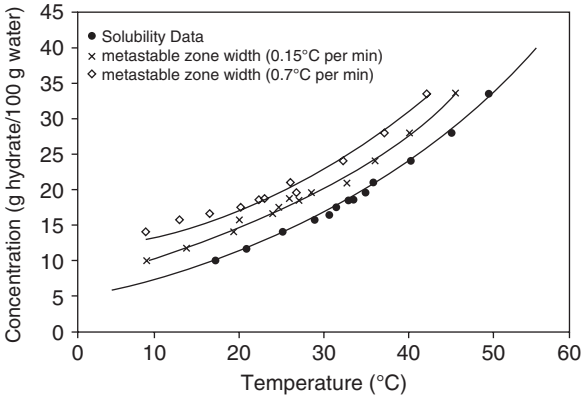


Fig. 5.5 Metastable zone width (MSZW) of aluminum potassium sulfate in water at different cooling rates reported by Barrett and Glennon. (Permission obtained, Copyright © 2002 The Institution of Chemical Engineers. Published by Elsevier B.V)

crystallization processes. In such cases, secondary nucleation inducing seeded operation is still preferred to promote secondary nucleation, while avoiding primary nucleation. While there are existing theories in attempts to explain primary and secondary nucleation mechanisms (Botsaris 1976; Garside and Davey 1980; Garside 1985; McCabe and Stevens 1951) such as *classical nucleation theory* (Becker and Döring 1935; Volmer 1939; Gibbs 1948), *two-step theory* (Erdemir et al. 2009; Chakraborty and Patey 2013; Davey et al. 2013), *dust breading* (Ting and McCabe 1934; Strickland-Constable and Mason 1963), *needle breading* (Strickland-Constable 1968), and *collision breading* (Strickland-Constable 1968), the exact mechanisms are not well understood (Mullin 2001b; Myerson 2002). A general rate expression does not exist for either primary or secondary nucleation. Instead, an empirical Arrhenius-type power law expression is often used to describe nucleation:

$$\text{Primary Nucleation}: J = k_j \exp\left(\frac{j}{T^3\left(\ln S\right)^2}\right) \tag{5.1}$$

$$\text{Secondary Nucleation}: B = k_b S^b \exp\left(-\frac{E_b}{RT}\right) M^m \tag{5.2}$$

where J and B are the primary and secondary nucleation rates, respectively, with the unit of number of particles per time per slurry volume, S represents supersaturation ratio and can be switched to absolute supersaturation ΔC or relative supersaturation s, T is the temperature, M denotes the solid concentration, R is the gas constant, and k_j, j, k_b, b, E_b, and m are kinetic parameters that can be estimated from experiments.

Growth refers to the enlargement of crystals, the rate of which significantly impacts the final CSD. Growth takes place at a lower supersaturation than nucleation in a more controlled manner (Davey and Garside 2000; Mullin 2001a). It is often preferred to generate larger crystals which can be isolated more efficiently compared to fine particles. In practice, techniques like high seed loading, small seed size, and slow supersaturation generation rate can be applied to promote growth and

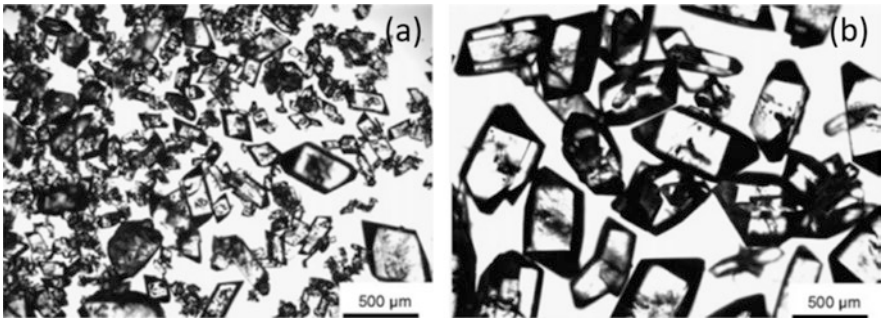


Fig. 5.6 Microscopic images of paracetamol product of (**a**) unseeded versus (**b**) seeded cooling crystallization. (Permission obtained from Nagy et al. Copyright © 2007 Elsevier Ltd.)

suppress nucleation to produce larger crystals. A comparison of the products of a growth-dominated, seeded batch crystallization experiment against an unseeded experiment is illustrated in Fig. 5.6 where seeded products are much larger in size and thus much easier to isolate. The mechanism of crystal growth is complex as crystal structures are complex: it is generally considered as a two-step process involving diffusion followed by the incorporation (equivalent to a reaction) at the surface, while other more complex two-dimensional growth theories have been explored (Bennema 1969; Burton et al. 1951; Frank 1949; Ghez and Gilmer 1974; Nyvlt et al. 1985; Ohara and Reid 1973; Wilcox 1971). Different models corresponding to these different theories have been developed mathematically, but a generic power law is often used in practice:

$$G = \frac{dL}{dt} = k_g S^g \exp\left(-\frac{E_g}{RT}\right) \mathrm{fn}(L) \tag{5.3}$$

where L denotes the characteristic size of the crystal, S represents supersaturation ratio which can be switched to absolute supersaturation ΔC or relative supersaturation s, and k_g, g, and E_g are kinetic parameters that can be estimated from experiments. Some function of L, fn(L), can be added to express size-dependent growth rate. It is important to note here that there are multiple ways to describe crystal size since crystals are three-dimensional structures. The characteristic length L is often defined as the volume equivalent sphere diameter (Fig. 5.7). Even though nucleation rate expressions Eqs. (5.1) and (5.2) and growth rate expression Eq. (5.3) are generalized expressions, they are all not completely empirical. They can be derived from classical nucleation theory and diffusion-reaction growth theory, respectively. In addition to nucleation and growth, other more complex phenomenon are often present during crystallization (Garside 1985; Mullin 2001c) such as agglomeration (multiple particles clustering together to form one large aggregate, Fig. 5.8a), breakage (one particle breaking into two or more particles, Fig. 5.8b), attrition (fines "chipping off" from a particle surface as individual particles, Fig. 5.8c), and Ostwald

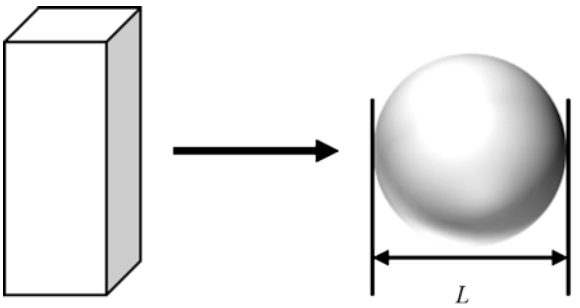


Fig. 5.7 Volume equivalent sphere diameter

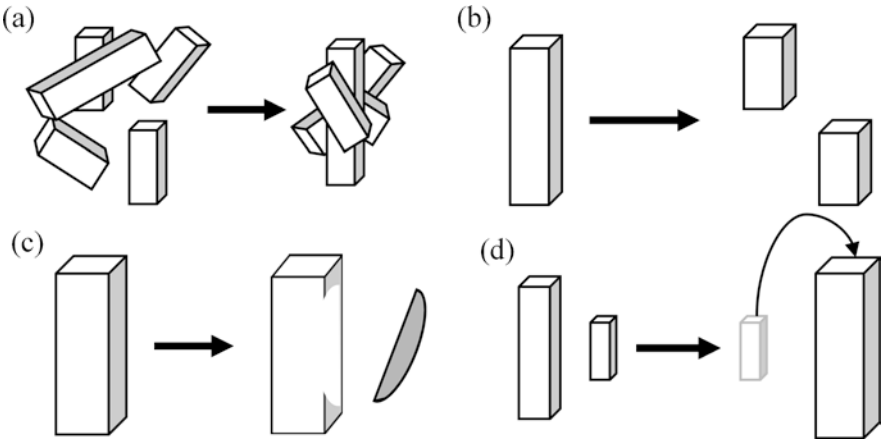


Fig. 5.8 Animated schematic of (**a**) agglomeration, (**b**) breakage, (**c**) attrition, and (**d**) Ostwald ripening mechanisms

ripening (fine particles dissolving despite supersaturation, Fig. 5.8d), which will not be discussed in detail in this chapter due to length constraints but can be important in industrial practices.

In order to achieve the desired crystal CQAs, one must design a suitable operating system, operating mode, supersaturation generation profile, and seeding strategy to manipulate nucleation and crystal growth (and sometimes agglomeration and breakage). If upstream processes allow, a proper solvent should be chosen to avoid high toxicity and increase solvent power so that changing temperature or antisolvent ratio results in higher changes in solubility. Crystal habits should also be considered when selecting a solvent to avoid difficult shape (e.g., needle shape), polymorphic impurity, and undesirable coloring (Gu et al. 2004; Karunanithi, Achenie, and Gani 2006; Kolář et al. 2002; Myerson, Decker, and Fan 1986). The selection of solvent is also coupled with choosing the type of crystallization such as cooling, antisolvent, or reactive crystallization for an optimal design (Chianese et al. 1984; Agnew

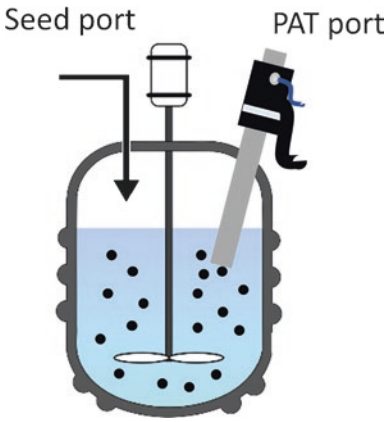


Fig. 5.9 Schematic of a typical batch crystallizer with overhead ports for PAT tools and seed addition

et al. 2016, 2017; Borsos et al. 2016; Chen et al. 2011; Frawley et al. 2012; Lindenberg et al. 2009). It is important to obtain the phase diagram and metastable zone limit so that an operating curve can be developed by experimental experience or optimization algorithms. Seeding is commonly practiced in industry to suppress nucleation and promote growth. Typically for batch crystallizations, 0.5–5 wt% seed loading is used and added when the system is close to solubility (¼ to ½ into the MSZW) (Bohlin and Rasmuson 1992; Aamir et al. 2010; Wey and Karpinski 2001). An empirical design equation can be used to select the seed size based on crystal size and shape (Davey and Garside 2000; Mersmann and Rennie 1970):

$$m_s = m_c \left(\frac{d_s}{d_c} \right)^i \tag{5.4}$$

where m_s and d_s stand for mass and size of seed, m_c and d_c denote mass and size of crystals, and i is a shape index, 1 for needles, 2 for plates, and 3 for cubes/spheres. In addition to operating conditions and seeding techniques, the equipment design is also crucial to implement a well-designed crystallization process. As shown in Fig. 5.9, a temperature-controlled vessel equipped with an overhead agitator and various ports for material addition and process analytical technology (PAT) tool installment is often used for crystallization in labarotary scale. PAT tools are important instrumentations that measure and monitor the process status and product quality. It plays a key role in improving pharmaceutical manufacturing, and they will be discussed in detail in Sect. 5.3. The material, dimension, and mixing scheme of the crystallizer must be properly designed to ensure efficient heat and mass transfer as well as effective suspension of particles.

The application of mathematical modeling can reduce the number of experiments needed to optimize a crystallization process. A population balance model (PBM) is often used to describe crystallization as it tracks the population of particles of different sizes in suspension as well as concentration in solution. The basic concept of a population balance is to balance the number of particles generated/

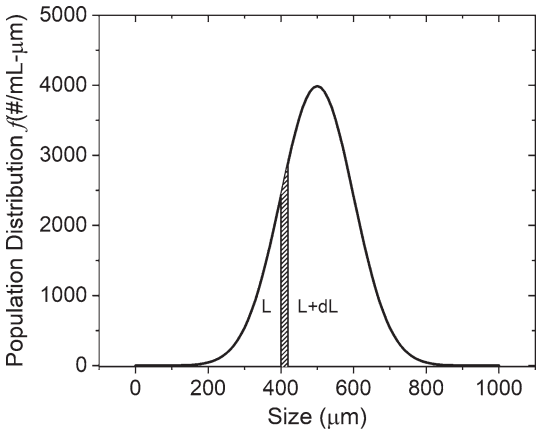


Fig. 5.10 Population distribution function where shaded area represents the number of particles between size L and $L + dL$ in a unit slurry volume

destroyed by nucleation, agglomeration, and/or breakage as well as to track the flux of particles travelling between size bins by crystal growth. To describe this idea mathematically, let us introduce a (univariate) population distribution function *f(L,t)* such that *f(L,t)dL* describes the number of particles between size L to $L + dL$ per unit slurry volume at time t (Fig. 5.10). For a batch crystallization process, population balance is written as:

$$\frac{\partial f}{\partial t} + \frac{\partial (Gf)}{\partial L} = Q_{\text{nuc}} + Q_{\text{agg}} + Q_{\text{break}} \tag{5.5}$$

where Q denotes particle generation or disappearance mechanisms and nuc, agg, and break subscripts stand for "nucleation," "agglomeration," and "breakage," respectively (Litster 2016; Ramkrishna 2000). If a seeded process is considered without agglomeration and breakage, Eq. (5.5) can be simplified as:

$$\frac{\partial f}{\partial t} + \frac{\partial (Gf)}{\partial L} = B\delta\left(L - L_0\right) \tag{5.6}$$

where $\delta(L - L_0)$ is the Dirac delta function that engages nucleation only at nucleus size L_0 (usually taken to be very small or 0):

$$\begin{aligned} &\int_{\infty}^{-\infty} \delta(L)\, dL = 1 \quad \text{if } L = L_0 \\ &\delta = 0 \qquad\qquad\quad \text{if } L \neq L_0 \end{aligned} \tag{5.7}$$

To close the system of equations, a mass balance can be written to relate solid- and liquid-phase concentration:

$$\frac{dC}{dt} = -3k_v\rho_c \int_{\infty}^{L_0} GL^2 f dL \tag{5.8}$$

where C is the mass concentration of the solute in solution, k_v represents the shape factor (1 for cube, 0.5 for sphere, 10 for needle shape, etc.), and ρ_c is the crystal density. An energy balance can also be written which is especially important for strongly exothermic or endothermic crystallization processes:

$$\frac{dT}{dt} = \frac{-hA\left(T - T_w\right)}{V\rho_{sol}C_p} - \frac{\Delta H_c 3k_v\rho_c}{\rho_{sol}C_p} \int_{\infty}^{L_0} GL^2 f dL \tag{5.9}$$

where h represents heat transfer coefficient, A is the heat-exchanging area of the crystallizer, T_w is the temperature at the heat-exchanging wall, V stands for the slurry volume, ρ_{sol} is the density of the solution, C_p denotes the specific heat of the solution, and ΔH_c represents heat of crystallization. Solving Eqs. (5.6), (5.8) and (5.9) together yields particle population distribution, solution concentration, and bulk temperature profile over time (Berglund 2002; Ramkrishna 2000). If heat of crystallization is minor, energy balance can be neglected. Initial concentration (and temperature) and seed population must be obtained as the initial conditions to solve the system of equations. Method of moment (MOM) is a common PBM solving method that describes the population distribution by its moments:

$$\mu_j = \int_{\infty}^{L_0} f\left(t,L\right) L^j dL \tag{5.10}$$

where μ_j denotes the jth moment of the population distribution f, μ_0 is the total number of particles, μ_1 represents the total length, μ_2 is the total surface, and μ_3 is the total particle volume per unit volume of the slurry. The population balance Eq. (5.6) can then be converted into a series of ordinary differential equation using moments:

$$\begin{aligned} \frac{d\mu_0}{dt} &= B \\ \frac{d\mu_1}{dt} &= G\mu_0 \\ \frac{d\mu_2}{dt} &= 2G\mu_1 \end{aligned} \tag{5.11}$$

Mass and energy balance can be written as:

$$\frac{dC}{dt} = -3\rho_c k_v G\mu_2 \tag{5.12}$$

$$\frac{dT}{dt} = \frac{-hA\left(T - T_w\right)}{V\rho_{sol}C_p} - \frac{\Delta H_c 3k_v\rho_c}{V\rho_{sol}C_p}G\mu_2 \tag{5.13}$$

The PBM is now converted as a system of ordinary differential equations which is relatively easy to solve. Alternatively, a finite volume method can be applied to numerically solve for the full population distribution which cannot be obtained using MOM, as well as concentration (and temperature). The details of the finite volume method as well as other solution techniques will not be discussed here, but references such as Kumar and Ramkrishna (1996), Gunawan et al. (2004), and Kubota et al. (2001) are recommended if interested. Many commercially available software packages are capable of solving population balance models with finite volume methods.

5.1.2 Batch and Continuous Crystallization

Crystallization of pharmaceuticals has been largely carried out in batch mode (Cao et al. 2017; Chung et al. 1999; Feng and Berglund 2002; Hu et al. 2005; Nagy and Braatz 2012; Peña et al. 2017a; Puel et al. 2003; Saleemi et al. 2012). Batch operations are "recipe-based," simple, and relatively low maintenance (Ferguson et al. 2013; Peña and Nagy 2015; Tung et al. 2008). Carefully designed batch crystallization operating curves (black dotted line in Fig. 5.11) can achieve relatively uniform CSD and maximum yield; however, it usually requires a very long batch time, and consequently batch crystallizers are often very large in size (on the order of 1 m^3). Such large size leads to high capital cost, high operating cost, elaborate scale-up practice (DiMasi et al. 2003; Suresh and Basu 2008; Teoh et al. 2016), and heavy consequence for failed batches which is a prevailing issue due to batch-to-batch variations (Plumb 2005). Furthermore, local poorly mixed spots are often present in

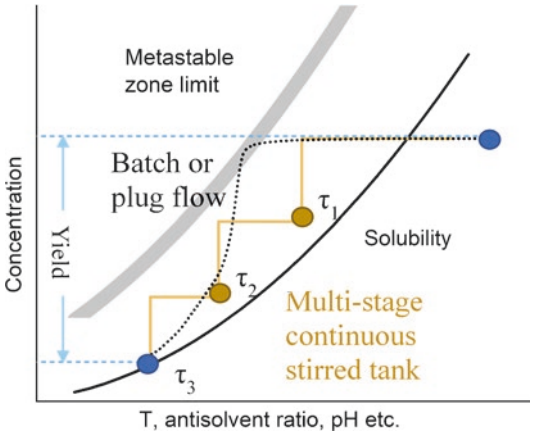


Fig. 5.11 Typical operating profile of batch, multistage continuously stirred tank, and plug flow crystallization in the phase diagram

large reactors where supersaturation is exceptionally high triggering primary nucleation despite seeding (Green and Myerson 2001). Large number of fines are produced as a result leading to possible fouling, filtration failure, excessive agglomeration, and undesired polymorph formation among other undesired issues.

Continuous crystallization on the other hand operates at "steady state" or, more accurately, at a "state-of-control" operation improving process robustness and product consistency (Nagy et al. 2013; Wood et al. 2019). It is crucial to incorporate PAT during continuous crystallization to maintain the consistent state-of-control operation. Continuous operations usually allow smaller-sized reactors because of its ability to operate for a long period of time without interruption, enabling on-demand manufacturing that can be easily relocated and reconfigured between different locations. More importantly, continuous crystallization may be scaled by simply increasing the flow rate with little to no equipment size change which significantly simplifies or potentially eliminates the traditional scale-up practices associated with batch crystallization (Nagy and Aamir 2012; Wang et al. 2017; Zhang et al. 2017). However, there are challenges associated with continuous crystallization despite the inherent advantages. These challenges mainly center on the issue of fouling and difficult slurry transfer: continuous crystallization inherently requires supersaturation throughout the whole operation which tends to cause encrustation on equipment surfaces and in transfer lines. Coupled with low flow rates required for most pharmaceutical processes, fouling and blockage can easily occur, and the state-of-control operation may subsequently be interrupted (Acevedo et al. 2019; Yang et al. 2017). Additionally, the continuous supersaturation present in the crystallizer inevitably leads to lowered yield compared to batch which is another obstacle for shifting to continuous from an economical perspective. Continuous crystallization is by no means a new concept. It has been an established process at large-scale production in industries like sugar processing and mineral refining, but it becomes challenging at smaller scales where pharmaceutical industry usually operates. It is often not a trivial decision to operate in batch or continuous mode when developing a crystallization process. Factors like demand, material property, time constraint, budget, and level of technology expertise must be considered during the decision-making process. When such a decision of continuous operation is beneficial, we hope the following discussion on continuous crystallization will provide readers with some insights on the design of a successful continuous crystallization process.

5.2 Continuous Crystallizers

There are two general types of continuous crystallizers: mixed-suspension-mixed-product removal (MSMPR) crystallizers where material is fed and removed from a well-mixed vessel producing wide residence time distribution (RTD) and near plug flow crystallizers (PFCs) where the material "flows" through a tubular reactor with the RTD being near uniform. MSMPR operation allows utilization of existing equipment, while PFC produces tight RTD enhancing product consistency. Each type can be the appropriate choice for different crystallization systems.

5.2.1 Mixed-Suspension-Mixed-Product Removal (MSMPR) Crystallizers

MSMPR operation consists of one or more well-mixed vessel(s) in series with continuous feed and slurry removal. Successful employment of continuous crystallization in MSMPRs relies on sufficient mixing and proper slurry withdrawal schemes to obtain products in the outlet representative of the bulk slurry. A single MSMPR crystallizer operates at one point in the phase diagram (e.g., τ_1 on Fig. 5.11), but a batch-like operating curve can be replicated by employing a multistage MSMPR system (yellow solid line on Fig. 5.11). For many APIs, pilot-scale crystallizers are enough to meet current demand in continuous mode without further scale-up. The possibility to utilize existing equipment and its simplicity of operation are major advantages of MSMPR systems over other continuous crystallization systems. Another advantage is the ability to operate at long residence times without particle settling issues which is especially beneficial for slow-growing compounds (Wood et al. 2019). During MSMPR mode operation, some slurry elements immediately exit the crystallizer after entrance; some elements end up never leaving the vessel, while most elements are somewhere in between, making up a broad RTD profile. This can be a disadvantage of MSMPR systems because broad RTD may lead to broad CSD.

Equipment and Scale

The most commonly used MSMPR crystallizer is a stirred tank crystallizer (STC) as it is readily available in almost every lab or pilot plant and it is simple and relatively cheap to build. Similar to a batch crystallizer, an overhead stirrer is often used to provide mixing. The agitator type, position, and agitation speed are all important design factors to consider for a well-mixed STC. Axial flow impellers such as pitched blade or retreat curve impellers are often used in crystallization. The material of the impeller must be durable and chemically compatible. Particles present in pharmaceutical crystallization processes tend to be heavier than its solvent; thus the impeller is often positioned closer to the bottom. Sufficient agitation speed is also needed to keep crystal particles suspended. This speed can be determined by experiments and/or computational fluid dynamics (CFD) which can be costly. Previously developed empirical equations are often used in practice to determine the minimal agitation speed. One of the most commonly used equations is known as the Zwietering correlation (Zwietering 1958):

$$N_{js} = Z_c \upsilon^{0.1} d_p^{0.2} \left[\frac{g_c \left(\rho_c - \rho_{sol} \right)}{\rho_{sol}} \right]^{0.15} X^{0.13} D_{im}^{-0.85} \tag{5.14}$$

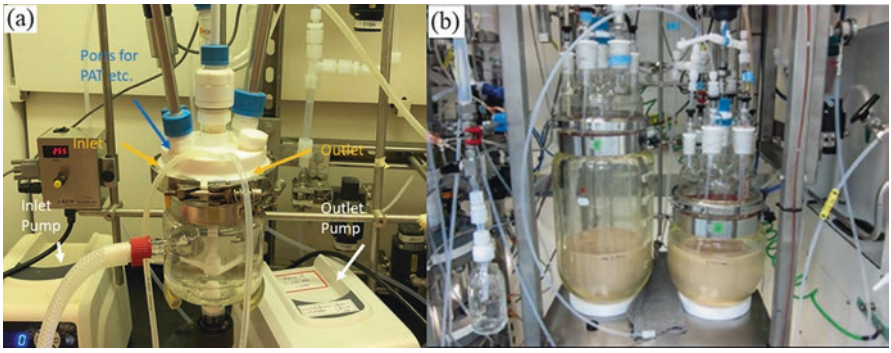


Fig. 5.12 Equipment setup of (**a**) a lab-scale (500 mL) stirred tank MSMPR. (Permission obtained from Liu et al. © 2019a Elsevier B.V.) and (**b**) a kilo-scale stirred tank MSMPR. (Permission obtained from Cole, K. P. et al. Copyright © 2017, Copyright © 2017 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works)

where N_{js} denotes the minimum agitation rate required for particles of mass mean size d_p; Z_c is the Zwietering constant, unique to the geometric characteristics of the agitation system, which can be experimentally determined or found in literature and handbooks; ν is the kinematic viscosity of the solution; g_c is the gravitational acceleration constant (9.81 m/s^2); X is the weight percentage of solid to liquid content; and D_{im} denotes the impeller diameter. An added complexity in crystallization is that agitation speed has a significant impact on crystal habits by affecting crystallization kinetics. While higher agitation speed improves mass transfer which improves local supersaturation, it can also induce particle breakage (Mazzarotta 1992), attrition (Biscans 2004), and agglomeration (Peña and Nagy 2015) which are often undesirable. An optimal mixing scheme design relies on experiments, modeling, and engineering experience.

Small bench-scale MSMPR STCs range from 10s mL to 100s mL (an example of 500 mL STC is shown in Fig. 5.12a) often equipped with ports for PAT tools (Ferguson et al. 2013; Garside and Tavare 1985; Garside and Shah 1980; Power et al. 2015; Su et al. 2015b; Yang and Nagy 2015a, b; Zhang et al. 2012). Pilot plant scale of a few liters (an example is shown in Fig. 5.12b) to 15 liters in volume have also been demonstrated to successfully produce crystal products at the rate of several kilograms of API per day (Polster et al. 2014; Cole et al. 2017). Larger equipment is not necessarily required to meet production demand at pharmaceutical scales, and laboratory/development equipment is capable of matching batch-scale production at equivalent or shorter timescale (Polster et al. 2014; Vetter et al. 2014; Wood et al. 2019) assuming that continuous crystallization can be sustainably operated for long periods of time, but it may be challenging. One of, if the not the most challenging aspects of maintaining a sustainable continuous crystallization in an MSMPR, is slurry transfer.

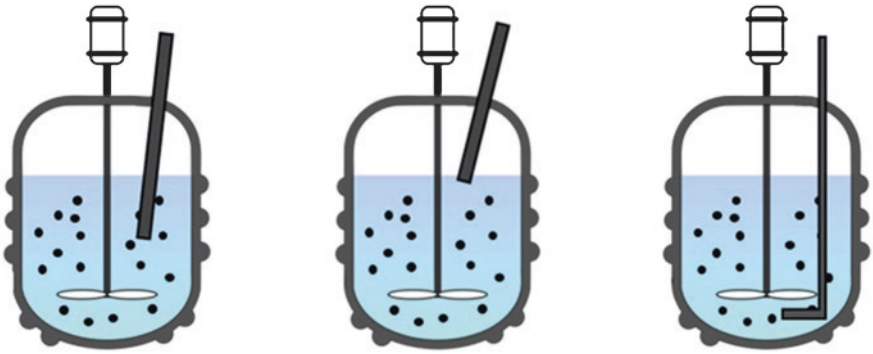


Fig. 5.13 Animation illustration of MSMPR drawing schemes: skimming (left), submerged (middle), and metal pipe-guided bottom drawing (right) (Yang et al. 2017b. Permision obtained form Yang et al. Copyright ©) 2017 American Chemical Soceity)

Slurry Transfer

Continuous crystallization is a two-phase process where a mixture of solid and liquid must be transferred from one vessel to the next which presents a host of challenges. Slurry transfer failure, including inhomogeneous slurry removal and transfer tubing blockage, has been reported to be the top reason causing premature termination during continuous crystallization studies. Representative or homogeneous product removal refers to the practice where the product removed at the outlet has the same composition (and CSD) of the bulk slurry. While it needs further discussion whether representative product removal is strictly required, representative slurry removal is often preferred for a more stable steady-state continuous operation. To ensure representative slurry removal, not only sufficient suspension of particles in the vessel and sufficient removal mechanisms are needed; the location of the transfer line is also important. Three common slurry removal locations are demonstrated in Fig. 5.13. Due to the nature of agitational mixing dynamics, particles tend to accumulate under the impeller; the inlet of the product removal line is recommended to be placed lower than the impeller if possible. A drain may be used as the product removal outlet to further improve this issue.

A bigger challenge is fouling and blockage of transfer lines (Fig. 5.14). A wide removal tubing, which is often required to prevent blockage, coupled with slow removal rates that are needed to maintain a certain residence time, poses a challenge to sustain turbulence in the transfer line (Cui et al. 2016; Hou et al. 2014). Lack of turbulence may cause particles to settle and/or block the transfer line completely. Moreover, unconsumed supersaturation that is naturally present during MSMPR operation can cause further crystallization in the transfer line which worsens particle settling as well as onsets encrustation on the tubing wall. This in turn can result in lowered yield, line blockage, and in some cases process failure. Therefore, transfer lines should only be as short as possible without unnecessary kinks and pinches and slurry transfer should take place as quickly as possible. Temperature control can

Fig. 5.14 A blocked slurry transfer line of a lab-scale MSMPR (Yang X. et al. Permision obtained form Yang X. et al. Copyright © 2017 American Chemical Soceity)

also be implemented on the transfer line to prevent further crystallization. In this section, a few slurry removal strategies will be discussed in effort to prevent slurry transfer failure.

Pump-driven slurry transfer is commonly used in conjunction with soft Teflon tubing to transfer slurries as shown in the MSMPR setup in Fig. 5.12a (Acevedo et al. 2018; Acevedo, Kamaraju, et al. 2017; Johnson et al. 2012; Li et al. 2016; Liu et al. 2019a, b; Zhang et al. 2014). Programmable peristaltic pumps also enable a simple yet effective slurry transfer scheme to prevent transfer line fouling: intermittent product removal. It has been reported that intermittent removal significantly improves transfer line fouling compared to continuous operation. During intermittent operation, the slurry transfer pump remains idle for some time followed by high flow rate removal of the accumulated volume. This accumulation is recommended to be less than 10% of the total volume in the vessel to avoid significant disturbances. It has been shown to produce similar products as true continuous operation both by experiments and by mathematical modeling (Powell et al. 2015b; Su et al. 2017). However, the level fluctuation caused by intermittent operation can lead to worsen encrustation on the vessel wall and agitation shaft due to evaporation. In addition, particles may get left behind in the tubing or "fall" back into the crystallizer during the idle period of the pump. Thus, the peristaltic pump transfer scheme may not prevent transfer line fouling for every system and can cause material back mixing that further widens the already broad RTD of MSMPR systems.

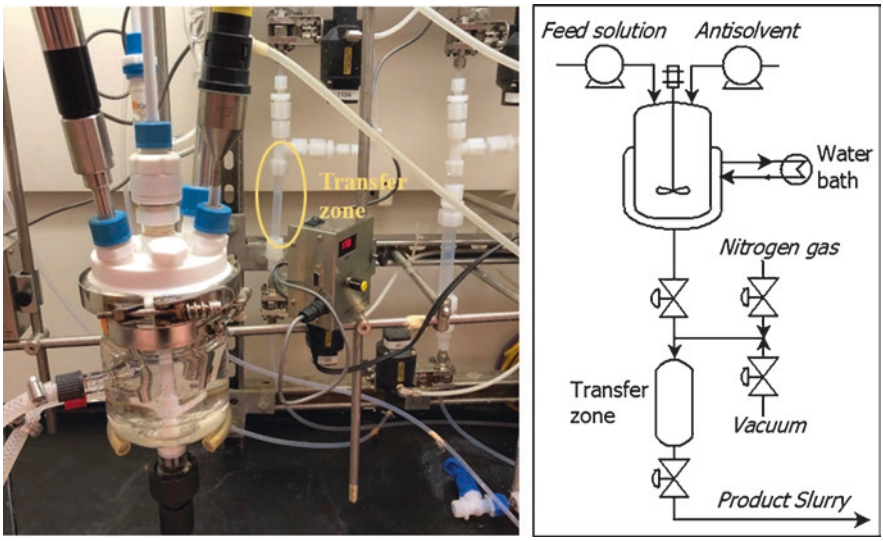


Fig. 5.15 Laboratory MSMPR slurry transfer zone and its schematics of the vacuum/nitrogen-driven transfer zone setup. (Permission obtained from Acevedo et al. © 2016 Elsevier B.V)

Employing a transfer zone is a more sophisticated intermittent slurry transfer strategy to address transfer line fouling by eliminating the accumulation of slurry in the line. It involves a slurry holding cell and a controlled vacuum/pressure system. Operated intermittently, the transfer zone mechanism periodically "rests" and "activates." During activation, slurry is quickly extracted from the vessel via a dip pipe and propelled to the holding cell by the coordination of valves and vacuum. The content in the holding cell is then purged into the next stage by nitrogen pressure. An additional line purge is often performed after each transfer to remove any remaining slurry. An example transfer zone setup is shown in Fig. 5.15. Transfer zone application is a superior strategy to effectively prevent transfer line fouling. Furthermore, vacuum as a driving force is more effective than the pump gear for slurry suction, and representative sampling is often guaranteed. However, it adds extra complexity that vacuum and pressure operating sequences must be robustly automated for sustainable operation. A further simplification of the transfer zone strategy can be achieved by implementing pressure-driven transfer schemes. The basic concept of pressure-driven slurry transfer is to simply remove the transfer cell and employ airtight MSMPR vessels. Intermittent slurry transfer is achieved by manipulating pressure and vacuum via valves opening/closing. A schematic of such a transfer system is shown in Fig. 5.16. It can be costly to construct such a robust system that undergoes frequent pressuring and depressuring during operation.

Transfer zones and pressure-driven transfers enable rapid extraction of representative samples, while maintaining little to no settling or fouling issues in the slurry transfer line. Their impact on particle size is also reduced compared to pumps as pump gears can cause particle breakage. But it cannot be said with certainty that it

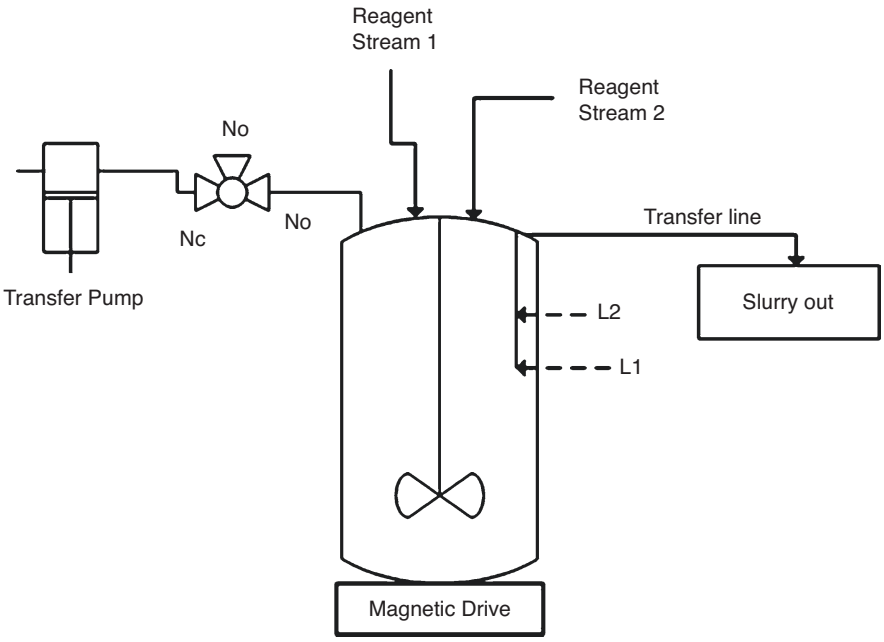


Fig. 5.16 Schematic of a pressure-driven MSMPR slurry transfer system. (Permission obtained from Cui et al. Copyright © 2016, American Chemical Society)

is always superior to implement complex transfer zones or pressure-driven transfer strategies because of their costly construction and complex operation. The fact that the implementation of simple pump transfer versus the utilization of complex slurry transfer techniques is comparable in popularity suggests that some systems are less sensitive to slurry transfer issues than others. Thus, the need for a sophisticated slurry transfer strategy varies on a case-to-case basis.

Start-Up Strategies

Start-up of continuous MSMPR operation is an important aspect of achieving a desirable state-of-control operation. Improper start-up procedures could lead to fouling, prolonged dynamics and even polymorph impurities. The start-up process should be optimized to minimize product loss and operational difficulties while shortening the time to reach the state-of-control. Continuous MSMPR operation can start with a clear solution or slurry. It is however not advisable to start the MSMPR crystallizer dry as it may result in excessive nucleation and fouling. If a clear solution is used during start-up, primary nucleation will take place, and the primary particles can serve as seed. A slurry solution can also be used as start-up which serves as a seed bed. It is important to note that unless seed is carried in continuously with feed, residence time must be long enough to avoid excessive particle

washout which in turn prevents periodic primary nucleation. A batch crystallization can be performed first to create the crystalline slurry for start-up, but an improper batch design can lead to significant fouling and blockage. A saturated solution with premade seed has been demonstrated for cooling continuous crystallization of paracetamol in water to produce the shortest time to reach steady state compared to batch cooling start-up mode or clear saturated solution start-up (Power et al. 2015). While start-up procedure might not affect the final state-of-control products for many systems, improper start-up can cause unstable steady-state operation where the system can easily deviate from a controlled state to seemingly produce different products with different start-up strategies. It has also been observed that reproducible results are not guaranteed in some experimental start-up studies of continuous crystallization and they need to be improved in future studies. Mathematical modeling can be especially helpful for designing an optimal start-up strategy.

State-of-Control Operation

The phrase "state-of-control" has been used in place of "steady-state operation" in this chapter. State-of-control is a more accurate description of stable continuous crystallization operation where small deviations from steady state may occur but system dynamics and/or applied control strategies can correct such a deviation without causing significant changes in product CQAs (Nagy and Braatz 2012; Acevedo et al. 2019; Lee et al. 2015; Woodcock 2014; U.S. Food and Drug Adminiatration 2019). Such a dynamic or sometimes oscillatory state is considered a steady and stable operation and is more appropriately called a "state-of-control" operation. During state-of-control operation, product quality attributes such as CSD, solution concentration, impurity level, and polymorphic form remain constant or in an acceptable range. The final state-of-control product qualities are often governed by the operating conditions; for example, if the temperature of the MSMPR is low during a cooling continuous crystallization process aiming to improve yield, CSD is likely to be small because of excessive nucleation but is observed to improve if residence time increases (Powell et al. 2016; Power et al. 2015). Incorporation of milling or ultrasound also affects state-of-control CSD by inducing particle breakage and/or accelerating nucleation. Yield is another important property to consider when designing a suitable state-of-control operation. It is often of economic interest to maximize yield. Recycling, being an obvious choice, has been shown to significantly improve the overall yield in MSMPR operation (Griffin et al. 2010; Alvarez et al. 2011; Ferguson et al. 2014; Vartak and Myerson 2017; Wong et al. 2012). Alvarez et al. (2011) showed that implementing a recycling stream increased the yield from 71% to 87%, while decreasing purity by ~2% for the cooling crystallization of cyclosporine (Fig. 5.17). Thus, recycling may be an unacceptable strategy for systems of high purity requirement. Lowering the MSMPR crystallizer temperature has also been shown to improve yield, but it may result in smaller crystal size (Quon et al. 2012; Rashid et al. 2017; Yang et al. 2015b). Inclusion of wet mill (Fig. 5.18) and sonication (Fig. 5.19) can improve particle uniformity, but it is a

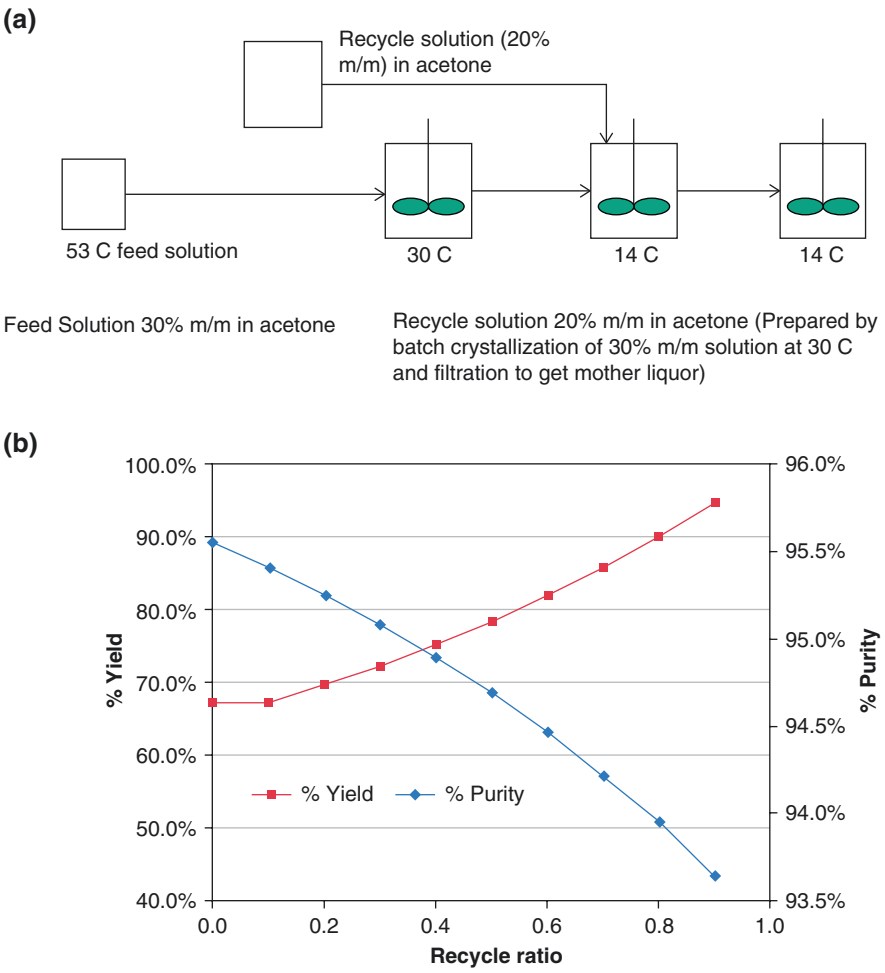


Fig. 5.17 (**a**) Schematic of multistage MSMPR continuous crystallization of cyclosporine with recycle; (**b**) effect of recycle ratio on product purity and process yield. (Permission obtained from Alvarez et al. Copyright © 2011, American Chemical Society)

trade-off with crystal size (Narducci et al. 2011; Yang et al. 2015a; Yang et al. 2016a). Increasing residence time can improve yield without compromising CSD and often times improving state-of-control crystal size (Ferguson et al. 2013; Su et al. 2015b; Wierzbowska et al. 2008). However, it will require lowering the flow rate, thus decreasing throughput or requiring more stages which increases cost. Depending on the kinetics, a combination of the aforementioned techniques can be applied to obtain an optimal state-of-control.

Polymorph (and chirality) control is another important but perhaps more subtle advantage that comes with MSMPR operation. Unlike batch, one MSMPR vessel state-of-control operation fixates on a single point in the phase diagram consistently

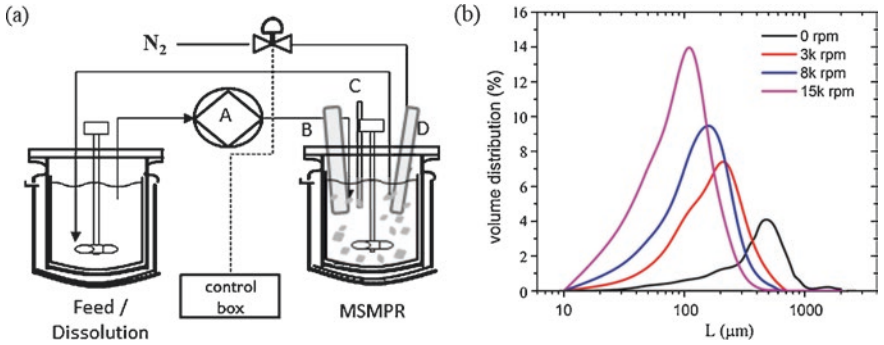


Fig. 5.18 (**a**) Schematic of in situ wet mill in the MSMPR experimental setup and (**b**) volume-based size distribution at different wet mill rpm. (Permission obtained from Acevedo et al. Copyright © 2017, American Chemical Society)

producing the polymorphic form favored at that point. A batch operation curve may inevitably pass through regions in the phase diagram that favor an undesired polymorphic form. Thus, maintaining a specific state-of-control operation can achieve tight control of the polymorphic (chiral) form (Qamar et al. 2012; Köllges and Vetter 2017; Lai et al. 2015; Powell et al. 2015a). Lai et al. (2014) demonstrated this aspect by manipulating the temperature and residence time of the MSMPR operation to selectively produce a β-form of L-glutamic acid that is relatively difficult to obtain under batch crystallization. Similarly, Steendam and Horst (2017) established a single-stage MSMPR continuous process to consistently produce chirally pure crystals by manipulating residence time, feed concentration, and start-up strategy to tune for a desirable state-of-control operation.

The aforementioned fouling issue remains a challenge to maintain the state-of-control operation. As of now, MSMPR continuous crystallization of pharmaceuticals remains in research stage. Fouling and blockage are often vaguely reported or kept off records. In addition, the robustness and reproducibility of these studies are not tested rigorously. Therefore, the issue of fouling may take place more frequently and cause more serious consequences than what current literature suggests. To achieve industrialization of continuous crystallization, the risk of fouling and blockage must be addressed. While choosing a set of operating conditions that will result in a state-of-control less likely to foul can be helpful, feedback control strategies provide a much more reliable solution that resolves fouling at its onset instead of solely relying on prevention. A possible strategy is automated direct nucleation control (ADNC) which is a model-free control strategy to control the number of particles present in the crystallizer by manipulating the supersaturation level, usually by changing the temperature in the crystallizer. It requires a PAT sensor to monitor particle count, usually focused beam reflectance measurement (FBRM). ADNC has been implemented by Yang, Song, and Nagy (2015b) in a two-stage MSMPR continuous crystallization process to improve product CSD. ADNC may be used to correct fouling by detecting the change in particle count in the crystallizer caused

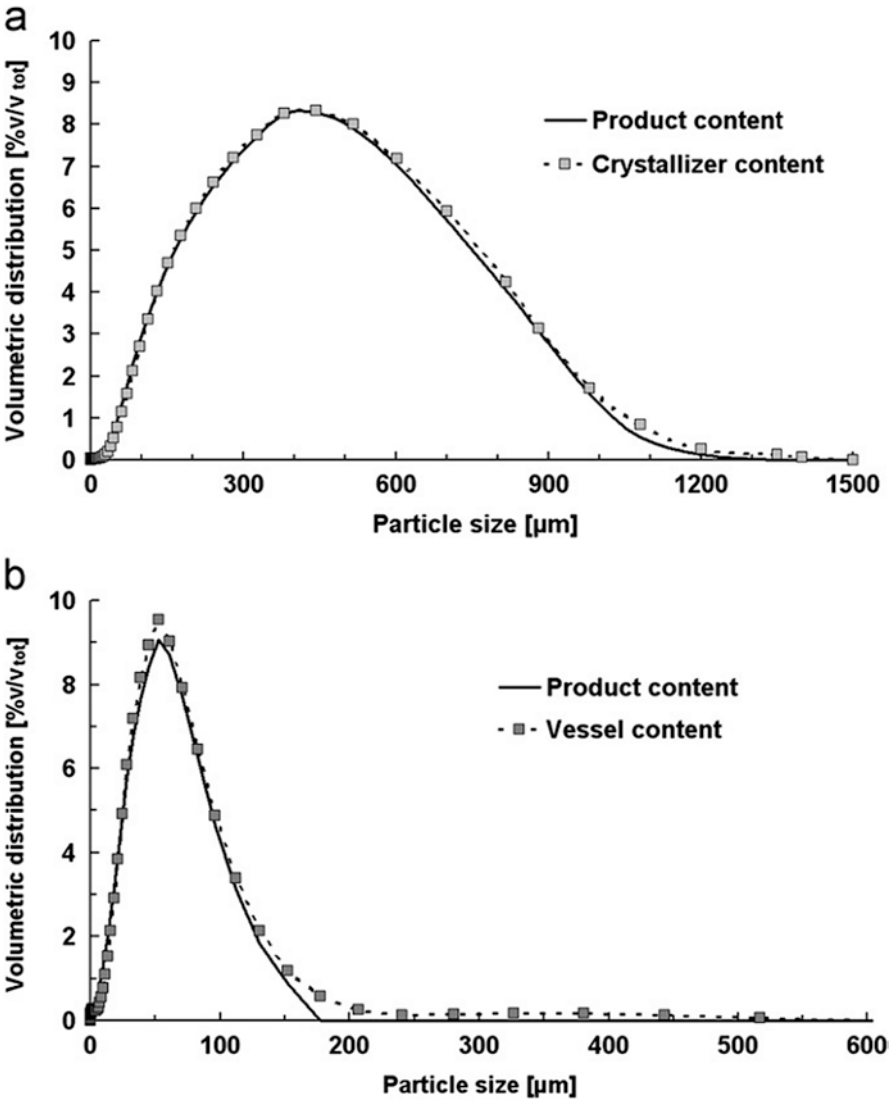


Fig. 5.19 Volume-based size distribution of particles in crystallizer implemented with sonication and product vessel at (**a**) silent conditions and (**b**) 40% power amplitude sonication. (Permission obtained from Narducci et al. Copyright © 2011 Elsevier Ltd.)

by fouling. Surely more intelligent control schemes like model predictive control (MPC) can be implemented to better address this issue (Mesbah et al. 2017; Su et al. 2019), but the main challenge lies in the detection of fouling. Particle count change is not necessarily a sign of fouling nor does fouling necessarily result in particle count change in bulk. A more reliable method must be developed to detect the onset of fouling more accurately and quickly. Methods such as ultrasound and

conductivity measurement that have been applied in other industries to detect encrustation may be borrowed to detect fouling in continuous crystallization. Image analysis utilizing intensive algorithm and high-resolution cameras is a more direct measurement of fouling, but it may be technically and/or economically unrealistic. Even though challenging, the risks associated with fouling can be limited to an acceptable level by a combination of proper transfer line design, optimized operation condition selection, and suitable control strategy implementation. Further study of the robustness of continuous crystallization in MSMPR mode is still needed to establish a systematic process development strategy to achieve a robust state-of-control operation.

Oscillatory MSMPR Crystallizers

While stirred tank crystallizers (STCs) are convenient and relatively simple to use, it has some inherent drawbacks such as poor local mixing and high shear rate near the agitator. There are research efforts focused on alternative MSMPR vessels, one of which is an oscillatory baffle reactor (OBR). It is also called a dynamic baffle crystallizer (DBC) where instead of an agitator, a set of oscillating baffles are used to provide mixing (Fig. 5.20). Oscillatory mixing provides more uniform fixing by inducing fully developed vortices (Fig. 5.21a) distributing mechanical energy more evenly throughout the reactor space (Mackley et al. 1990; Mackley and Ni 1991; Mackley and Stonestreet 1995; Ni et al. 1995a, b, 1997; Ni and Gao, 1996). As a result, heat and mass transfer is significantly improved in oscillatory mixing systems compared to agitational vessels (an example study is shown in Fig. 5.21b). Efficient mixing is especially advantageous for continuous crystallization

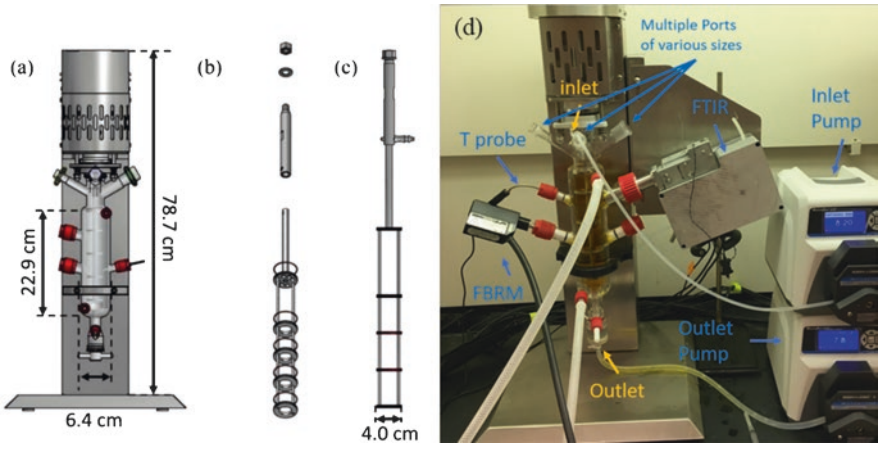


Fig. 5.20 (**a**) Overall schematic of the OBR/DBC, (**b**) an exploded view drawing of the baffle shaft, (**c**) side view of the baffle shaft, and (**d**) a picture of an experimental continuous OBR setup. (Permission obtained from Liu et al. © 2019a Elsevier B.V)

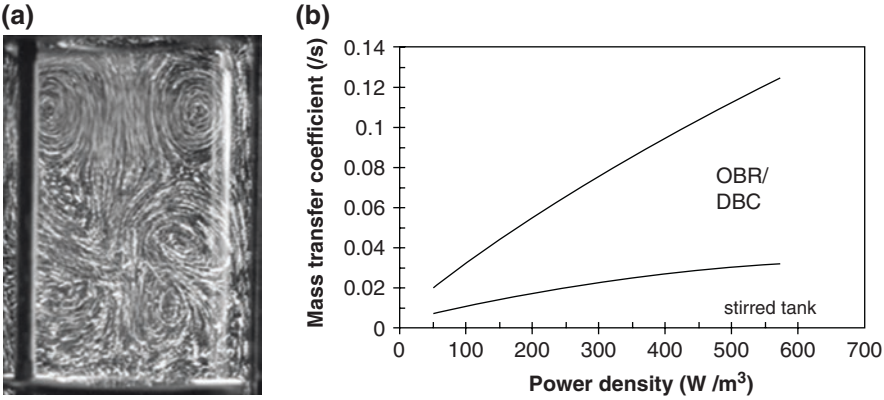


Fig. 5.21 (**a**) Image of fully developed vortices in an OBR/DBC (permission obtained from Ni et al. Copyright © 2003 The Institution of Chemical Engineers. Published by Elsevier B.V.); (**b**) improved mass transfer coefficient between water and air in an oscillatory system and a stirred tank. (Permission obtained from Ni and Gao, Copyright © 1996 Elsevier B.V)

applications because it will significantly improve local supersaturation which lowers the chance of unwanted primary nucleation and promotes growth (Brown et al. 2014, 2015). An OBR usually consists of a jacketed elongated vessel and a set of baffles. The baffles can be tightly fitted to the vessel leaving no space between the baffles and the wall, or they can be loosely fitted leaving some space which is less common. The baffles are low-constriction "donut"-shaped plates with the orifice taking up 20–30% of the whole plate area. The baffles are usually placed at equal distance which is about 1.5 times the plate diameter (Brunold et al. 1989; Dickens et al. 1989; Ni et al. 2003). The degree of mixing can be tightly controlled by adjusting frequency and/or amplitude of the oscillation. Oscillatory Reynolds number (Re_o) can be used to describe the mixing intensity (Fitch 2003; Ni and Gough 1997):

$$Re_o = \frac{2\pi\omega x_o \rho D}{\mu} \tag{5.15}$$

where ω is the frequency of oscillations, x_o is the center to peak oscillation amplitude, D is the inner diameter of the tube, ρ is the fluid density, and μ is the fluid viscosity. Strouhal number (St) can be used to quantify eddy propagation (Fitch 2003; Ni and Gough 1997):

$$\mathrm{St} = \frac{D}{4\pi x_o} \tag{5.16}$$

Shear rate is generally lower in an OBR in comparison to a stirred tank crystallizer which can be beneficial for crystallization (Brown and Ni 2012; Chew and Ristic 2005; Hewgill et al. 1993; McGlone et al. 2015; Ni et al. 2000a, b, 2004; Ni

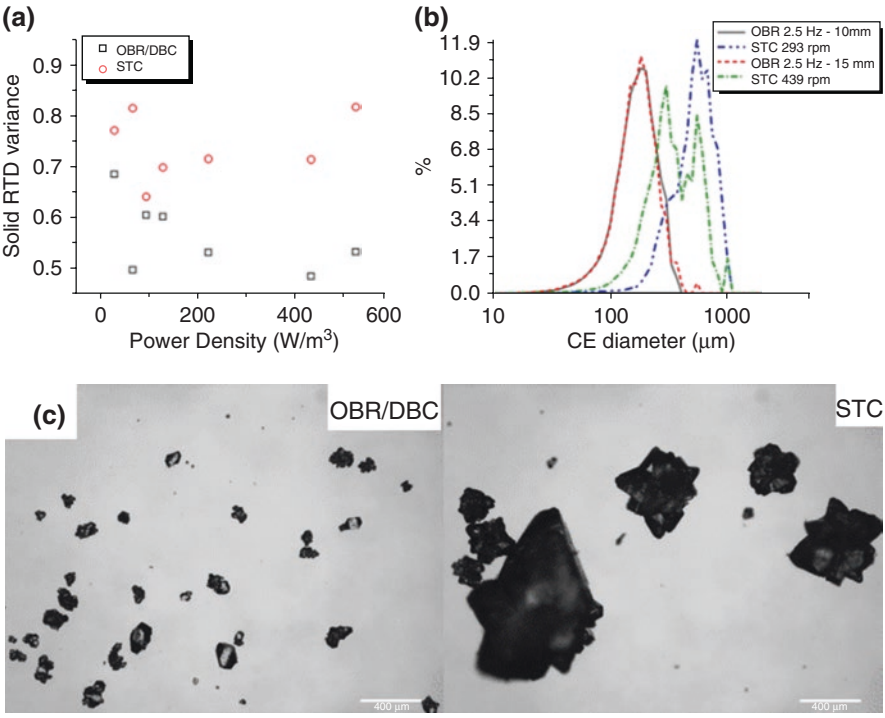


Fig. 5.22 (**a**) RTD variance improvement in the OBR/DBC compared to STC at the same power density, (**b**) volume-based CSD of OBR/DBC and STC, and (**c**) microscopic images of state-of-control product of OBR/DBC and STC at the same power density. (Permission obtained from Liu et al. © 2019a Elsevier B.V)

and Mackley 1993; Ristic 2007). Liu et al. (2019a) carried out a comparative study of the OBR and the stirred tank. The authors observed that the RTD is narrower in the OBR as shown in Fig. 5.22a. It is also observed that state-of-control product of continuous crystallization of paracetamol was more uniform and less aggregated in the OBR than in the stirred tank at the same power density as shown in Fig. 5.22b, c. However, the OBR is a slightly more complex system and is not readily available. Incorporating PAT sensors is also more difficult than the STC where they can simply be inserted overhead. The setup used in the aforementioned study features side opening ports for PATs. However, airtight fittings had to be custom made for PAT sensors of various sizes. These ports can pose as dead zones without proper fittings causing particle settling and accumulation in the ports. A more optimized design is needed to improve the geometry of PAT ports.

Modeling of MSMPR Continuous Crystallization

The power of PBM is especially significant when designing an MSMPR crystallization process including start-up/process optimization and MPC implementation (Benyahia et al. 2012; Lakerveld et al. 2013; Morris et al. 2015; Yang and Nagy 2015a, b). To represent an MSMPR operation, the crystallizer is assumed to be well mixed; in other words, the product removed has the same composition including solution concentration and solid CSD as the bulk material in the crystallizer. Therefore, accumulation of particles = number of particles in – number of particles out + number particles generated by crystallization. The population balance of a single-stage MSMPR can be written as:

$$\frac{\partial(Vf)}{\partial t}+\frac{\partial(VGf)}{\partial L}=\dot{Q}_{\text{in}}f_{\text{in}}-\dot{Q}_{\text{out}}f+V\left(Q_{\text{nuc}}+Q_{\text{agg}}+Q_{\text{break}}\right) \tag{5.17}$$

where V is the slurry volume in the crystallizer, $Q_{\cdot\text{in}}$ and $Q_{\cdot\text{out}}$ represent the inlet and outlet volumetric flow rate, and f_{in} is the population distribution of the seed carried into the system with feed (taken as zero if the system is not continuously seeded, i.e., there are no particles entrained with feed). If agglomeration and breakage are negligible, Eq. (5.17) can be simplified as:

$$\frac{\partial(Vf)}{\partial t}+\frac{\partial(VGf)}{\partial L}=\dot{Q}_{\text{in}}f_{\text{in}}-\dot{Q}_{\text{out}}f+VB\delta\left(L-L_0\right) \tag{5.18}$$

A mass balance can also be written for a single-stage MSMPR operation:

$$\frac{d(VC)}{dt}=\dot{Q}_{\text{in}}C_{\text{in}}-\dot{Q}_{\text{out}}C-3Vk_{\text{v}}\rho_{\text{c}}\int_{\infty}^{L_0}GL^2f\,dL \tag{5.19}$$

where C_{in} denotes the feed concentration. Solving Eqs. (5.18) and (5.19) together yields the population and concentration over the entire time of operation. Initial conditions can be written based on the start-up method:

$$\begin{aligned}&f\left(@\,t=0\right)=\text{population of intial seed, if seeded}\\&f\left(@\,t=0\right)=0\,\text{for all}\,L,\,\text{if unseeded}\\&C\left(@\,t=0\right)=\text{initial concentration}\end{aligned} \tag{5.20}$$

At constant volume steady state, the first term of Eqs. (5.18) and (5.19) becomes zero, and the solution gives steady-state concentration and particle population. Therefore, the population balance of a constant volume single-stage MSMPR crystallization without continuous seeding at steady state becomes:

$$\frac{\partial(Gf)}{\partial L}+\frac{f}{\tau}-B\delta\left(L-L_0\right)=0 \tag{5.21}$$

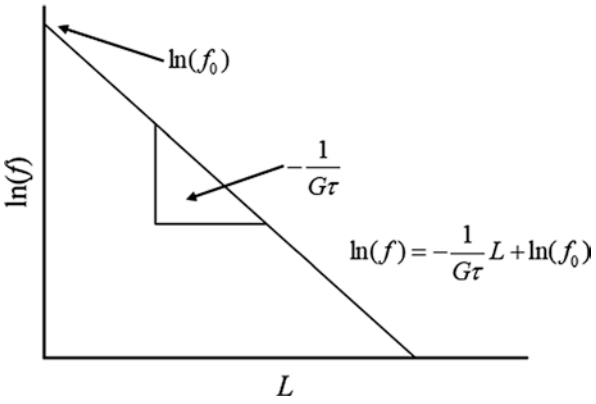


Fig. 5.23 Semilogarithmic population density versus size plot to predict growth rate

where τ is the residence time. Equation (5.21) is an ODE and can be solved to estimate size-independent linear growth rate:

$$\frac{df}{dL} = -\frac{f}{G\tau}$$
$$\text{B.C.} f_0 = \frac{B}{G} = \text{population of nuclei} \tag{5.22}$$

Solving the ODE analytically:

$$\ln\left(f\right) = -\frac{1}{G\tau}L + \ln\left(f_0\right) \tag{5.23}$$

Linear growth rate can be estimated by plotting logarithmic steady-state population f against size L as shown in Fig. 5.23 (Garside and Shah 1980).

A similar system of equations can be written for multistage MSMPR operations. Consider the ith stage of a multistage MSMPR system; the inlet flow of the ith stage is equal to the outlet flow of the $i–1$ th stage:

$$\frac{\partial\left(V_i f_i\right)}{\partial t} + \frac{\partial\left(V_i G f_i\right)}{\partial L} = \dot{Q}_{i-1} f_{i-1} - \dot{Q}_i f_i + V_i B\delta\left(L - L_0\right)$$
$$\frac{d\left(V_i C_i\right)}{dt} = \dot{Q}_{i-1} C_{i-1} - \dot{Q}_i C_i - 3V_i k_v \rho_c \int_{\infty}^{L_0} GL^2 f_i dL \tag{5.24}$$

Equation (5.24) can be solved for each stage with $i–1$th stage being the feed for when $i = 1$. Initial conditions can be written similarly to Eq. (5.20) based on the start-up procedure.

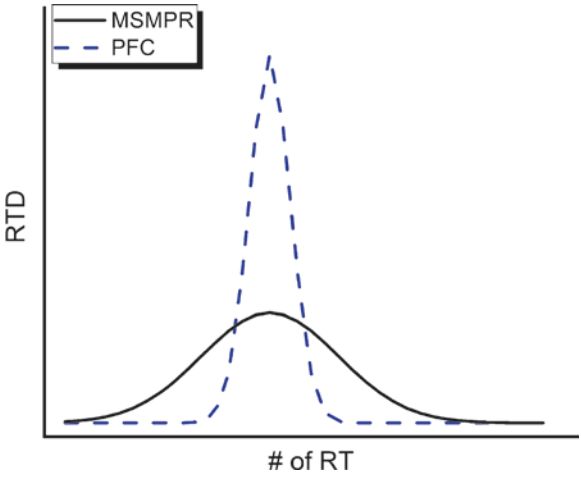


Fig. 5.24 Typical RTD of MSMPR versus PFC

5.2.2 Plug Flow Crystallizers (PFC)

Plug flow crystallizers (PFCs) are tubular reactors in which the content flows at a near-constant velocity. The obvious appeal of a PFC over an MSMPR vessel is that the axial mixing is minimal in the PFC producing a near-uniform RTD compared to the broad RTD in an MSMPR crystallizer as illustrated by Fig. 5.24. RTD uniformity may significantly improve product consistency in the PFC. In addition, superior heat and mass transfer is achieved in the tubular configuration of the PFC compared to a MSMPR vessel, and scale-up can be easily achieved by extending the length of the PFC (scaling-out). When a larger scale is needed, the PFC may be scaled up linearly because unlike a stirred tank MSMPR where mixing intensity is unevenly distributed in the reactor space, local mixing in a PFC is much more uniform. The spatial gradient (i.e., the length) of a PFC is equivalent to that of batch time which would require several MSMPR crystallizers in series to achieve the same operating profile. However, in order to maintain turbulence in the PFC, a high flow rate is usually required resulting in short residence times or extended reactor length. Encrustation and fouling are another challenge in maintaining the state-of-control in the PFC. Cleaning procedures can also be very complex for long PFCs. In this section, several PFC and near-PFC systems will be discussed as well as the issue of fouling. Mathematical modeling will also be discussed briefly in this section.

Equipment and Operation

While plug flow reactors have been successfully applied for single-phase reaction processes, the added complexity of particle suspension poses a challenge for successful crystallization operations in them. It is difficult to produce high enough

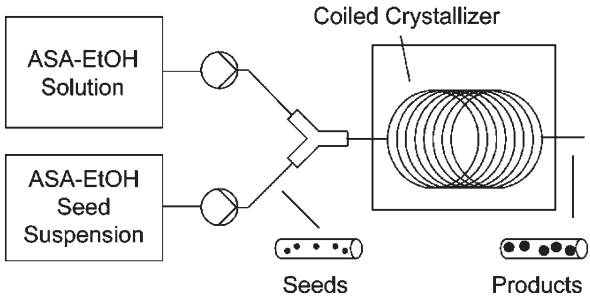


Fig. 5.25 Mesoscale plug flow crystallizer setup. (Permission obtained from Eder et al. Copyright © 2010, American Chemical Society)

turbulence in PFCs to ensure homogeneous mixing of the liquid phase and the solid phase. Therefore, a bench-scale PFC is rarely demonstrated to be successful. However, there are a few nonconventional PFCs that incorporated unique designs to enhance mixing for successful continuous crystallization operations.

The first common strategy to improve plug flow mixing for crystallization is to employ smaller-scale reactors. A mesoscale plug flow crystallizer of 2 mm inner diameter, 15 m in length, has been demonstrated by Eder et al. (2010) for continuously seeded cooling crystallization of acetylsalicylic acid (Fig. 5.25). The authors observed heavy agglomeration and blockage during their operation; however proper seeding and higher flow rates significantly improved this issue. Immediate blockage was observed without seeding which suggests that excessive nucleation causes serious fouling and must be avoided in this mesoscale system. Another group improved on this design (3.1 mm inner diameter 15.2 m length) by generating air segmented slug flow. Schematics and photographs of the slug flow setup are shown in Fig. 5.26. Each liquid slug behaves like a well-mixed batch crystallizer traveling along the length of the tube. The circulatory hydrodynamics in slug flow significantly improved fouling previously seen in non-segmented systems. It largely reduces stagnation near the wall which lowers the chance of clogging even at low flow rates, suitable for slow-growing compounds that require a long residence time. However, slug flow pattern can only be generated under certain flow rates with certain fluids. It may not be suitable for every process. Microscale slug flow crystallizers have also been demonstrated to successfully generate uniform crystals mainly for the purpose of fundamental kinetic studies. Superior mixing dynamics and improved heat and mass transfer are clear advantages of small-scale tubular reactors. Valuable experimental results can be obtained in these small-scale reactors in early development studies (Dombrowski et al. 2007; Eder et al. 2010; Lindenberg and Mazzotti 2011; Jiang et al. 2014; Lu et al. 2015).

Another approach to improve mixing in a PFC is implementing additional mixing geometry to create more turbulence. A static mixer is an example of such an approach. Alvarez and Myerson demonstrated the use of a Kenics-type static mixer (details shown in Fig. 5.27) to crystalize ketoconazole, flufenamic acid, and L-glutamic acid via antisolvent addition. The PFC is 0.6 m in length with an internal diameter of 12.7 mm and a volume of 76 mL. Small and uniform product was obtained at various flow rates, and larger particle size was obtained at higher flow

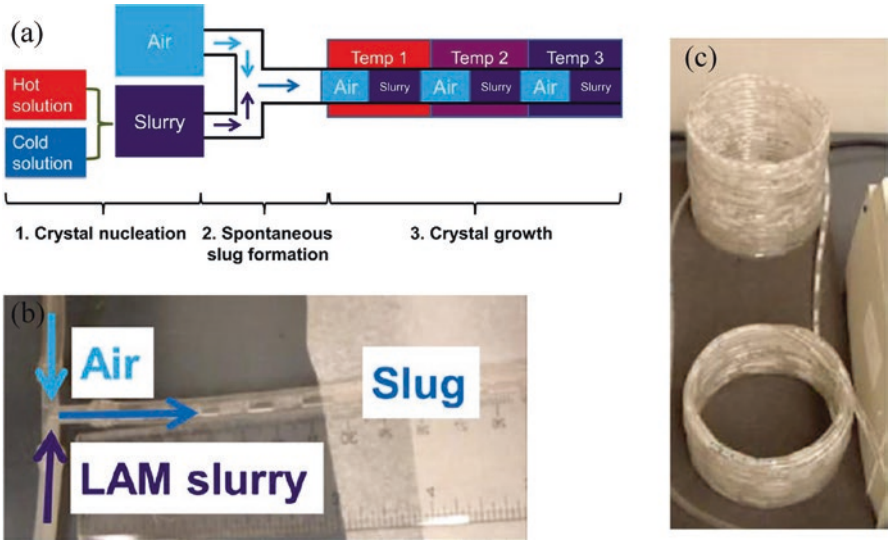


Fig. 5.26 (**a**) Schematic of the slug flow crystallizer setup, (**b**) image of slug flow T-junction, and (**c**) image of the slug flow crystallizer coil. (Permission obtained from Jiang et al. Copyright © 2014, American Chemical Society)

rates where yield was found to be lower. However, the residence time was less than 1 minute for all the flow rates tested which is mainly suitable to produce small and narrow CSD products via antisolvent addition but not suitable for the purpose of achieving large crystals. The addition of baffles on the tubing wall is also an effective approach to induce more turbulence in a PFC.

A variation of the statics addition approach to improve turbulence is to apply additional mechanical energy to the PFC. A combination of this approach with the aforementioned "statics" approach gave rise to a variation of PFC that has been gaining popularity in recent years: an oscillatory flow baffled crystallizer (OFBC) also known as a continuous oscillatory baffled crystallizer (COBC) or simply oscillatory baffled crystallizer (OBC) (Brown and Ni 2011; Lawton et al. 2009; Ni and Liao 2010; Peña et al. 2017b; Sang-Il Known et al. 2014; Siddique et al. 2015; Su et al. 2015a). It superimposes oscillatory motion onto the flow by a plunger in a baffled tubular reactor to aid particle suspension as shown in Fig. 5.28. The baffles are usually created by periodic pinches in the tubular diameter along the tubular reactor to create extra eddies and turbulences. A combination of oscillatory motion and static baffles generates sufficient turbulence at laminar net flow conditions attaining far longer residence times (10–30 minutes) compared to static PFCs (up to a few minutes). Oscillatory motion inevitably generates axial dispersion; in other words, the RTD in an OFBC is predictably less uniform than a PFC. To quantify axial dispersion, a dimensionless number, dispersion number (D/vL) can be calculated based on experimentally measured dispersion coefficient D, net flow velocity v, and the OFBC length L. Dispersion number of an ideal PFR is zero and infinity

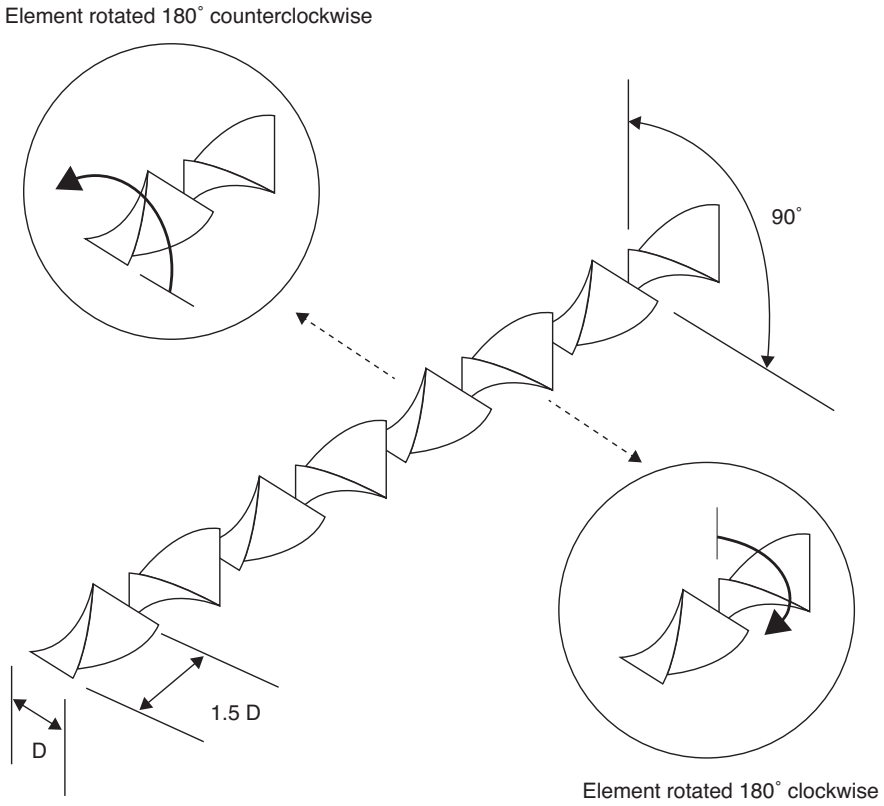


Fig. 5.27 Details of the Kenics-type static mixer. (Permission obtained from Alvarez & Myerson. Copyright © 2010, American Chemical Society)

for an ideal CSTR. Dispersion number of OFBC has been experimentally measured to be on the order of 10^{-2}, very close to a PFC (Kacker et al. 2017). Compared to an MSMPR, OFBC produces significantly more uniform products at comparable residence times that are not achievable in smooth or static PFCs. Oscillatory mixing also brings additional advantages such as improved heat and mass transfer and low shear as mentioned in section "Oscillatory MSMPR crystallizers". The same dimensionless groups including oscillatory Reynolds number and Strouhal number defined in Eq. (5.15) and Eq. (5.16) can be used to quantify the mixing conditions in an OFBC. Net Reynolds number (Re_n) is also of significance here:

$$Re_n = \frac{\rho v D}{\mu} \tag{5.25}$$

The OFBC is commercially available in different sizes made with glass or stainless steel. A picture of a commercially available OFBC unit is shown in Fig. 5.29. The tubular reactor is jacketed for temperature control and offers multiple injection

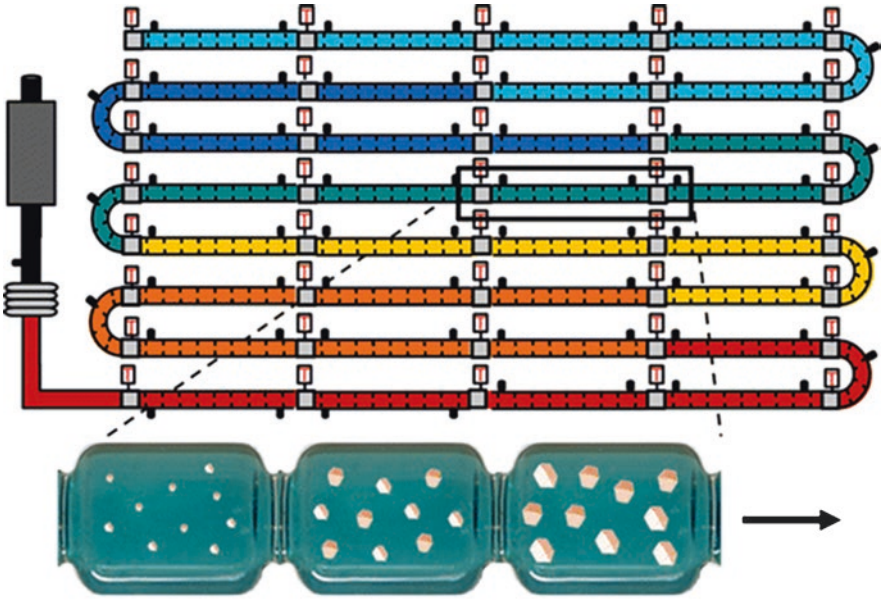


Fig. 5.28 Schematic of the tubular OFBC and a zoomed in view on a baffled tube segment. (Permission obtained from McGlone et al. Copyright © 2015, American Chemical Society)

Fig. 5.29 A commercially available glass model of OFBC unit designed by NiTech® solutions

ports for material inlet. It is also PAT capable by using special ports and fittings. Lawton et al. (2009) were among the first to demonstrate successful continuous crystallization of an API in an OFBC (Fig. 5.30). In the same study, a cost analysis was carried out which suggested a potential £300 k saving annually compared to

Fig. 5.30 A custom-built stainless steel OFBC unit. (Permission obtained from Lawton et al. Copyright © 2009, American Chemical Society)

Table 5.1 Cost analysis of a continuous operation in an OFBC compared to a traditional batch isolation process. (Lawton et al. 2009)

	Potential saving (£) from traditional batch operation
New build	20% lower
Operating costs	300 k per annum
Crystal engineering without milling	50% lower +300 K per annum

equivalent batch operations (Table 5.1) providing financial incentives to switch to continuous crystallization.

Fouling

A common issue that have been observed and reported for continuous crystallization in PFCs, including the OFBC, is fouling and encrustation (Acevedo et al. 2019). During continuous crystallization, nucleation can be triggered by existing particles as well as foreign objects like equipment walls, the latter initiate encrustation on reactor walls. An example of encrustation on the PFC wall is shown in Fig. 5.31. Once encrustation onsets, it may grow over the course of operation time altering process conditions and causing variability in product quality attributes. Since growth rates of bulk material and of encrust are similar, the severity of encrustation

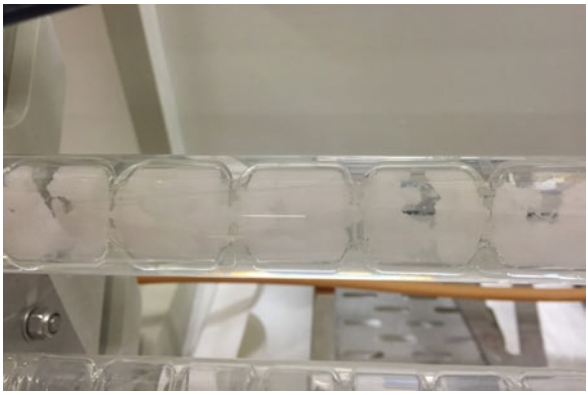


Fig. 5.31 Image of an encrusted tube section of an OFBC

largely depends on the kinetics of the two competing nucleation mechanisms. Encrustation layer consumes supersaturation which lowers yield and reduces heat transfer efficiency which worsens the yield. It may even completely block the PFC causing process failure. In addition, severe encrustation may pose serious safety risks for systems requiring effective heat transfer for heat relief.

Encrustation is affected by many factors including equipment surface properties (e.g., surface energy, roughness, and topology), crystallization system properties (e.g., interfacial tension, solubility, and contact angle), and process conditions (e.g., supersaturation level and mixing conditions) (Acevedo et al. 2019). To mitigate encrustation, there are two general strategies: onset prevention and reactive control strategies. In some systems optimization of crystallizer and process design may prevent encrustation (Majumder and Nagy 2015). Such strategies like treating the reactor surface, increasing temperature, and seeding the process can suppress encrustation kinetics so substantially that the timescale for significant encrust formation becomes much longer than the operation timescale thus completely preventing encrustation. However, in other circumstances thick encrust layer will form during operation regardless of crystallizer design or process condition optimization. In such situations, control strategies must be implemented to react to encrustation onset. Two control strategies can be used: PAT-based feedback control and model-based predictive control. Feedback control relies on PAT to detect fouling and initiate control actions to remove the encrust layer (e.g., heating cycles). Acoustic transducer, conductivity meter, and temperature probe can be used to detect encrustation; however, they cannot quickly detect encrustation onset. It takes some time for the encrust layer to grow to a certain thickness that causes detectable changes in acoustic, conductivity, or temperature measurements. A more direct detection method is to use a camera to recognize encrustation visually. Depending on the camera specs and image analysis capability, encrustation onset may be detected with very short delays. However, the use of camera may not be suitable for non-transparent crystallizers. The complex system dynamics of crystallization and encrustation can also be very difficult for a simple feedback PID controller. A model-based control strategy on the other hand can utilize PBM to predict

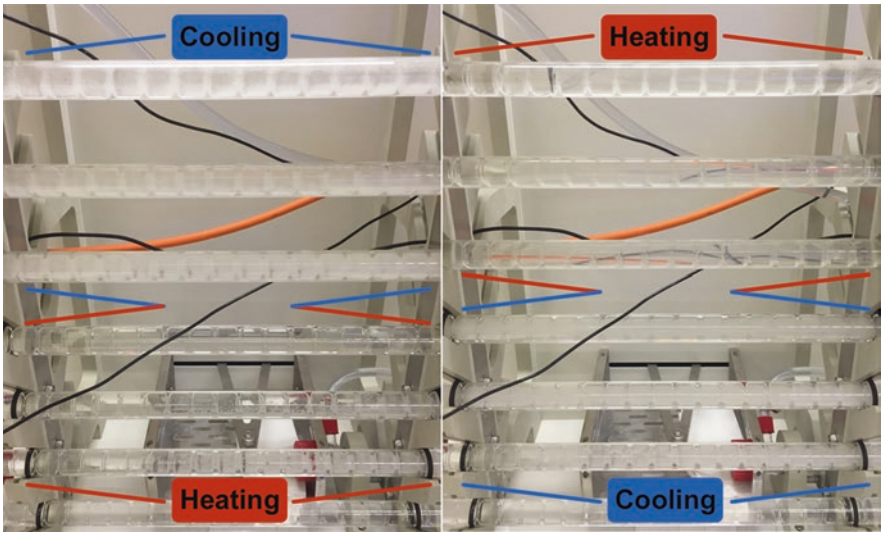


Fig. 5.32 Periodic heating/cooling cycles for AFC control developed by Koswara and Nagy in an OFBC

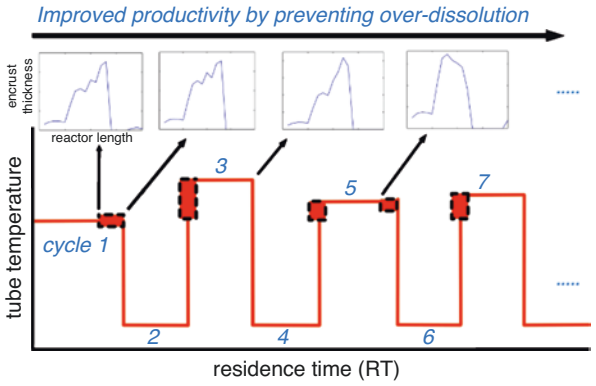


Fig. 5.33 Encrust thickness response to antifouling feedback controller (heating/cooling cycles) in an OFBC. (Permission obtained from Koswara and Nagy. Copyright © 2017a, b, IEEE)

encrustation and take actions to prevent or mitigate encrustation with little to no delays. PAT measurements can also be utilized for real-time model parameter estimation for continuous model improvement. Koswara and Nagy (2015 and 2017a, b) studied and patented a model-based antifouling control (AFC) strategy in a PFC by dividing its tube segments into two symmetric parts: one for maximizing crystal growth and the other for encrust dissolution while crystal dissolution is minimized. The two parts are periodically cycled (Fig. 5.32) under AFC supervision to achieve a periodic state-of-control operation without encrustation (Fig. 5.33). A challenge remains that off-spec products produced during control dynamics must be identified, separated, and preferably recycled to improve yield.

Modeling of PFC Continuous Crystallization

PBM can also be used to simulate continuous crystallization in the PFC with an added dependence along the length z of the crystallizer. f then becomes dependent on t, L, and z (i.e., $f = f(t, L, z)$). Assuming crystal nucleation and growth are the only two significant mechanisms, the population balance equation can be written as:

$$\frac{\partial(f)}{\partial t} + \frac{\partial(Gf)}{\partial L} + \frac{\partial(vf)}{\partial z} = B\delta\left(L - L_0\right) \tag{5.26}$$

where v is the velocity of the particles which equals to volumetric flow rate divided by cross-sectional area. Initial and boundary conditions must be written to solve Eq. (5.26). There are several seeding strategies commonly used in PFC operations: no seeding, initial seeding only, and continuous seeding.

If no seed:

$$@\, t = 0; f = 0 \text{ for all } z \text{ and } L \tag{5.27}$$

If initial seed only:

$$@\, t = 0, z = 0; f = \text{fn}\left(L\right) = \text{population distribution of seed} \tag{5.28}$$

If continuous seeding:

$$@\, z = 0; f = \text{fn}\left(L\right) = \text{population distribution of seed for all } t \tag{5.29}$$

Similarly, concentration also becomes dependent on both t and z:

$$\frac{\partial C}{\partial t} + \frac{\partial(vC)}{\partial z} = -3k_v\rho_c \int_{\infty}^{L_0} GL^2 f dL \tag{5.30}$$

Depending on the concentration profile in the PFC during start-up, initial condition can be written as:

$$\begin{aligned} &\text{if PFC starts with solution}, @\, t = 0, C = C_i\left(z\right) \\ &\text{if PFC starts with solvent only}, @\, t = 0, C = 0 \text{ for all } z \end{aligned} \tag{5.31}$$

Boundary condition is as follows:

$$@\, z = 0, C = C_{\text{in}} \text{ for all } t \tag{5.32}$$

where C_{in} represents feed concentration which is assumed to be constant.

During steady state, the first term of population balance Eq. (5.26) and mass balance Eq. (5.30) can be eliminated; the coupled system of equations becomes:

$$\begin{aligned}\frac{\partial(Gf)}{\partial L}+\frac{\partial(vf)}{\partial z}&=B\delta(L-L_0)\\ \frac{\partial(vC)}{\partial z}&=-3k_v\rho_c\int_{\infty}^{L_0}GL^2fdL\end{aligned}\tag{5.33}$$

from which the particle size distribution and concentration traveling through the length of the PFC can be solved.

5.3 Process Analytical Technology (PAT) for Crystallization

Process analytical technology is essential in the paradigm shift to the quality by design regulatory concept promoted by regulatory agencies including the US Food and Drug Administration (2004). The incorporation of process analytical technology provides a framework to optimize the process by enhancing the understanding of the process, identifying and diagnosing off-spec processes, and enabling feedback control strategies. PAT is especially critical to continuous processes to maintain a state-of-control operation. PATs can be in situ (or online), at-line, and off-line. In situ instruments provide real-time monitoring of the process quantitatively or qualitatively with almost no delay. In situ PAT sensors are usually invasive (i.e., must be inserted into the slurry system) and nondestructive (i.e., does not destroy samples). At-line instruments are located at close proximity to the process that poses a short delay but usually gives more accurate qualitative and/or quantitative results than in situ instruments. Off-line PATs are usually characterization instruments that provide difficult-to-measure qualitative information. At-line and off-line PAT tools are usually destructive.

Concentration is one of the most important quality attributes of a crystallization process. There are many in situ PAT tools that provide real-time concentration monitoring such as attenuated total reflection-ultraviolet/visible (ATR-UV/Vis) spectroscopy, attenuated total reflection-Fourier transform infrared (ATR-FTIR) spectroscopy, attenuated total reflection-near-infrared (ATR-NIR) spectroscopy, and Raman spectroscopy. ATR-UV/Vis is very useful in tracking single-component solution concentration. ATR-UV/Vis applies UV/Vis light and measures the reflectance. Solution of different concentrations reflects UV light differently. The calibration model is usually quite simple containing temperature and the absorbance of a single significant peak. Thus, the following calibration equation is often used:

$$C=a_1Abs+a_2T+a_3Abs\bullet T+b\tag{5.34}$$

where a_i denotes fitted coefficients, *Abs* is the absorbance at a certain wavenumber, and b is the fitted intercept. ATR-NIR and ATR-FTIR measure the reflectance of IR or near-IR light. Raman spectroscopy measures the Raman scattering of a laser at different Raman shifts to identify and quantify a compound. IR and Raman

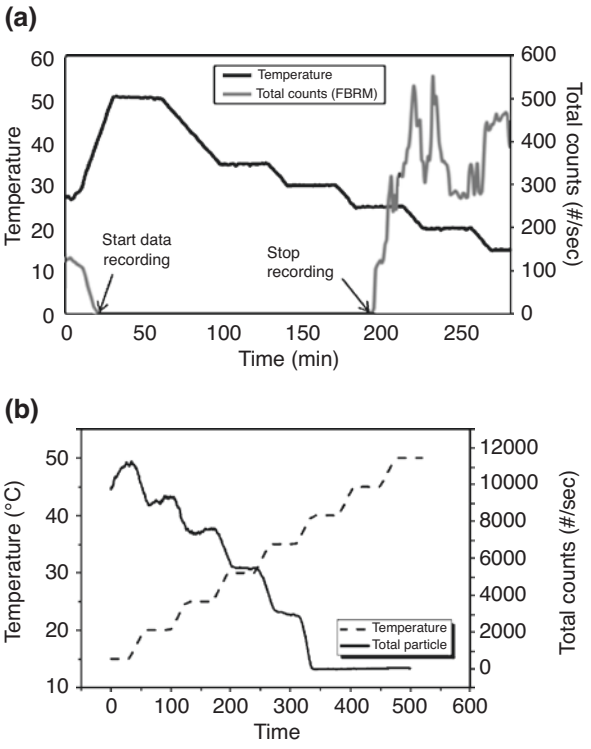


Fig. 5.34 Temperature and particle count profile of (**a**) detailed solution concentration calibration. (Permission obtained from Simone et al. Copyright © 2014 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim) and (**b**) rapid solution concentration calibration. (Permission obtained from Liu et al. © 2019a Elsevier B.V)

spectroscopy usually require a chemometric calibration model. IR spectrum is usually less affected by solid particle presence; however, experiments should still be carried out to test that solid concentration does not significantly affect the reading. In contrast, solid concentration usually has a significant impact on Raman scattering; thus certain Raman shifts that are heavily correlated solid concentration changes can be used to track solid concentration in slurry suspensions. IR spectrum sometimes can also be used to track solid concentration if correlated peaks can be identified.

To calibrate quantitative PAT tools such as UV, IR, and Raman for solution concentration measurement, a series of experiment should be carried out to reflect the effect of concentration, temperature, and sometimes solid concentration on the spectra. In other words, solutions of known concentrations should be measured at different temperatures to construct such calibration models. To do so efficiently, the following experiment is repeated for several concentrations (Simone et al. 2014): solution is first heated to ensure complete dissolution and maintained at 10 °C above its solubility temperature for 30 min. Then the temperature is decreased stepwise until nucleation with 10–30 min hold at each step (Fig. 5.34a) until nucleation. Nucleation can be detected by PAT tools such as the FBRM or the human eye which is less accurate. Per experiment, the absorbance is obtained at different temperatures

of the same concentration. Once the same experiment is repeated for different concentrations, a calibration model can be developed to calculate concentration at a certain temperature from the spectrum reading. Additional experiments can be carried out to vary solid concentration at fixed concentration to observe the solid concentration effects on the spectrum. A faster but not as accurate method is also commonly used to calibrate concentration measurements. This method assumes that the solubility curve against temperature is known. The solution is first held at the lower limit of the temperature range of interest with excess solute material (i.e., the system starts as a slurry). Then the solution is heated stepwise slowly with 30 min hold at each step to ensure equilibrium (Fig. 5.34b). Thus, the solution concentration at each temperature step can be assumed as the solubility concentration (Liu et al. 2019a). This method is commonly known as the rapid calibration method. Its disadvantage is that it does not decouple temperature and concentration effects on the spectrum nor does it consider the solid effect on the spectra reading. However, it can be sufficient if the spectrum is sensitive to concentration change without much interference from solid suspension. If an antisolvent crystallization process is intended, the solvent-to-antisolvent ratio (SASR) should be varied instead of temperature.

Solid properties including size, shape, and polymorphic form are important for crystallization processes. Focused beam reflectance measurement (FBRM), turbidity meter, ultrasound measurement, acoustic measurement, and endoscopes are useful PAT tools that give quantitative and/or qualitative information about the solid properties.

FBRM is a calibration-free tool that measures solid particle count and particle chord length. It has gained popularity over the last decade for its monitoring and control applications. It consists of a rotating laser optics in a sapphire window (Fig. 5.35a) to measure laser back scattering corresponding to the chord length distribution (Fig. 5.35b) and the particle count per unit time of its rotation (Fig. 5.35c). It is important to notice that FBRM measurement of particle count and chord length distribution does not quantitatively describe the number of particles in the crystallizer or the solid particle size distribution, but the trend is indicative of the changes in particle population. For example, sudden increase of particle counts accompanied by mean chord length decrease indicates nucleation event (Fig. 5.36a); mean chord length increasing overtime suggests possible crystal growth (Fig. 5.36b); FBRM coupled with concentration gives great insight into the process that is otherwise difficult to observe. FBRM can also be used for ADNC where particle count is monitored via FBRM and controlled by manipulating supersaturation. An example of ADNC application in batch crystallization of paracetamol is shown in Fig. 5.37. ADNC can also be used to maintain a state-of-control in continuous crystallization processes as mentioned in section "State-of-Control Operation".

Turbidity meter is another tool that measures solid properties. Turbidity is a phenomenon where a solution loses its transparency due to the presence of suspended solids in the slurry (Fig. 5.38). The turbidity meter does not give detailed particle size distribution (PSD) or chord length information and is often used qualitatively

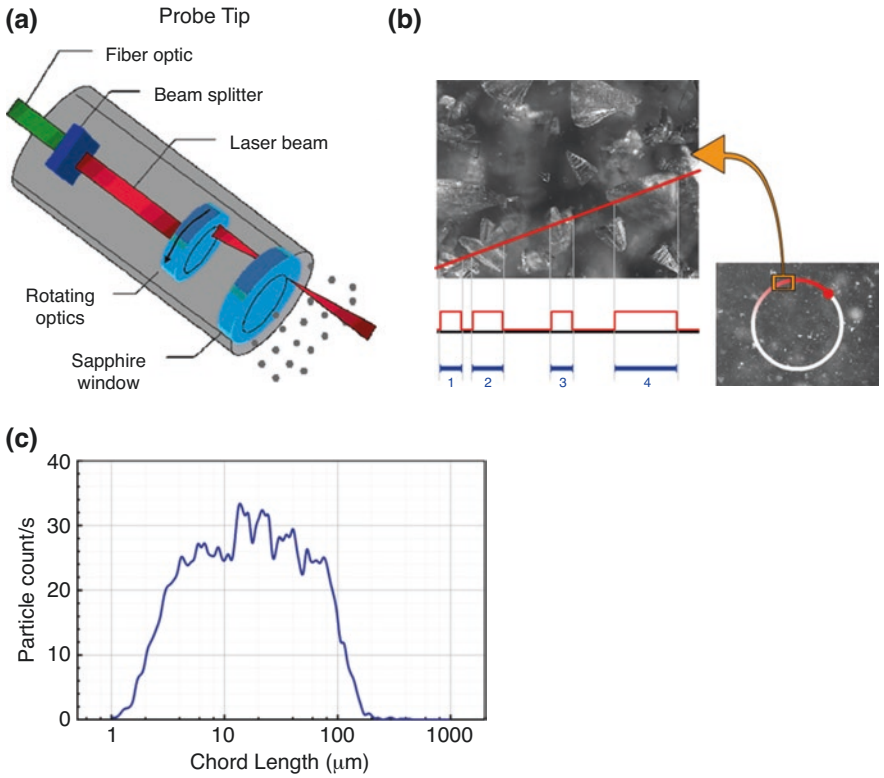


Fig. 5.35 FBRM probe (**a**) schematics, (**b**) chord length measurement mechanism, and (**c**) typical chord length distribution measurement

to identify the presence of solids during metastable zone measurement experiments (Liang et al. 2004). There are three types of turbidity measurements: adsorption (fixed sample volume, medium-high solid concentration), forward scattering (fixed sample volume, low solid concentration), and backward scattering (simple design, open measuring zone). Acoustic emission (AE) measurement is another PAT tool that can monitor solid properties by measuring the elastic energy change of the acoustic wave induced by dynamic changes such as crystallization.

In contrast to the other online PATs, AE sensing method is noninvasive as shown in Fig. 5.39a. In other words, it does not come into contact with the slurry content because it does not require an observation window. It allows analysis of opaque samples that are difficult to monitor with FBRM or other laser technologies. It is intrinsically safe and relatively inexpensive, and it can be useful in both MSMPR and PFC operations. However, AE is a less commonly applied technology in the pharmaceutical manufacturing industry mostly due to its largely multivariate measurement data (Fig. 5.39b) which requires advanced multivariate statistical method to analyze and calibrate.

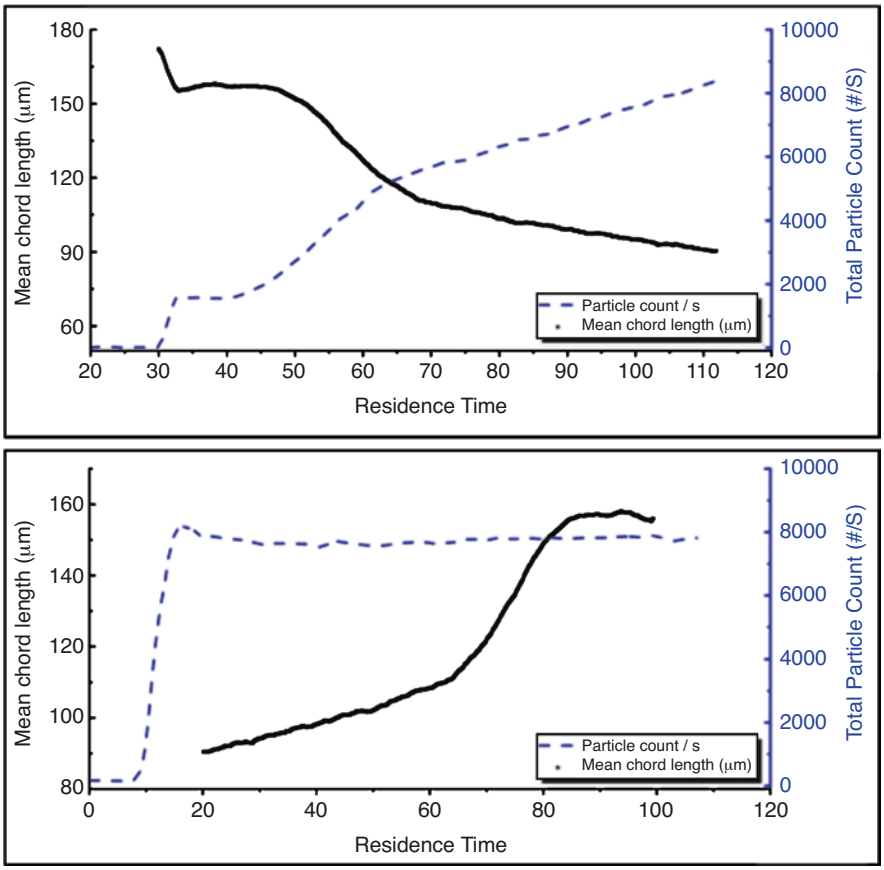


Fig. 5.36 Typical FBRM profile of nucleation dominated crystallization (top) and growth dominated crystallization (bottom)

Endoscopy or in situ video monitoring is another useful PAT tool that gives useful visual information of the system. Particle vision measurement (PVM) is a commonly used probe that gives real-time microscopic images (examples are shown in Fig. 5.40) of the crystals that can be used to visually detect nucleation, growth, agglomeration, and polymorphic transformation among other events (Barrett and Glennon 2002; Hou et al. 2014; Schöll et al. 2006; Simon et al. 2009). Image analysis can be applied to provide some quantitative information; however it is difficult to obtain reliable quantitative information when solid concentration is at medium to high level as the microscopic images become overcrowded and particle images become overlapped (Borsos et al. 2017; Zhou et al. 2009).

The aforementioned Raman spectroscopy can also be used to monitor solid properties including solid concentration and solid polymorphic form in addition to solution concentration. Different polymorphic forms of crystalline particles yield different Raman readings due to differences in their molecular rotational or

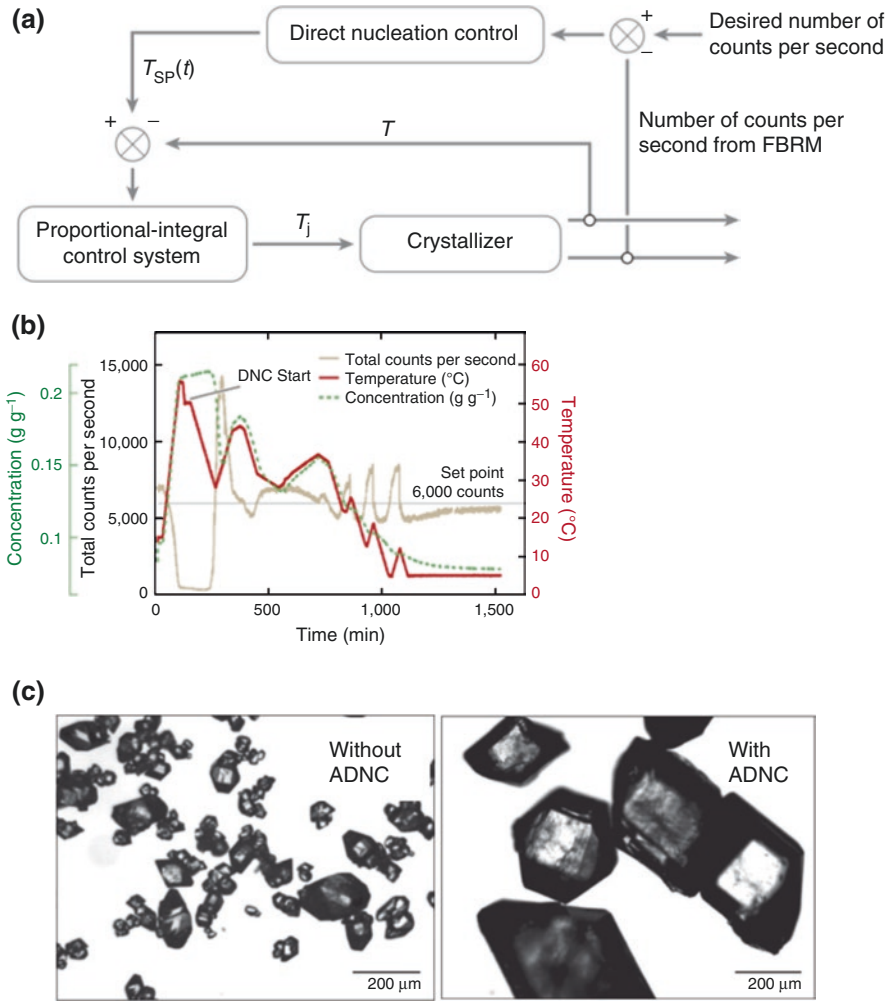


Fig. 5.37 (**a**) Control diagram of ADNC, (**b**) operating profile of an ADNC crystallization process, and (**c**) microscopic images of paracetamol batch crystallization products with and without ADNC. (Permission obtained from Saleemi et al. Copyright © 2012, American Chemical Society)

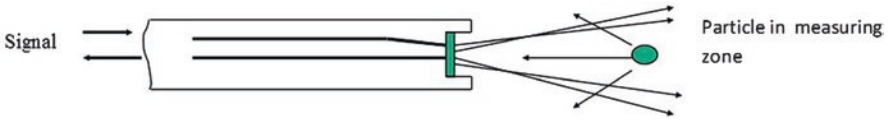


Fig. 5.38 Turbidity meter measurement mechanism

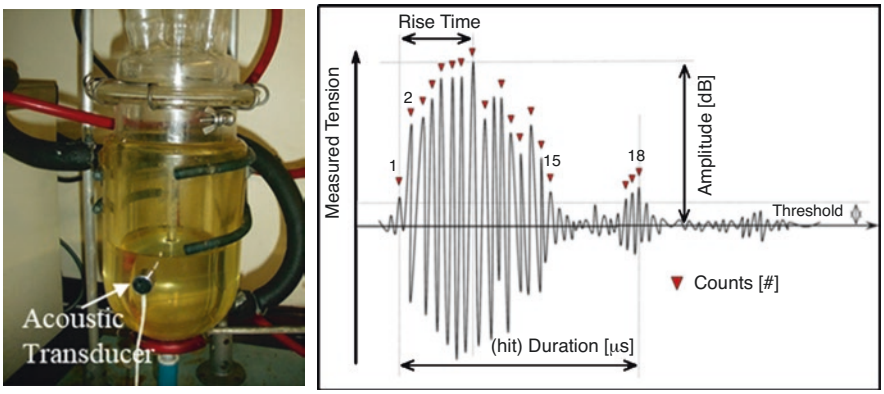


Fig. 5.39 A noninvasive acoustic transducer implemented on a laboratory-scale crystallizer and its main characteristic parameters of a typical AE hit. (Permission obtained from Gherras et al. Copyright © 2012 Elsevier B.V)

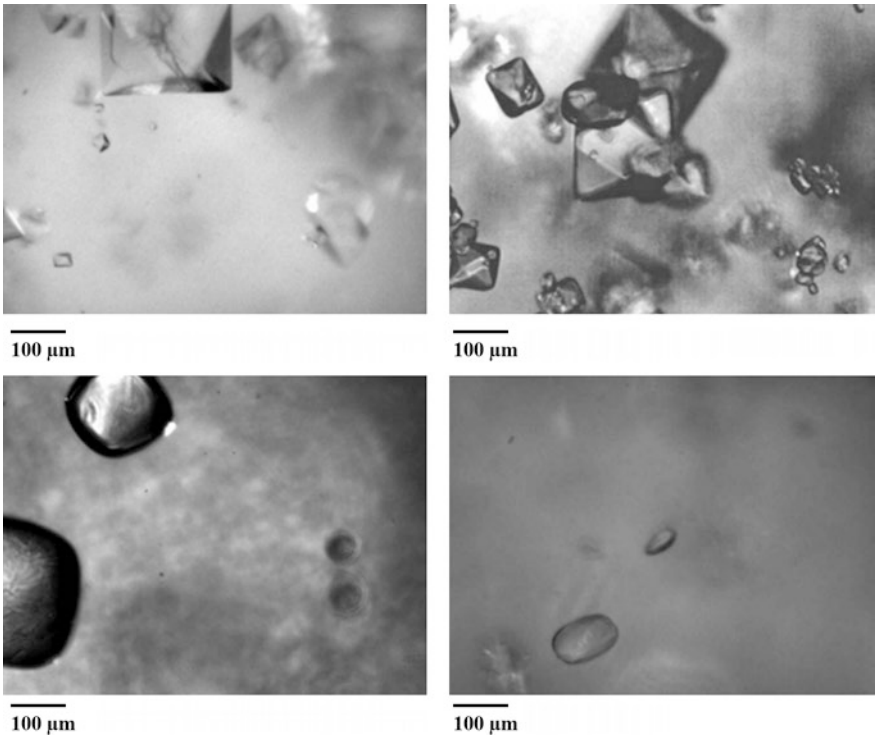


Fig. 5.40 PVM images of particles during crystallization and dissolution. (Permission obtained from Barrett and Glennon, Copyright © 2002 The Institution of Chemical Engineers. Published by Elsevier B.V)

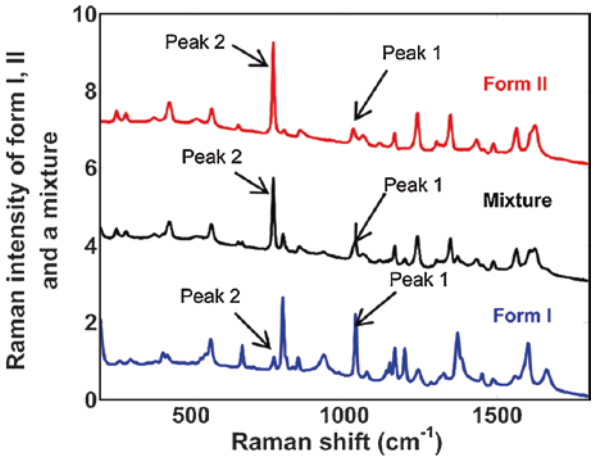


Fig. 5.41 In situ Raman spectrum of form I, form II, and mixture of ortho-aminobenzoic acid (OABA). (Permission obtained from Simone et al. Copyright © 2014, American Chemical Society)

Table 5.2 Summary of in situ PAT tools

PAT	Measurement	Properties monitored
UV	UV spectrum	Solution concentration
FTIR, NIR	(Near) IR spectrum	Solution concentration, solid concentration, polymorphism
Raman	Raman scattering	Solution concentration, solid concentration, polymorphism
FBRM	Laser reflectance	Qualitative particle counts and chord length distribution
Turbidity	Turbidity	Solid concentration
AE	Acoustics	Solid concentration
PVM	Imaging	Particle shape, size, and agglomeration

vibrational modes which can be picked up by in situ Raman probes (Fig. 5.41) thus useful for polymorphic monitoring and control during crystallization (Acevedo et al. 2018; Simone et al. 2014; Simone et al. 2017). IR spectroscopy can also identify polymorphism of certain compounds. In situ PAT sensors are extremely useful during continuous crystallization operations. They can be incorporated in MSMPR systems in a similar fashion as in batch systems. It is important to carefully place the sensors at a well-mixed spot to avoid fouling, while ensuring complete submersion throughout the entirety of the operation. In the case of PFC, it is more difficult to place PAT tools due to a lack of readily available ports, space limitation, and potential disturbance of the flow behavior. Noncontact sensors such as noncontact Raman and acoustic measurement are preferred. A summary of commonly used in situ PAT tools for crystallization is listed in Table 5.2.

Relative complex analytical tools that are unable to fit in the crystallizer yet have a relatively short measurement time (usually in a matter of minutes) are usually placed at-line of the process. These tools may analyze solution concentration, identify impurities, and/or measure solid properties. Ultra-performance liquid

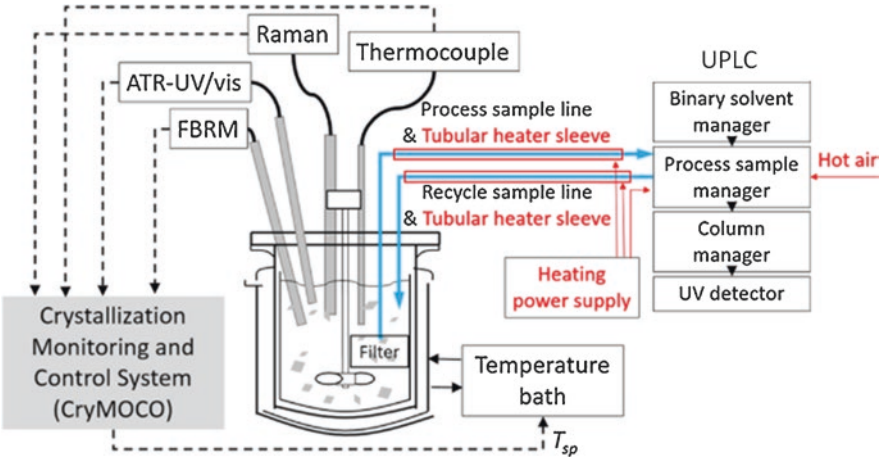


Fig. 5.42 At-line concentration monitoring of a crystallization by UPLC. (Permission obtained from Yang et al. Copyright © 2016a, b, American Chemical Society)

chromatography (UPLC), mass spectrometry (MS), and gas chromatography (GC) are examples of chemical analytical tools with at-line measurement capabilities. They have been well developed in the field of analytical chemistry, and there are many handbooks and literature on their method development. Nevertheless, it is usually difficult to maintain their automatic at-line operations because they are intricate instruments and are prone to blockage. Particles and/or gas bubble may get entrained into the instruments from the sampling channel which will likely cause measurement failure and even instrument damage. As a result, they are demonstrated more successful for liquid only upstream reaction process monitoring and are not commonly used for crystallization processes. However, there have been studies that demonstrated the usage of a PATROL UPLC by Waters for at-line concentration monitoring of crystallization processes. Y. Yang et al. studied a UPLC setup (Fig. 5.42) that was equipped with a heated autosampling line with a filter placed at the inlet to prevent particles from entering the lines. The authors successfully established process monitoring and feedback control strategies with at-line UPLC for crystallization process in two separate studies (Yang et al. 2017, b; Yang et al. 2016a, b). Issues with blockage were not discussed in those studies. Manual sampling is also possible which largely prevents particle entrainment in the sampling line by properly treating the sample before measurement but significantly prolongs sampling time, while introducing hard-to-monitor human errors.

Laser diffraction wet dispersion sizing technology such as Malvern Mastersizer and Sympatec is also capable of at-line analysis mostly by manual sampling. They are nondestructive instruments that measure volume-based PSD by measuring the laser relative transmission of the dispersed sample in an insoluble liquid as shown in Fig. 5.43. Laser diffraction measurement only provides volume-based PSD which does not directly correspond to the population distribution function f which is number based. Volume-based PSDs can be converted to number-based PSD under

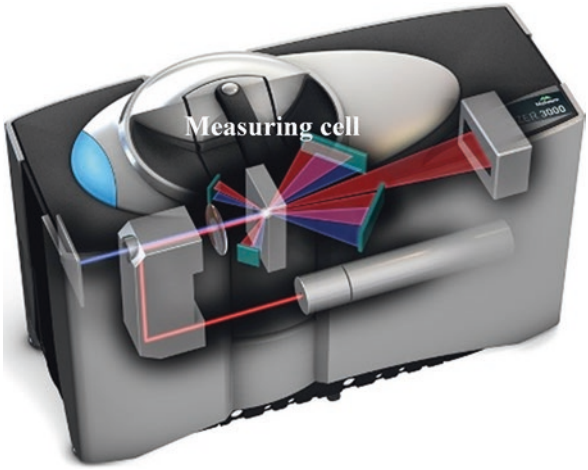


Fig. 5.43 Laser diffraction PSD measurement mechanism

certain assumptions, but it is not very accurate in the small/fine particle range. It is also important to note that laser diffraction sizing technology is fundamentally different from FBRM and does not yield comparable results as the FBRM: laser diffraction instruments measure solid volume concentration, and FBRM measures chord length. Some wet dispersion models offer flow cell capability for automatic at-line or in-line monitoring; however, it is usually difficult to maintain stable sample obscurity for reliable at-line reading.

Off-line tools are instruments that usually provide comprehensive and accurate information about the system but take longer to process or are very large in size. They are excellent tools to measure properties that are otherwise difficult to measure in situ or at-line such as multicomponent concentration, molecular structure, and thermostability, whose results can also serve as validation data for in situ PAT sensors. Off-line tools that are helpful for crystallization processes include high-performance liquid chromatography (HPLC) for purities, X-ray diffraction (XRD) for polymorph identification, microscopy (such as scanning electron microscopy SEM) for imaging, nuclear magnetic resonance (NMR) for impurity identification, dry dispersion laser diffraction sizing measurement and image-based sizing technology for PSD, and so on. At-line instruments can be used off-line as well. A summary of at-line and off-line analytical tools is listed in Table 5.3.

5.4 Integration of Continuous Crystallization and Downstream Operations

Operated as a slurry, crystallization is inseparable from its immediate unit operations such as filtration, drying, and granulation. Integration of continuous crystallization with downstream operations is an important step toward end-to-end continuous manufacturing. However, there are many challenges associated with the integration of multiple continuous unit operations including slurry transfer, spatial

Table 5.3 Summary of at-line and off-line PAT tools

PAT	Measurement	Properties Monitored
HPLC, UPLC	Liquid chromatography	Multicomponent identification and analysis
MS-GC	Gas chromatography	Multicomponent identification and analysis
Malvern Mastersizer	Laser diffraction	Volume-based PSD
Sympatec	Laser diffraction	Volume-based PSD
Malvern Morphologi	Imaging	Number-based PSD
XRD	X-ray diffraction	Solid structure, polymorphism
NMR	Magnetic resonance	Compound identification
SEM	Microscopy	Particle microscopic structure

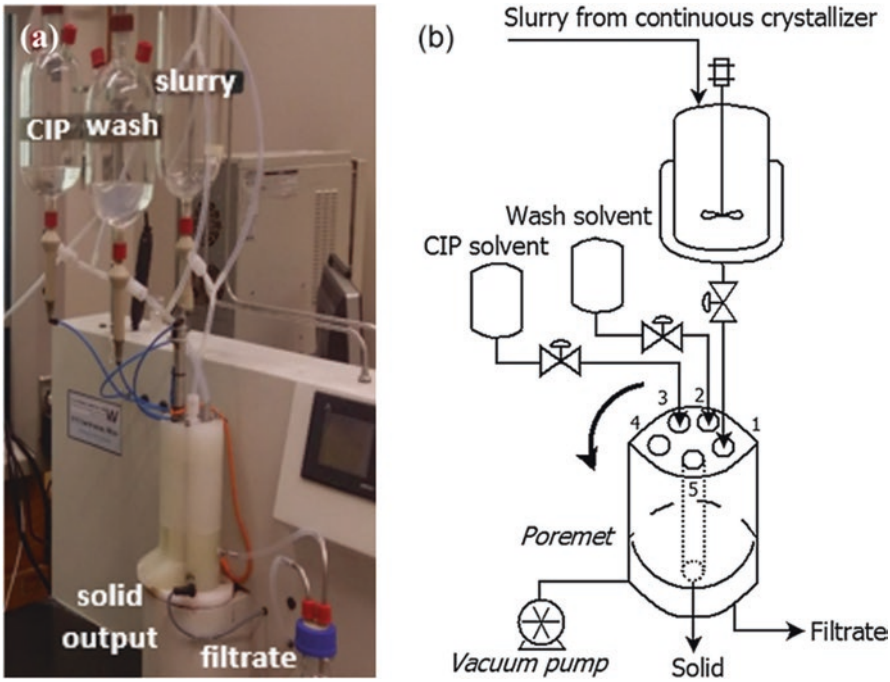


Fig. 5.44 (**a**) A picture of CFC manufactured by AWL and (**b**) schematics of the CFC. (Permission obtained from Acevedo et al. © 2016 Elsevier B.V)

constraints, and scheduling. The immediate unit operation that follows crystallization is filtration. Filtration is a difficult unit operation to run continuously, and most continuous crystallization-filtration studies apply an alternating semi-batch filtration system. Acevedo et al. (2016) were among the first to demonstrate a commercially available continuous filtration carousel (CFC system from Alconbury Weston Ltd., AWL) which is shown in Fig. 5.44. It consists of a wash solvent tank, a cleaning in place (CIP) solvent tank, and a five-port filtration carousel. A Poremet metal

Fig. 5.45 A continuous filtration drying unit developed by AWL. (Permission obtained from Ottoboni et al. © 2019 American Pharmacists Association®. Published by Elsevier Inc.)

mesh filter is installed at the bottom of the carousel covering port 1, 2, 3, and 4, while leaving port 5 uncovered. The CFC withdraws slurry either directly from the crystallizer or from a holdup tank and dispenses the slurry into port 1; wash solvent and CIP solvent are dispended into port 2 and 3. Port 4 inlet is blanked. Vacuum is applied under the filter mesh to remove filtrate/solvent and air-dry the filter cake residing in port 1–4. Then the carousel is rotated one index counterclockwise, i.e., port 1 becomes port 2 for wash, port 2 becomes port 3 for CIP, etc. Port 5 is equipped with a piston at the top to push the filter cake into the collector vessel at the bottom. The coupling of an MSMPR crystallizer and a CFC was demonstrated in the study to obtain filtered paracetamol particles of an average moisture content of 22%. A similar unit with the addition of drying abilities has also been developed by AWL, and an initial study comparing to a batch bench-scale vacuum filtration unit has been carried out by Ottoboni et al. (2019) (Fig. 5.45). Liu et al. (2019b) later sumnmarized a risk diagram for the implementation of continuous fitlration.

Process intensification is another technique to integrate downstream operations into crystallization and combine into a single step. Yazdanpanah et al. (2016) studied a novel falling film crystallization technique that combines crystallization and

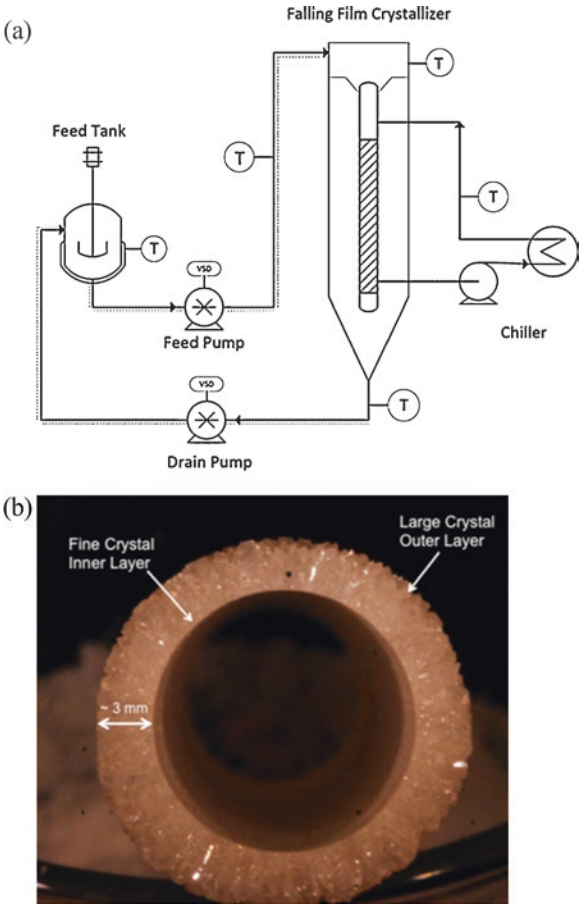


Fig. 5.46 (**a**) Schematic of falling film crystallizer setup and (**b**) crystal film deposited. (Permission obtained from Yazdanpanah et al. Copyright © 2016, American Chemical Society)

removal of solvent in one step that eliminates a subsequent filtration step as shown in Fig. 5.46a. The film (Fig. 5.46b) can then be redissolved for the next step. However, it is difficult to isolate the filtered solid in this setup and should only be used for intermediate purification crystallization steps. Another process intensification technique combines crystallization with granulation to produce spherical agglomerates in one step. Combining crystallization with agglomeration can be useful for difficult-to-filter crystals such as needle/plate-shaped crystals and/or slow-growing compounds like proteins. By adding a bridging agent, spherical agglomerates of small nuclei are produced at the end of the crystallization process which significantly improves the filterability of the compound. Peña et al. (2015 and 2017a, b) demonstrated continuous spherical agglomeration of benzoic acid in a two-stage MSMPR system where reverse antisolvent (i.e., primary nucleation dominated) crystallization took place in the first stage and the addition of bridging liquid occurred in the second stage to produce spherical agglomerates as shown in Fig. 5.47a. The properties of the agglomerates can be tailored by optimizing process

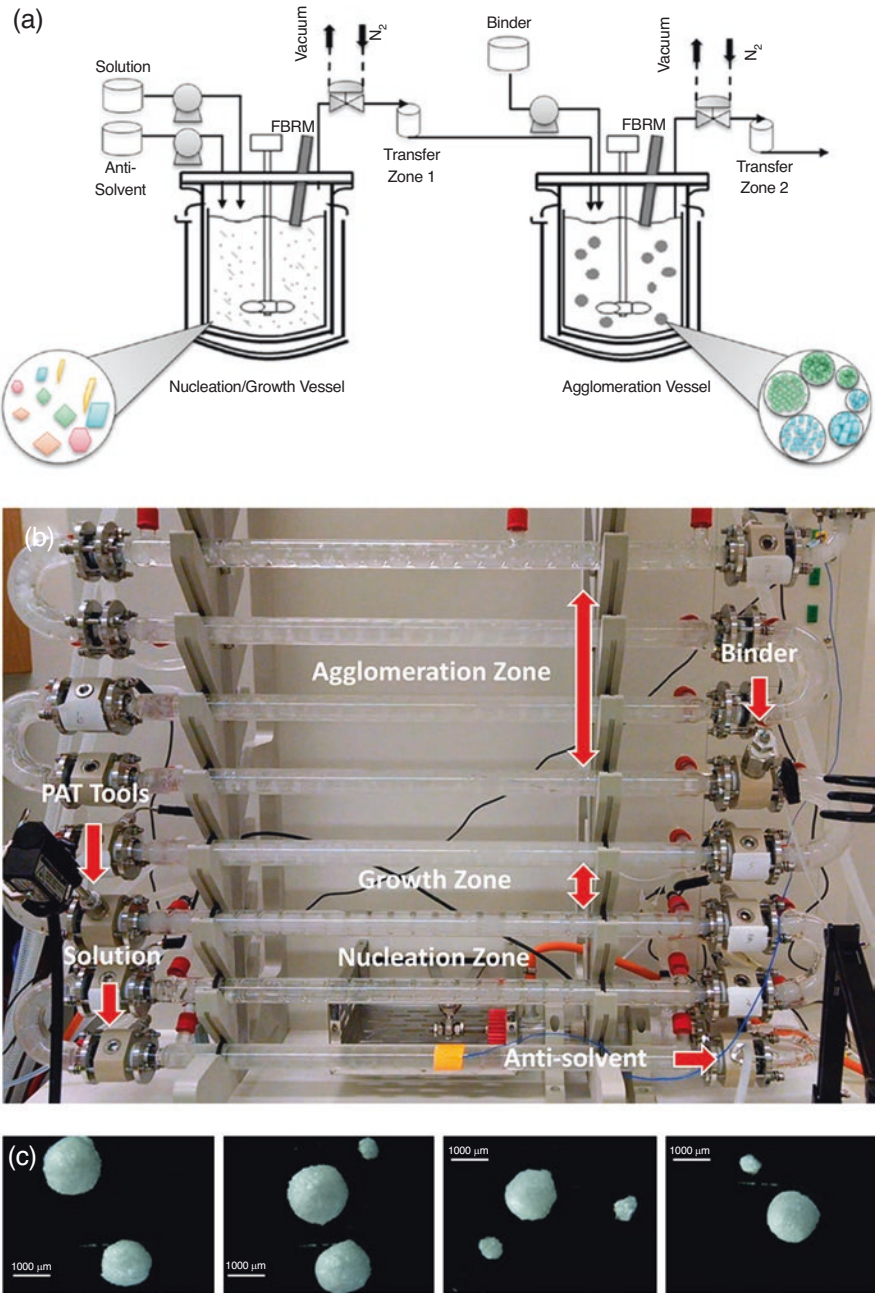


Fig. 5.47 Process intensification through continuous spherical agglomeration in (**a**) two-stage MSMPR. (Permission obtained from Pena and Nagy Copyright © 2015, American Chemical Society), (**b**) an OFBC, and (**c**) microscopic images of the benzoic acid spherical agglomerates. (Permission obtained from Peña et al. Copyright © 2017b, American Chemical Society)

parameters including feed concentration, SASR, bridging liquid to solute ratio, agitation rate, operating volume, and residence time. Later the same group demonstrated spherical agglomeration in the OFBC (Fig. 5.47b). Different zones were designated along the length of the OFBC: nucleation zone, growth zone, and agglomeration zone. The length of each zone was adjusted by changing the binder addition point location, essentially tailoring the residence time of crystallization and agglomeration separately. The primary particle size and the agglomerate size were controlled by manipulating feed concentration, antisolvent ratio, and residence times. Example microscopic images of these spherical agglomerates are shown in Fig. 5.47c.

One of the strongest arguments for the case of continuous crystallization development is the consistent API product properties, but it can be limited by returning to batch operation in subsequent filtration and drying steps, introducing batch-to-batch variations back into the process which negates the purpose of continuous crystallization. Coherent integration of continuous crystallization into the overall manufacturing process is crucial to achieving end-to-end CM which remains a challenge.

5.5 Conclusion

Crystallization is a common unit operation used in pharmaceutical manufacturing often as the final drug substance manufacturing step to purify and isolate APIs. Crystallization has substantial impact on downstream efficiency and final drug product quality. It is a critical but challenging step in developing end-to-end continuous manufacturing. Significant progress has been made in recent years on the process development of continuous crystallization of pharmaceuticals. MSMPR and PFC crystallizers are both excellent platforms for continuous crystallizations, each with their unique advantages and challenges. Solid transfer has been identified by researchers and engineers studying this topic as the main challenge to achieve sustainable continuous crystallization operations. The tendency to foul during continuous crystallization is caused by the low flow rates required to meet the pharmaceutical demand along with the inevitably present supersaturation. In MSMPR operations, solid fouling usually occurs in the transfer line; thus, the utilization intermittent transfer can largely improve this issue. Applying sophisticated transfer zone design may be necessary depending on the system. In PFCs, fouling usually occurs on the equipment wall (i.e., encrustation) causing process deviation and even complete blockage of the crystallizer. Depending on the kinetics and severity of encrustation, advanced control strategies may be required to ensure successful operation over time. Another challenge is the integration of continuous crystallization with downstream processes which requires further investigation. With the support of advanced PAT methods and population balance modeling, an optimized continuous process can be developed, and with further robust testing, achieving sustainable end-to-end continuous manufacturing is entirely achievable.

References

Aamir E, Nagy ZK, Rielly CD. Evaluation of the effect of seed preparation method on the product crystal size distribution for batch cooling crystallization processes. Cryst Growth Des [Internet]. 2010;10(11):4728–40. Available from: https://pubs.acs.org/doi/10.1021/cg100305w.

Acevedo D, Peña R, Yang Y, Barton A, Firth P, Nagy ZK. Evaluation of mixed suspension mixed product removal crystallization processes coupled with a continuous filtration system. Chem Eng Process Process Intensif [Internet]. 2016[cited 2017 Feb 21];108:212–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0255270116302926

Acevedo D, Yang X, Mohammad A, Pavurala N, Wei-Lee Wu, O'Connor TF, Nagy ZK, Cruz CN. Raman spectroscopy for monitoring the continuous crystallization of carbamazepine. Org Process Res Dev [Internet]. 2018;22(2):156–65

Acevedo D, Yang X, Liu YC, O'Connor TF, Koswara A, Nagy ZK, et al. Encrustation in continuous pharmaceutical crystallization processes—a review. Org Process Res Dev [Internet]. 2019;23(6):1134–42. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.9b00072

Agnew LR, Cruickshank DL, McGlone T, Wilson CC. Controlled production of the elusive metastable form II of acetaminophen (paracetamol): a fully scalable templating approach in a cooling environment. Chem Commun [Internet]. 2016;52(46):7368–71. Available from: http://xlink.rsc.org/?DOI=C6CC01032F.

Agnew LR, McGlone T, Wheatcroft HP, Robertson A, Parsons AR, Wilson CC. Continuous crystallization of paracetamol (Acetaminophen) Form II: selective access to a metastable solid form. Cryst Growth Des [Internet]. 2017;17(5):2418–27. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.6b01831.

Alvarez AJ, Myerson AS. Continuous plug flow crystallization of pharmaceutical compounds. Cryst Growth Des [Internet]. 2010;10(5):2219–228

Alvarez AJ, Singh A, Myerson AS, 64849 NL. Crystallization of cyclosporine in a multistage continuous MSMPR crystallizer. Growth Des [Internet]. 2011 [cited 2017 Apr 9];11(10):4392–400. Available from: http://pubs.acs.org/doi/pdf/10.1021/cg200546g.

Barrett P, Glennon B. Characterizing the metastable zone width using Lasentec FBRM and PVM. Chem Eng Res Des [Internet]. 2002;80(7):799–805. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0263876202722521.

Berglund KA. Analysis and measurement of crystallization utilizing the population balance. In: Handbook of Industrial Crystallization. Butterworth-Heinemann; 2002;101–13. https://doi.org/10.1016/B978-075067012-8/50006-9

Becker R, Döring W. Kinetische Behandlung der Keimbildung in übersättigten Dämpfen. Ann Phys [Internet]. 1935 [cited 2019 Mar 11];416(8):719–52. Available from: http://doi.wiley.com/10.1002/andp.19354160806.

Bennema P. The importance of surface diffusion for crystal growth from solution. J Cryst Growth. 1969;5:29–43.

Benyahia B, Lakerveld R, Barton PI. A plant-wide dynamic model of a continuous pharmaceutical process. Ind Eng Chem Res [Internet]. 2012;51(47):15393–412. Available from: http://pubs.acs.org/doi/10.1021/ie3006319.

Biscans B. (2004) Impact attrition in crystallization processes. Analysis of repeated impacts events of individual crystals. Powder Tech 2004;143-144:264–72

Bohlin M, Rasmuson ÅC. Application of controlled cooling and seeding in batch crystallization. Can J Chem Eng [Internet]. 1992;70(1):120–6. Available from: http://doi.wiley.com/10.1002/cjce.5450700117.

Borsos A, Szilágyi B, Agachi PS, Nagy ZK. Real-time image processing based online feedback control system for cooling batch crystallization. Org Process Res Dev [Internet]. 2017;21(4):511–19

Borsos A, Majumder A, Nagy ZK. Multi-impurity adsorption model for modeling crystal purity and shape evolution during crystallization processes in impure media. Cryst Growth Des [Internet]. 2016;16(2):555–68. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b00320.

Botsaris GD. Secondary nucleation — a review. In: Industrial crystallization [Internet]. Boston: Springer US; 1976 [cited 2019 Mar 11]. p. 3–22. Available from: http://link.springer.com/10.1007/978-1-4615-7258-9_1.

Brown CJ, Ni X-W. Evaluation of growth kinetics of antisolvent crystallization of paracetamol in an oscillatory baffled crystallizer utilizing video imaging. Cryst Growth Des [Internet]. 2011;11(9):3994–4000. Available from: https://pubs.acs.org/doi/10.1021/cg200560b.

Brown CJ, Ni X-W. Determination of metastable zone width, mean particle size and detectable number density using video imaging in an oscillatory baffled crystallizer. Cryst Eng Comm [Internet]. 2012;14(8):2944. Available from: http://xlink.rsc.org/?DOI=c2ce06628a.

Brown CJ, Lee YC, Nagy ZK, Ni X. Evaluation of crystallization kinetics of adipic acid in an oscillatory baffled crystallizer. Cryst Eng Comm [Internet]. 2014 [cited 2019 Mar 11];16(34):8008. Available from: http://xlink.rsc.org/?DOI=C4CE00192C.

Brown CJ, Adelakun JA, Ni X wei. Characterization and modelling of antisolvent crystallization of salicylic acid in a continuous oscillatory baffled crystallizer. Chem Eng Process Process Intensif [Internet]. 2015;97:180–6. Available from: https://doi.org/10.1016/j.cep.2015.04.012.

Brunold CR, Hunns JCB, Mackley MR, Thompson JW. Experimental observations on flow patterns and energy losses for oscillatory flow in ducts containing sharp edges. Chem Eng Sci. 1989;44(5):1227–44.

Burton WK, Cabrera N, Frank FC. The growth of crystals and the equilibrium structure of their surfaces. Philos Trans R Soc A Math Phys Eng Sci [Internet]. 1951;243(866):299–358. Available from: http://rsta.royalsocietypublishing.org/cgi/doi/10.1098/rsta.1951.0006.

Cao Y, Acevedo D, Nagy ZK, Laird CD. Real-time feasible multi-objective optimization based nonlinear model predictive control of particle size and shape in a batch crystallization process. Control Eng Pract [Internet]. 2017 [cited 2017 Nov 14];69:1–8. Available from: http://www.sciencedirect.com/science/article/pii/S0967066117301958.

Chakraborty D, Patey GN. Evidence that crystal nucleation in aqueous NaCl solution occurs by the two-step mechanism. Chem Phys Lett [Internet]. 2013;587:25–9. Available from: https://doi.org/10.1016/j.cplett.2013.09.054.

Chatterjee S. FDA perspective on continuous manufacturing. In: IFPAC annual meeting. Baltimore; 2012.

Gu C, Li H, Gandhi RB, Raghavan K. Grouping solvents by statistical analysis of solvent property parameters: implication to polymorph screening. Int J Pharm [Internet]. 2004;283(1-2):117–25

Chen J, Sarma B, Evans JMB, Myerson AS. Pharmaceutical crystallization. Cryst Growth Des [Internet]. 2011;11(4):887–95. Available from: https://pubs.acs.org/doi/10.1021/cg101556s.

Chew CM, Ristic RI. Crystallization by oscillatory and conventional mixing at constant power density. AIChE J [Internet]. 2005;51(5):1576–9. Available from: http://doi.wiley.com/10.1002/aic.10391.

Chianese A, Cave SD, Massarotta B. Investigation on some operating factors influencing batch cooling crystallization. In: Jancic SJ, Joiig EJ, editors., vol. 84; 1984. p. 443–6.

Chung SH, Ma DL, Braatz RD. Optimal seeding in batch crystallization. Can J Chem Eng [Internet]. 1999;77(3):590–6. Available from: http://doi.wiley.com/10.1002/cjce.5450770322

Cole KP, Groh JM, Johnson MD, Burcham CL, Campbell BM, Diseroad WD, et al. Kilogram-scale prexasertib monolactate monohydrate synthesis under continuous-flow CGMP conditions. Science (80-) [Internet]. 2017;356(6343):1144–50. Available from: http://www.sciencemag.org/lookup/doi/10.1126/science.aan0745.

Cui Y, O'Mahony M, Jaramillo JJ, Stelzer T, Myerson AS. Custom-built miniature continuous crystallization system with pressure-driven suspension transfeR. Org Process Res Dev [Internet]. 2016;20(7):1276–82. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.6b00113.

Davey R, Garside J. From molecules to crystallizers: an introduction to crystallization: Oxford Science; 2000. Oxford University press, New York. p. 1–52.

Davey RJ, Schroeder SLM, ter Horst JH. Nucleation of organic crystals-a molecular perspective. Angew Chemie Int Ed [Internet]. 2013;52(8):2166–79. Available from: http://doi.wiley.com/10.1002/anie.201204824.

Dickens AW, Mackley MR, Williams HR. Experimental residence time distribution measurements for unsteady flow in baffled tubes. Chem Eng Sci [Internet]. 1989 [cited 2017 Feb 21];44(7):1471–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/0009250989800235.

DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. J Health Econ [Internet]. 2003;22(2):151–85. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167629602001261.

Dombrowski RD, Litster JD, Wagner NJ, He Y. Crystallization of alpha-lactose monohydrate in a drop-based microfluidic crystallizer. Chem Eng Sci [Internet]. 2007;62(17):4802–10. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250907004162.

Eder RJP, Radl S, Schmitt E, Innerhofer S, Maier M, Gruber-Woelfler H, et al. Continuously seeded, continuously operated tubular crystallizer for the production of active pharmaceutical ingredients. Cryst Growth Des [Internet]. 2010;10(5):2247–57. Available from: https://pubs.acs.org/doi/10.1021/cg9015788.

Erdemir D, Lee AY, Myerson AS. Nucleation of crystals from solution: classical and two-step models. Acc Chem Res [Internet]. 2009;42(5):621–9. Available from: https://pubs.acs.org/doi/10.1021/ar800217x.

Feng L, Berglund KA. ATR-FTIR for determining optimal cooling curves for batch crystallization of succinic acid. Cryst Growth Des [Internet]. 2002;2(5):449–52. Available from: https://pubs.acs.org/doi/10.1021/cg025545e.

Ferguson S, Morris G, Hao H, Barrett M, Glennon B. Characterization of the anti-solvent batch, plug flow and MSMPR crystallization of benzoic acid. Chem Eng Sci [Internet]. 2013;104:44–54. Available from: https://doi.org/10.1016/j.ces.2013.09.006.

Ferguson S, Ortner F, Quon J, Peeva L, Livingston A, Trout BL, et al. Use of continuous MSMPR crystallization with integrated nanofiltration membrane recycle for enhanced yield and purity in API crystallization. Cryst Growth Des [Internet]. 2014;14(2):617–27. Available from: http://pubs.acs.org/doi/10.1021/cg401491y.

Fitch A. On the determination of axial dispersion coefficient in a batch oscillatory baffled column using laser induced fluorescence. Chem Eng J [Internet]. 2003;92(1–3):243–53. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1385894702002619.

Frank FC. The influence of dislocations on crystal growth. Discuss Faraday Soc [Internet]. 1949 [cited 2019 Mar 11];5(0):48. Available from: http://xlink.rsc.org/?DOI=df9490500048.

Frawley PJ, Mitchell NA, Ó'Ciardhá CT, Hutton KW. The effects of supersaturation, temperature, agitation and seed surface area on the secondary nucleation of paracetamol in ethanol solutions. Chem Eng Sci [Internet]. 2012;75:183–97. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250912002114.

Garside J. Industrial crystallization from solution. Chem Eng Sci [Internet]. 1985 [cited 2019 Mar 11];40(1):3–26. Available from: https://ac.els-cdn.com/0009250985850430/1-s2.0-0009250985850430-main.pdf?_tid=040284f9-da3c-40ca-b723-10622eec8f11&acdnat=1552349534_ad0c62f86d45d1da25f55bbbac56e905.

Garside J, Davey RJ. Invited review secondary contact nucleation: kinetics, growth and scale-up. Chem Eng Commun [Internet]. 1980;4(4–5):393–424. Available from: https://www.tandfonline.com/doi/full/10.1080/00986448008935918.

Garside J, Shah MB. Crystallization kinetics from msmpr crystallizers. Ind Eng Chem Process Des Dev [Internet]. 1980;19(4):509–14. Available from: http://pubs.acs.org/doi/abs/10.1021/i260076a001.

Garside J, Tavare NS. Mixing, reaction and precipitation: limits of micromixing in an MSMPR crystallizer. Chem Eng Sci. 1985;40(8):1485–93.

Ghez R, Gilmer GH. An analysis of combined surface and volume diffusion processes in crystal growth. J Cryst Growth [Internet]. 1974;21(1):93–109. Available from: https://linkinghub.elsevier.com/retrieve/pii/0022024874901572.

Gibbs JW. The collected works of J. Willard Gibbs. [Internet]. New Haven: Yale University Press; 1948 [cited 2019 Mar 11]. Available from: http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=agrono.xis&method=post&formato=2&cantidad=1&expresion=mfn=006899.

Griffin DW, Mellichamp DA, Doherty MF. Reducing the mean size of API crystals by continuous manufacturing with product classification and recycle. Chem Eng Sci [Internet]. 2010 [cited 2019 Mar 11];65(21):5770–80. Available from: https://www.sciencedirect.com/science/article/pii/S0009250910003246.

Green D, Myerson AS. Crystallizer Mixing: Understanding and Modeling Crystallizer Mixing and Suspension Flow. In: Handbook of Industrial Cyrstallization. Elsevier Science; 2001:182–199.

Gunawan R, Fusman I, Braatz RD. High resolution algorithms for multidimensional population balance equations. AIChE J [Internet]. 2004;50(11):2738–49. Available from: http://doi.wiley.com/10.1002/aic.10228.

Hewgill MR, Mackley MR, Pandit AB, Pannu SS. Enhancement of gas-liquid mass transfer using oscillatory flow in a baffled tube. Chem Eng Sci [Internet]. 1993;48(4):799–809. Available from: https://linkinghub.elsevier.com/retrieve/pii/000925099380145G.

Hou G, Power G, Barrett M, Glennon B, Morris G, Zhao Y. Development and characterization of a single stage mixed-suspension, mixed-product-removal crystallization process with a novel transfer unit. Cryst Growth Des. 2014;14(4):1782–93.

Hu Y, Liang JK, Myerson AS, Taylor LS. Crystallization monitoring by raman spectroscopy: simultaneous measurement of desupersaturation profile and polymorphic form in flufenamic acid systems. Ind Eng Chem Res [Internet]. 2005 [cited 2017 Nov 14];44(5):1233–40. Available from: http://pubs.acs.org/doi/pdf/10.1021/ie049745u.

Jang M, Zhu Z, Jimenez E, Papageorgiou CD, Waetzig J, Hardy A, et al. Continuous-flow tubular crystallization in slugs spontaneously induced by hydrodynamics. Cryst Growth Des [Internet]. 2014;14(2):851–60. Available from: http://pubs.acs.org/doi/10.1021/cg401715e.

Johnson MD, May SA, Calvin JR, Remacle J, Stout JR, Diseroad WD, et al. Development and scale-up of a continuous, high-pressure, asymmetric hydrogenation reaction, workup, and isolation. Org Process Res Dev [Internet]. 2012;16(5):1017–38. Available from: http://pubs.acs.org/doi/10.1021/op200362h.

Kacker R, Regensburg SI, Kramer HJM. Residence time distribution of dispersed liquid and solid phase in a continuous oscillatory flow baffled crystallizer. Chem Eng J [Internet]. 2017 [cited 2017 Mar 7];317:413–23. Available from: http://linkinghub.elsevier.com/retrieve/pii/S138589471730164X.

Karunanithi AT, Achenie LEK, Gani R. A computer-aided molecular design framework for crystallization solvent design. Chem Eng Sci [Internet]. 2006;61(4):1247–60

Kolář P, Shen J, Tsuboi A, Ishikawa T. Solvent selection for pharmaceuticals. Fluid Phase Equilibria 2002;194-197:771–782

Köllges T, Vetter T. Model-based analysis of continuous crystallization/reaction processes separating conglomerate forming enantiomers. Cryst Growth Des [Internet]. 2017;17(1):233–47. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.6b01487.

Koswara A, Nagy ZK. Anti-fouling control of plug-flow crystallization via heating and cooling cycle. IFAC-PapersOnLine [Internet]. 2015;28(8):193–8. Available from: https://doi.org/10.1016/j.ifacol.2015.08.180.

Koswara A, Nagy ZK. ON–OFF feedback control of plug-flow crystallization: a case of quality-by-control in continuous manufacturing. IEEE Life Sci Lett [Internet]. 2017a;3(1):1–4. Available from: http://ieeexplore.ieee.org/document/7837677/.

Koswara A, Nagy ZK. Systems with anti-fouling control and methods for controlling fouling within a channel of a plug flow crystallizer. United States; US 2017b / 0312795 A1, 2017.

Kubota N, Doki N, Yokota M, Sato A. Seeding policy in batch cooling crystallization. Powder Technol [Internet]. 2001;121(1):31–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0032591001003710.

Kumar S, Ramkrishna D. On the solution of population balance equations by discretization—I. A fixed pivot technique. Chem Eng Sci [Internet]. 1996;51(8):1311–32. Available from: https://linkinghub.elsevier.com/retrieve/pii/0009250996884892.

Lai T-TC, Ferguson S, Palmer L, Trout BL, Myerson AS. Continuous crystallization and polymorph dynamics in the <scp>l</scp> −glutamic acid system. Org Process Res Dev [Internet]. 2014;18(11):1382–90. Available from: http://pubs.acs.org/doi/10.1021/op500171n.

Lai T-TC, Cornevin J, Ferguson S, Li N, Trout BL, Myerson AS. Control of polymorphism in continuous crystallization via mixed suspension mixed product removal systems cascade design. Cryst Growth Des [Internet]. 2015;15(7):3374–82. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b00466.

Lakerveld R, Benyahia B, Braatz RD, Barton PI. Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. AIChE J [Internet]. 2013 [cited 2019 Mar 11];59(10):3671–85. Available from: http://doi.wiley.com/10.1002/aic.14107.

Lawton S, Steele G, Shering P, Zhao L, Laird I, Ni X-W. Continuous crystallization of pharmaceuticals using a continuous oscillatory baffled crystallizer. Org Process Res Dev [Internet]. 2009;13(6):1357–63. Available from: https://pubs.acs.org/doi/10.1021/op900237x.

Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov [Internet]. 2015;10(3):191–9. Available from: http://link.springer.com/10.1007/s12247-015-9215–8.

Li J, Trout BL, Myerson AS. Multistage continuous Mixed-Suspension, Mixed-Product Removal (MSMPR) crystallization with solids recycle. Org Process Res Dev [Internet]. 2016;20(2):510–6. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00306.

Liang K, White G, Wilkinson D, Ford LJ, Roberts KJ, Wood WML. An examination into the effect of stirrer material and agitation rate on the nucleation of -glutamic acid batch crystallized from supersaturated aqueous solutions. Cryst Growth Des [Internet]. 2004;4(5):1039–44

Lindenberg C, Mazzotti M. Continuous precipitation of L-asparagine monohydrate in a micromixer: estimation of nucleation and growth kinetics. AIChE J [Internet]. 2011 [cited 2019 Mar 11];57(4):942–50. Available from: http://doi.wiley.com/10.1002/aic.12326.

Lindenberg C, Krättli M, Cornel J, Mazzotti M, Brozio JJ, Krättli M, et al. Design and optimization of a combined cooling/antisolvent crystallization process. Cryst Growth Des [Internet]. 2009;9(2):1124–36. Available from: https://doi.org/10.1021/cg800934h%5Cnhttp://pubs.acs.org/doi/abs/10.1021/cg800934h.

Litster J. Design and processing of particulate products. Cambridge University Press Cambridge, UK. 2016.

Liu YC, Domokos A, Coleman S, Firth P, Nagy ZK, Development of continuous filtration in a novel continuous filtration carousel integrated with continuous crystallization. Org Process Res Dev [Internet]. 2019a;23(12):2655–65

Liu YC, Dunn D, Lipari M, Barton A, Firth P, Speed J, Wood D, Nagy ZK, A comparative study of continuous operation between a dynamic baffle crystallizer and a stirred tank crystallizer. Chem Eng Sci [Internet]. 2019b;367:278–94

Lu J, Litster JD, Nagy ZK. Nucleation studies of active pharmaceutical ingredients in an air-segmented microfluidic drop-based crystallizer. Cryst Growth Des [Internet]. 2015;15(8):3645–51. Available from: http://pubs.acs.org/doi/abs/10.1021/acs.cgd.5b00150.

Mackley MR, Ni X. Mixing and dispersion in a baffled tube for steady laminar and pulsatile flow. Chem Eng Sci [Internet]. 1991;46(12):3139–51. Available from: https://linkinghub.elsevier.com/retrieve/pii/000925099185017R.

Mackley MR, Stonestreet P. Heat transfer and associated energy dissipation for oscillatory flow in baffled tubes. Chem Eng Sci [Internet]. 1995;50(14):2211–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/000925099500088M.

Mackley MR, Tweddle GM, Wyatt ID. Experimental heat transfer measurements for pulsatile flow in baffled tubes. Chem Eng Sci. 1990;45(5):1237–42.

Majumder A, Nagy ZK. Dynamic modeling of encrust formation and mitigation strategy in a continuous plug flow crystallizer. Cryst Growth Des [Internet]. 2015;15(3):1129–40. Available from: http://pubs.acs.org/doi/10.1021/cg501431c.

Mazzarotta B. Abrasion and breakage phenomena in agitated crystal suspensions. Chem Eng Sci [Internet]. (1992);47(12):3105–111

McCabe WL, Stevens RP. Rate of growth of crystals in aqueous solutions. Chem Eng Prog. 1951;47(4):168–74.

McGlone T, Briggs NEBB, Clark CA, Brown CJ, Sefcik J, Florence AJ. Oscillatory Flow Reactors (OFRs) for continuous manufacturing and crystallization. Org Process Res Dev [Internet]. 2015;19(9):1186–202. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00225.

Mersmann A, Rennie FW. Seeding. In: Crystallization technology handbook; 1970. p. 401–82.

Mesbah A, Paulson JA, Lakerveld R, Braatz RD. Model predictive control of an integrated continuous pharmaceutical manufacturing pilot plant. Org Process Res Dev [Internet]. 2017;21(6):844–54. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.7b00058.

Morris G, Power G, Ferguson S, Barrett M, Hou G, Glennon B. Estimation of nucleation and growth kinetics of benzoic acid by population balance modeling of a continuous cooling mixed suspension, mixed product removal crystallizer. Org Process Res Dev [Internet]. 2015;19(12):1891–902. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00139.

Mullin JW. Crystal growth. In: Crystallization; Elsevier; 2001a. Oxford, UK p. 216–88.

Mullin JW. Nucleation. In: Crystallization; Elsevier; 2001b. Oxford, UK p. 181–215.

Mullin JW. Industrial techniques and equipment. Crystallization (Fourth Edition), Butterworth-Heinemann 2001c, Oxford 315–402.

Myerson AS, Decker SE, Fan W. Solvent selection and batch crystallization. Ind Eng Chem Res Pro Des Dev. 1986;25(4):925–29.

Myerson AS. Nucleation kinetics. In: Handbook of industrial crystallization. 2nd ed. Elsevier Science; 2002. p. 50–52.

Nagy ZK, Braatz RD. Advances and new directions in crystallization control. Annu Rev Chem Biomol Eng [Internet]. 2012 [cited 2017 Nov 14];3(1):55–75. Available from: http://www.annualreviews.org/doi/10.1146/annurev-chembioeng-062011-081043.

Nagy ZK, Fevotte G, Kramer H, Simon LL. Recent advances in the monitoring, modelling and control of crystallization systems. Chem Eng Res Des [Internet]. 2013 [cited 2017 Nov 14];91(10):1903–22. Available from: http://www.sciencedirect.com/science/article/pii/S0263876213002955.

Nagy ZK, Aamir E. (2012) Systematic design of supersaturation controlled crystallization processes for shaping the crystal size distribution using an analytical estimator. Che Eng Sci [Internet]. 2012;84:656–70.

Narducci O, Jones AG, Kougoulos E. Continuous crystallization of adipic acid with ultrasound. Chem Eng Sci [Internet]. 2011;66(6):1069–76. Available from: https://doi.org/10.1016/j.ces.2010.12.008.

Ni X, Gao S. Scale-up correlation for mass transfer coefficients in pulsed baffled reactors. Chem Eng J Biochem Eng J [Internet]. 1996;63(3):157–66. Available from: https://linkinghub.elsevier.com/retrieve/pii/S092304679603120X.

Ni X, Gough P. On the discussion of the dimensionless groups governing oscillatory flow in a baffled tube. Chem Eng Sci [Internet]. 1997;52(18):3209–12. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250997001048.

Ni X, Liao A. Effects of mixing, seeding, material of baffles and final temperature on solution crystallization of l-glutamic acid in an oscillatory baffled crystallizer. Chem Eng J [Internet]. 2010;156(1):226–33. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1385894709007372.

Ni X, Mackley MR. Chemical reaction in batch pulsatile flow and stirred tank reactors. Chem Eng J [Internet]. 1993;52(3):107–14. Available from: http://linkinghub.elsevier.com/retrieve/pii/030094679380059W.

Ni X, Gao S, Pritchard DW. A study of mass transfer in yeast in a pulsed baffled bioreactor. Biotechnol Bioeng [Internet]. 1995a;45(2):165–75. Available from: http://doi.wiley.com/10.1002/bit.260450211.

Ni X, Gao S, Cumming RH, Pritchard DW. A comparative study of mass transfer in yeast for a batch pulsed baffled bioreactor and a stirred tank fermenter. Chem Eng Sci [Internet]. 1995b [cited 2017 Feb 21];50(13):2127–36. Available from: https://linkinghub.elsevier.com/retrieve/pii/000925099500050F.

Ni X, Gao S, Santangeli L. On the effect of surfactant on mass transfer to water-glycerol solutions in a pulsed baffled reactor. J Chem Technol Biotechnol [Internet]. 1997;69(2):247–53. Available from: http://doi.wiley.com/10.1002/%28SICI%291097-4660%28199706%2969%3A2%3C247%3A%3AAID-JCTB693%3E3.0.CO%3B2-Q.

Ni X, Cosgrove J, Arnott A, Greated C, Cumming R. On the measurement of strain rate in an oscillatory baffled column using particle image velocimetry. Chem Eng Sci [Internet]. 2000a [cited 2017 Feb 25];55(16):3195–208. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250999005771.

Ni X, Nelson G, Mustafa I. Flow patterns and oil-water dispersion in a 0.38 m diameter oscillatory baffled column. Can J Chem Eng [Internet]. 2000b;78(1):211–20. Available from: http://doi.wiley.com/10.1002/cjce.5450780127.

Ni X, Mackley MR, Harvey AP, Stonestreet P, Baird MHI, Rama Rao NV. Mixing through oscillations and pulsations—a guide to achieving process enhancements in the chemical and process industries. Chem Eng Res Des [Internet]. 2003;81(3):373–83. Available from: https://linkinghub.elsevier.com/retrieve/pii/S026387620372320X.

Ni X, Valentine A, Liao A, Sermage SBC, Thomson GB, Roberts KJ. On the crystal polymorphic forms of <scp>l</scp> −glutamic acid following temperature programmed crystallization in a batch oscillatory baffled crystallizer. Cryst Growth Des [Internet]. 2004;4(6):1129–35. Available from: https://pubs.acs.org/doi/10.1021/cg049827l.

Nyvlt J, Sohnel O, Matuchaova M, Broul M. The kinetics of industrial crystallization. Amsterdam: Elsevier; 1985.

Ohara M, Reid RC. Modeling crystal growth rates from solutions. 1973.

Ottoboni S, Price CJ, Steven C, Meehan E, Barton A, Firth P, et al. Development of a novel continuous filtration unit for pharmaceutical process development and manufacturing. J Pharm Sci [Internet]. 2019;108(1):372–81. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022354918304052.

Peña R, Nagy ZK. Process intensification through continuous spherical crystallization using a two-stage Mixed Suspension Mixed Product Removal (MSMPR) system. Cryst Growth Des [Internet]. 2015;15(9):4225–36. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b00479.

Peña R, Burcham CL, Jarmer DJ, Ramkrishna D, Nagy ZK. Modeling and optimization of spherical agglomeration in suspension through a coupled population balance model. Chem Eng Sci [Internet]. 2017a [cited 2017 Nov 14];167:66–77. Available from: http://www.sciencedirect.com/science/article/pii/S000925091730235X.

Peña R, Oliva JA, Burcham CL, Jarmer DJ, Nagy ZK. Process intensification through continuous spherical crystallization using an oscillatory flow Baffled crystallizer. Cryst Growth Des [Internet]. 2017b [cited 2018 Feb 6];17(9):4776–84. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.7b00731.

Plumb K. Continuous processing in the pharmaceutical industry. Chem Eng Res Des [Internet]. 2005;83(6):730–8. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0263876205727556.

Polster CS, Cole KP, Burcham CL, Campbell BM, Frederick AL, Hansen MM, et al. Pilot-scale continuous production of LY2886721: amide formation and reactive crystallization. Org Process Res Dev [Internet]. 2014;18(11):1295–309. Available from: http://pubs.acs.org/doi/10.1021/op500204z

Powell KA, Bartolini G, Wittering KE, Saleemi AN, Wilson CC, Rielly CD, et al. Toward continuous crystallization of urea-barbituric acid: a polymorphic co-crystal system. Cryst Growth Des [Internet]. 2015a;15(10):4821–36. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b00599

Powell KA, Saleemi AN, Rielly CD, Nagy ZK. Periodic steady-state flow crystallization of a pharmaceutical drug using MSMPR operation. Chem Eng Process Process Intensif [Internet]. 2015b;97:195–212. Available from: https://doi.org/10.1016/j.cep.2015.01.002.

Powell KA, Saleemi AN, Rielly CD, Nagy ZK. Monitoring continuous crystallization of paracetamol in the presence of an additive using an integrated PAT array and multivariate methods. Org Process Res Dev [Internet]. 2016;20(3):626–36. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00373

Power G, Hou G, Kamaraju VK, Morris G, Zhao Y, Glennon B. Design and optimization of a multistage continuous cooling mixed suspension, mixed product removal crystallizer. Chem Eng Sci [Internet]. 2015;133:125–39. Available from: https://doi.org/10.1016/j.ces.2015.02.014.

Puel F, Févotte G, Klein JP. Simulation and analysis of industrial crystallization processes through multidimensional population balance equations. Part 1: a resolution algorithm based on the method of classes. Chem Eng Sci [Internet]. 2003;58(16):3715–27. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250903002549.

Qamar S, Peter Elsner M, Hussain I, Seidel-Morgenstern A. Seeding strategies and residence time characteristics of continuous preferential crystallization. Chem Eng Sci [Internet]. 2012 [cited 2019 Mar 11];71:5–17. Available from: https://doi.org/10.1016/j.ces.2011.12.030.

Quon JL, Zhang H, Alvarez A, Evans J, Myerson AS, Trout BL. Continuous crystallization of Aliskiren Hemifumarate. Cryst Growth Des [Internet]. 2012;12(6):3036–44. Available from: http://pubs.acs.org/doi/10.1021/cg300253a.

Ramkrishna D. Population balances: Theory and applications to particulate systems in engineering. Elsevier, 2000.

Rashid A, White E, Howes T, Litster J, Marziano I. From raw data to process: the path to a batch or a continuous crystallizer design for Ibuprofen. Org Process Res Dev [Internet]. 2017;21(2):187–94. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.6b00339.

Ristic RI. Oscillatory mixing for crystallization of high crystal perfection pharmaceuticals. Chem Eng Res Des [Internet]. 2007;85(7):937–44. Available from: http://www.sciencedirect.com/science/article/pii/S0263876207731283.

Saleemi AN, Rielly CD, Nagy ZK. Comparative investigation of supersaturation and automated direct nucleation control of crystal size distributions using ATR-UV/vis spectroscopy and FBRM. Cryst Growth Des [Internet]. 2012;12(4):1792–807. Available from: http://pubs.acs.org/doi/10.1021/cg201269c.

Sang-Il Kwon J, Nayhouse M, Orkoulas G, Christofides PD. Crystal shape and size control using a plug flow crystallization configuration. Chem Eng Sci [Internet]. 2014;119:30–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0009250914004035.

Schöll J, Bonalumi D, Vicum L, Mazzotti M, Müller M. In situ monitoring and modeling of the solvent-mediated polymorphic transformation of -glutamic acid. Cryst Growth Des [Internet]. 2006;6(4):881–891

Siddique H, Brown CJ, Houson I, Florence AJ. Establishment of a continuous sonocrystallization process for lactose in an oscillatory baffled crystallizer. Org Process Res Dev [Internet]. 2015;19(12):1871–81. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00127.

Simon LL, Nagy ZK, Hungerbuhler K. Endoscopy-based in situ bulk video imaging of batch crystallization processes. Org Process Res Dev [Internet]. 2009;13(6):1254–61

Simone E, Saleemi AN, Tonnon N, Nagy ZK. Active polymorphic feedback control of crystallization processes using a combined raman and ATR-UV/Vis spectroscopy approach. Cryst Growth Des [Internet]. 2014;14(4):1839–50

Simone E, Szilagyi B, Nagy ZK. Systematic model identification and optimization-based active polymorphic control of crystallization processes. Chem Eng Sci [Internet]. 2017;174:374–86

Steendam RRE, ter Horst JH. Continuous total spontaneous resolution. Cryst Growth Des [Internet]. 2017;17(8):4428–36. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.7b00761.

Strickland-Constable RF. Kinetics and mechanism of crystallization from the fluid phase and of the condensation and evaporation of liquids [Internet]. London: Academic Press; 1968 [cited 2019 Mar 11]. Available from: http://cds.cern.ch/record/269001.

Strickland-Constable RF, Mason REA. Breeding of nuclei. Nature [Internet]. 1963 [cited 2019 Mar 11];197(4870):897–8. Available from: http://www.nature.com/articles/197897b0

Su Q, Benyahia B, Nagy ZK, Rielly CD. Mathematical modeling, design, and optimization of a multisegment multiaddition plug-flow crystallizer for antisolvent crystallizations. Org Process Res Dev [Internet]. 2015a;19(12):1859–70. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00110.

Su Q, Ganesh S, Moreno M, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. Comp Chem Eng 2019;125:216–231

Su Q, Nagy ZK, Rielly CD. Pharmaceutical crystallisation processes from batch to continuous operation using MSMPR stages: modelling, design, and control. Chem Eng Process Process Intensif [Internet]. 2015b;89:41–53. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0255270115000033.

Su Q, Rielly CD, Powell KA, Nagy ZK. Mathematical modelling and experimental validation of a novel periodic flow crystallization using MSMPR crystallizers. AIChE J [Internet]. 2017;63(4):1313–27. Available from: http://doi.wiley.com/10.1002/aic.15510.

Suresh P, Basu PK. Improving pharmaceutical product development and manufacturing: impact on cost of drug development and cost of goods sold of pharmaceuticals. J Pharm Innov [Internet]. 2008;3(3):175–87. Available from: http://link.springer.com/10.1007/s12247-008-9043-1.

Teoh SK, Rathi C, Sharratt P. Practical assessment methodology for converting fine chemicals processes from batch to continuous. Org Process Res Dev [Internet]. 2016;20(2):414–31. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.5b00001.

Ting HH, McCabe WL. Supersaturation and crystal formation in seeded solutions. Ind Eng Chem [Internet]. 1934 [cited 2019 Mar 3];26(11):1201–7. Available from: http://pubs.acs.org/doi/abs/10.1021/ie50299a017.

Tung H-H, Paul EL, Midler M, McCauley JA. Crystallization of organic compounds: an industrial perspective: Wiley; 2008.

U.S. Food and Drug Administration. Guidance for industry PAT: a framework for innovative pharmaceutical development, manufacturing, and quality assurance [Internet]. FDA official document. 2004. Available from: http://www.fda.gov/CDER/guidance/6419fnl.pdf.

U.S. Food and Drug Administration. Quality considerations for continuous manufacturing guidance for industry DRAFT GUIDANCE [Internet]. 2019. Available from: https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

Vartak S, Myerson AS. Continuous crystallization with impurity complexation and nanofiltration recycle. Org Process Res Dev [Internet]. 2017;21(2):253–61. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.6b00438.

Vetter T, Burcham CL, Doherty MF. Regions of attainable particle sizes in continuous and batch crystallization processes. Chem Eng Sci [Internet]. 2014 [cited 2017 Nov 14];106:167–80. Available from: https://ac.els-cdn.com/S0009250913007379/1-s2.0-S0009250913007379-main.pdf?_tid=f3858f58-c991-11e7-910e-00000aacb362&acdnat=1510701616_1c1d4d7662fe5426d25afafd44c8d330.

Volmer M. Kinetik der Phasenbildung. 1939.

Wang T, Lu H, Wang J, Xiao Y, Zhou Y, Bao Y, et al. Recent progress of continuous crystallization. J Ind Eng Chem [Internet]. 2017;54:14–29. Available from: https://doi.org/10.1016/j.jiec.2017.06.009.

Wey JS, Karpinski PH. Batch crystallization. In: Handbook of industrial crystallization: Elsevier Science & Technology; 2001. Cambridge University Press, Cambridge, UK

Wierzbowska B, Hutnik N, Piotrowski K, Matynia A. Continuous mass crystallization of vitamin C in scpl/scp(+)-Ascorbic acid−ethanol−water system: size-independent growth kinetic model approach. Cryst Growth Des [Internet]. 2011;11(5):1557–65. Available from: https://pubs.acs.org/doi/10.1021/cg101521k.

Wilcox WR. Movement of crystal inclusions in a centrifugal field. J Appl Phys. 1971;42(5):1823–7.

Wong SY, Tatusko AP, Trout BL, Myerson AS. Development of continuous crystallization processes using a single-stage mixed-suspension, mixed-product removal crystallizer with recycle. Cryst Growth Des [Internet]. 2012;12(11):5701–7. Available from: http://pubs.acs.org/doi/10.1021/cg301221q.

Wood B, Girard KP, Polster CS, Croker DM. Progress to date in the design and operation of continuous crystallization processes for pharmaceutical applications. Org Process Res Dev [Internet]. 2019 [cited 2019 Feb 5];acs.oprd.8b00319. Available from: https://pubs.acs.org/sharingguidelines.

Woodcock J. Modernizing pharmaceutical manufacturing – continuous manufacturing as a key enabler. In: MIT-CMAC international symposium on continuous manufacturing of pharmaceuticals. Cambridge, MA; 2014.

Yang Y, Nagy ZK. Advanced control approaches for combined cooling/antisolvent crystallization in continuous mixed suspension mixed product removal cascade crystallizers. Chem Eng Sci [Internet]. 2015a [cited 2017 Apr 11];127:362–73. Available from: http://www.sciencedirect.com/science/article/pii/S0009250915000895

Yang Y, Nagy ZK. Combined cooling and antisolvent crystallization in continuous mixed suspension, mixed product removal cascade crystallizers: steady-state and startup optimization. Ind Eng Chem Res [Internet]. 2015b [cited 2017 Apr 11];54(21):5673–82. Available from: http://pubs.acs.org/doi/abs/10.1021/ie5034254.

Yang Y, Song L, Gao T, Nagy ZK. Integrated upstream and downstream application of wet milling with continuous mixed suspension mixed product removal crystallization. Cryst Growth Des [Internet]. 2015a;15(12):5879–85. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b01290.

Yang Y, Song L, Nagy ZK. Automated direct nucleation control in continuous mixed suspension mixed product removal cooling crystallization. Cryst Growth Des [Internet]. 2015b [cited 2017 May 9];15(12):5839–48. Available from: http://pubs.acs.org.ezproxy.lib.purdue.edu/doi/pdfplus/10.1021/acs.cgd.5b01219.

Yang Y, Song L, Zhang Y, Nagy ZK. Application of wet milling-based automated direct nucleation control in continuous cooling crystallization processes. Ind Eng Chem Res [Internet]. 2016a;55(17):4987–96. Available from: http://pubs.acs.org/doi/10.1021/acs.iecr.5b04956.

Yang Y, Zhang C, Pal K, Koswara A, Quon J, McKeown R, et al. Application of ultra-performance liquid chromatography as an online process analytical technology tool in pharmaceutical crystallization. Cryst Growth Des [Internet]. 2016b;16(12):7074–82. Available from: http://pubs.acs.org/doi/abs/10.1021/acs.cgd.6b01302.

Yang X, Acevedo D, Mohammad A, Pavurala N, Wu H, Brayton AL, et al. Risk considerations on developing a continuous crystallization system for carbamazepine. Org Process Res Dev [Internet]. 2017;21(7):1021–33. Available from: http://pubs.acs.org/doi/10.1021/acs.oprd.7b00130.

Yang Y, Pal K, Koswara A, Sun Q, Zhang Y, Quon J, et al. Application of feedback control and in situ milling to improve particle size and shape in the crystallization of a slow growing needle-like active pharmaceutical ingredient. Int J Pharm [Internet]. 2017b [cited 2017 Nov 14];533(1):49–61. Available from: https://linkinghub.elsevier.com/retrieve/pii/S037851731730916X.

Yazdanpanah N, Ferguson ST, Myerson AS, Trout BL. Novel technique for filtration avoidance in continuous crystallization. Cryst Growth Des [Internet]. 2016;16(1):285–96. Available from: http://pubs.acs.org/doi/10.1021/acs.cgd.5b01231.

Zhang H, Quon J, Alvarez AJ, Evans J, Myerson AS, Trout B. Development of continuous antisolvent/cooling crystallization process using cascaded mixed suspension, mixed product

removal crystallizers. Org Process Res Dev [Internet]. 2012;16(5):915–24. Available from: http://pubs.acs.org/doi/10.1021/op2002886.

Zhang H, Lakerveld R, Heider PL, Tao M, Su M, Testa CJ, et al. Application of continuous crystallization in an integrated continuous pharmaceutical pilot plant. Cryst Growth Des [Internet]. 2014;14(5):2148–57. Available from: http://pubs.acs.org/doi/10.1021/cg401571h.

Zhang D, Xu S, Du S, Wang J, Gong J. Progress of pharmaceutical continuous crystallization. Engineering [Internet]. 2017;3(3):354–64. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2095809917304277.

Zhou Y, Srinivasan R, Lakshminarayanan S. Critical evaluation of image processing approaches for real-time crystal size measurements. Comp Chem Eng. 2009;33(5):1022–35

Zwietering TN. Suspending of solid particles in liquid by agitators. Chem Eng Sci [Internet]. 1958;8(3–4):244–53. Available from: https://linkinghub.elsevier.com/retrieve/pii/0009250958850319.

Chapter 6
Continuous Feeding-Blending in Pharmaceutical Continuous Manufacturing

Qinglin Su, Gintaras V. Reklaitis, and Zoltan K. Nagy

Abstract Pharmaceutical continuous manufacturing has steadily progressed from the proof of concept to the pilot and industrial production in the past two decades, some of which have recently been approved by the US Food and Drug Administration (FDA), resulting in a greater demand on experience in process design and operation in pharmaceutical continuous manufacturing. Unlike many of the individual unit operations that are themselves continuous operations, such as roller compaction, tableting, etc., and have been well studied previously, only the characterization of a continuous feeding-blending system will be discussed in detail in this chapter, which undergoes the most substantial change with a transition from batch to continuous manufacturing.

Keywords Feeding · Blending · Continuous manufacturing · Process design · Pharmaceutical

6.1 Introduction

In pharmaceutical industry, drug dosages are most commonly delivered in an oral solid tablet or capsule form. These solid dosages can contain multiple powder ingredients such as active pharmaceutical ingredients (API), excipient, lubricant, disintegrants, and absorption enhancers or agents that slow down and control absorption. Ingredients in a specific formulation are chosen depending on the material properties of the API and the desired nature of its release so that proper powder flow characteristics, API dissolution profile, and bioavailability can be ensured. It is not

Q. Su · G. V. Reklaitis ·
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Z. K. Nagy (✉)
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Department of Chemical Engineering, Loughborough University, Loughborough, UK
e-mail: zknagy@purdue.edu

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_6

uncommon for a drug formula to contain up to six or more ingredients. Therefore, the feeding and blending of these ingredients form the basis of pharmaceutical manufacturing processes.

A pharmaceutical manufacturing process usually proceeds with an initial powder feeding step, which defines the amount of each ingredient entering, simultaneously or sequentially, the manufacturing process, and thus determines the final content of a formulation. This is followed by a blending step to achieve a required homogeneity in powder. The feeding and blending are so important that the API is often a small part of a total blend (i.e., 10% by weight) and its content and uniformity are examined at the scale of scrutiny of in-process dosage unit, viz., a tablet or capsule as it is formed in the manufacturing process before it is coated or packaged. Inconsistent feeding and segregated ingredients can lead to quality failures such as out of specification dosage form assay and content uniformity. However, the FDA withdrew in August 2013 its draft guidance on blend uniformity (BU) – Guidance for Industry: Powder Blends and Finished Dosage Units-Stratified In-Process Dosage Unit Sampling and Assessment – due to concerns that the results of the test according to USP <905> uniformity of dosage units did not provide sufficient assurance that future samples taken from the batch would also comply with the same acceptance criteria. Further proposals were tabled by the International Society for Pharmaceutical Engineering. There is also an ongoing discussion about the correct methods of sampling materials to ensure that adequate blending is achieved. Nevertheless, the rigorousness and uncompromising in the determination of blend content uniformity also provide a glimpse of the current challenge in feeding-blending steps as well as research opportunity for developing modernized approaches to enforcing the Good Manufacturing Practice (GMP) requirements in this area (Cholayudth 2017).

At the meantime, the pharmaceutical industry has shown the trend in modernizing the industry with continuous manufacturing technologies in the last two decades (Lee et al. 2015). The conventional pharmaceutical batch manufacturing, by processing raw material stepwise through feeding, blending, and tableting with a batch size of around 100 kg or more, faces many challenges in scaling-up, manufacturing cost, product quality variance, etc., while the continuous manufacturing, defined as processing of raw materials without interruption and with continuity of production over a sustained period of time, offers many benefits in minimization of uncertainty in scale-up, reduced space and capital requirement, improved quality with consistent operations, etc. (Pernenkil 2008). These advantages in continuous manufacturing have also drawn the increasing interest in developing continuous feeding and blending unit operations. The pharmaceutical continuous manufacturing has since been identified as an emerging technology advanced by FDA and reinforced with its pharmaceutical quality-by-design (QbD) framework which required a deeper understanding of pharmaceutical manufacturing processes, for example, how ingredients are added and blended, how blending progresses through stages, and how to maintain consistent blending performance.

Fundamental research is still ongoing into continuous feeding-blending system, e.g., feeding mechanism (Weinekotter and Gericke 2000), powder flows in blender (Aissa et al. 2012), and addition order of individual components (Warman 2011); a

primary focus of this chapter is on the implementation of process system engineering tools in continuous feeding-blending system to gain better process understanding and achieve consistent blend quality.

6.2 Continuous Feeding-Blending System

Continuous manufacturing with powder-based unit operation is not uncommon in industries of mineral processing, polymers, food, etc.; however, it is relatively new to the pharmaceutical manufacturing industry, which is also subject to the unparalleled challenges in supervisory process control and real-time quality assurance. For example, the feeding and blending unit operations are the most common practice to feed different solid powder ingredients into a holdup, within which the individual ingredient is mixed uniformly with the others. Although continuous blending been practiced decades in other industries, only until last few years, continuous blenders that claim compliance to CFR (Code of Federal Regulations, US FDA) and CGMP (Current Good Manufacturing Practice, US FDA) have become available for pharmaceutical continuous manufacturing. It should be noted that the world's first commercial continuous blending process was successfully commissioned until 2015 by Vertex Pharmaceuticals for the cystic fibrosis drug (Rockoff 2015).

The challenges of continuous feeding-blending system in pharmaceutical manufacturing of oral solid drugs can be understood and evaluated as individual unit operation or within an integrated continuous manufacturing process.

In current pharmaceutical production, many feeders have already been operated in an automated way to deliver a weighted amount of ingredient with a defined material flow rate to be loaded into a manufacturing process. They are equally designed for continuous operation and make no difference in the basic principle for either batch or continuous manufacturing. However, the critical processing parameter (CPP) identified for feeding operation is very different. In a batch manufacturing, the mass of each ingredient needed for an entire batch is considered as the main CPP variable; a feeder has to deliver, reproducibly, the correct amount of each ingredient such that its mass composition, one of the critical quality attributes (CQA), agrees with the formulated specification as a whole. For example, an automatic feeder dispensing 10% of active pharmaceutical ingredient (API) to a 100 kg batch process simply has to deliver a total amount of 10 kg of API accumulatively from the start to the end of feeding, during which the feeding flow rate is flexible to accelerate or slow down as desired. While in a continuous manufacturing process, the feeding flow rate is identified instead as the main CPP variable which affects the CQA variable of ingredient mass fraction, as well as the total mass flow rate, at downstream unit operations. For a continuous manufacturing process producing 20 kg/hr, the automatic feeder is now required to deliver strictly 2 kg/hr of API constantly with adequate precision to meet requirements around API mass composition. The change of feeding CPP variable in continuous manufacturing has posted a great challenge to design a mechanically reliable feeding system. It is even worse if the API or excipient mass fraction is lower, e.g., a typical lubricant mass fraction of 0.2% equates to 0.04 kg/hr or rather 40 g/hr (Warman 2011).

However, a continuous blending unit operation differs from the batch one literally in their operation mode. A batch blending unit operation using a V shell or bin blender starts with the discharge of individual components, simultaneously or sequentially, into the blender, and the entire batch reaches a single end point when all the ingredients are blended homogeneously. While in a continuous blending unit operation, the blender or mixer most often features a tubular design, by which input ingredients are continuously being fed in one end and output homogeneous powder continuously withdrawn at the other end. The apparent advantage of continuous blending compared to the batch one is that the reduced size of equipment and amount of material handling empower the steady and instant uniformity of each component, a blending residence time of ~ 1 minute compared to a typical ~ 20 minutes batch blending time. However, the advantage of short residence time in turn conversely prevents the reach of mixing uniformity. For example, impeller rotation speed is one of the important CPP variable for continuous blending which affects the mixing uniformity of each ingredient, the CQA variable. Interestingly, intermedia impeller rotation speeds of a Gericke GCM 250 blender were observed for all blade configurations to achieve the best blending performance (Vanarase and Muzzio 2011), meaning there is an optimal but narrow operating window for continuous blending. This is because a higher impeller rotation speed tends to push forward the power ingredients more aggressively and shortens the residence time for powder to mix, on the one hand; it also has the detrimental effect to segregate the ingredients, on the other hand. A lower rotation speed may not sufficiently blend the powder despite extended residence time. Furthermore, it is not simply just a matter of residence time since a long blending time can also actually have a negative effect on the uniformity of the blended material, causing ingredients to segregate or demix, viz., overblending. The situation is getting worse when there are multiple ingredients involved in the continuous blender, and each ingredient possesses unique mixing performance in response to the impeller rotation speed. Though similar concerns are present in batch blending unit operation, the post-batch testing relaxes the quality control strategy compared to the real-time quality assurance in continuous manufacturing. Nevertheless, a convergence in continuous blender configuration and operation has not yet been achieved to ensure thorough mixing of powder ingredients, which has seen the recent surge in interest in industry and academia to research a variety of continuous blender design (Pernenkil and Cooney 2006; Oka and Muzzio 2012; Oka et al. 2017).

It is also important to understand the challenge of feeding-blending operation within an integrated pharmaceutical continuous manufacturing process. The feeding-blending subsystem is the first step in solid-dosage continuous manufacturing, which is also the last step at which major changes in ingredient compositions can be introduced, mitigated, or managed. Beyond this step, the formulation is largely fixed since final product or intermediate streams cannot be reworked (Pernenkil 2008). The feeding-blending operation is the major source of variance in the final product quality in continuous manufacturing. Therefore, the API mass composition and its uniformity in the powder must be strictly monitored and controlled within this subsystem, which are identified as the two CQA variables at this

step. For example, when a disturbance occurs in the feeding step, how will it affect the blending step and subsequent unit operations? Besides, one of the CPP variables that needs to be closely controlled is the total powder flow rate at the exit of blender. This variable affects the production rate of the entire processing line and is extremely important when one considers the limited holdup volume of the downstream equipment. Given the impact that the variability from these sources has on in-process materials, downstream processing, and drug product quality, it is clearly desirable to shift control efforts upstream to the feeding-blending system with a focus to reduce the reliance on end-product testing (Lee et al. 2015; FDA 2009). However, there are limited studies in the literature on the detailed development of systematic control strategy and risk assessment of the feeding-blending subsystem situated in a continuous manufacturing process (Pernenkil 2008; Zhao et al. 2013; Previdi et al. 2011; Rehrl et al. 2016).

Overall, because of the lack of experience in applying continuous feeding and blending to pharmaceutical manufacturing processes, it is thus important to review and summarize accordingly the previous research work and experience and analyze the design and monitoring of a resilient feeding-blending system with fault-tolerant control strategy and risk assessment for this processing subsystem in pharmaceutical continuous manufacturing (Su et al. 2017).

6.3 Equipment and Process Design

6.3.1 Continuous Feeder

Pharmaceutical continuous manufacturing requires a consistent, accurate, and reliable feed stream of formula ingredient to produce on-spec powder blend. Hence, the ability to feed powder consistently and continuously, specifically the most often cohesive API particles, is regarded as one of the critical requirements for continuous feeders.

A loss-in-weight (LIW) feeder is the most commonly used automatic feeding method in pharmaceutical industry, which continuously dispenses granular powder from an emptying hopper (Osorio et al. 2015). Generally, a LIW feeder consists of three parts, namely, (1) a volumetric feeder, (2) a weighting platform, and (3) a gravimetric controller. The volumetric feeder consists of a single screw or a pair of screws situated at the base of a feed hopper. The material enters the flight of the rotating screws due to gravity and is transported to the other end of the barrel which houses the screw. The whole volumetric feeder is placed on top of the weighting platform which measures the total mass of the feeder and the powders contained within. As powder is dispensed via the feeding device, the gravimetric controller acquires a signal from the weighting platform and derives the instantaneous feeding flow rate (Oka et al. 2017). The controller then compares the obtained feeding flow rate with a desired set point and adjusts the feed rate to achieve the set point by changing the speed of rotating screws.

Besides, many feeders can have agitators, microcentrifugal feeders, or even screens incorporated within their design. For example, agitators are typically in the form of rotating blade situated at the base of the hopper and help to ensure that material gets pushed into the feed screws. They also prevent ratholing of hoppers and prevent lump formation. It should be noted that the agitation results in mixing of powder along the depth of the hopper and it becomes critical for material traceability (Oka et al. 2017). A screen of various mesh sizes can be fitted at the exit of the screw barrel, which helps to de-lump the outgoing material and ensures a steadier flow.

In continuous manufacturing, a volumetric feeder needs to be refilled periodically due to a limited hold up in the hopper. During refilling, the fluctuations in hopper level have a direct impact on the feeding performance. One reason for this impact is the feeder switching to volumetric mode during refill. In this mode, the feeder is blind to changes in screw filling and changes in powder density. Another potential source of variation is material aeration during refill causing it to behave like a liquid and flood the feed screws (Oka et al. 2017). Hence, an accurate and consistent refill of these feeders is also equally important. However, the method of refill, refill control algorithm, reaction time of the refill device, and the size of the refill hopper have not been well investigated in designing a continuous manufacturing system (Nowak 2016; Engisch and Muzzio 2015).

Recently, different feeding mechanisms are discussed, and the importance of understanding the physical properties of the raw materials and its impact on the feeding process is reviewed (Blackshields and Crean 2018). It is advised that prior knowledge of materials will provide an initial indication of how the powders behave through processing and facilities in the selection of the most suitable feeder capacity, feeding mechanism, and screw type. Through careful evaluation of the feeding process, system layout, and material characteristics, not only a continuous feeding and refilling system can be optimized (Nowak 2016), but also the associated risk with its system disturbances and uncertainties to the downstream unit operations can be managed and controlled.

6.3.2 Continuous Blender

A tubular design of continuous blenders has become the most popular class for pharmaceutical blending applications to obtain the best homogenization of two or more particulate ingredients. There are two major technologies in which powders are distributed homogeneously in the entire powder bed, viz., (1) high shear, which includes high-speed mixing and paddle or baffle mixers, and (2) low shear, which includes the mixers that are operated at lower speeds or where the mixer outer shell revolves to create a particular movement of the powder inside (Velázquez et al. 2018; Bridgwater 2012).

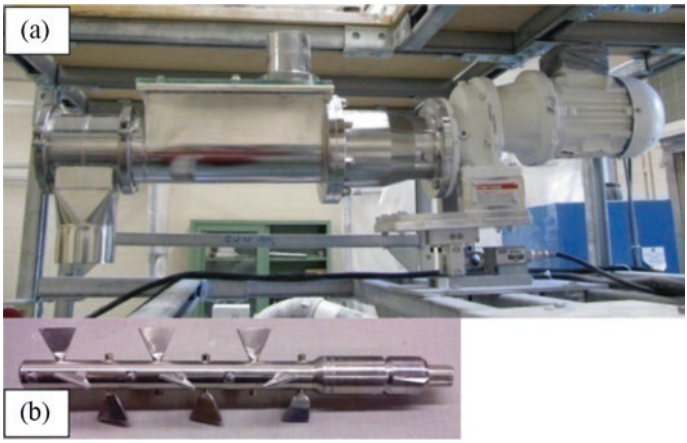


Fig. 6.1 A Gericke GCM 250 continuous blender (Reprint permission obtained from Vanarase and Muzzio 2011. Copyright © 2011 Elsevier)

For a high-shear blender, a cylindrical tubular section with a diameter ranging from 3 to 6 inches and axial length of 6–36 inches is usually observed (Oka et al. 2017). A motor-driven agitator is fitted along the axial centerline of this tubular section, with a number of blades distributed along its length, for example, a Gericke GCM 250 continuous blender is shown in Fig. 6.1. Besides changing the agitator rotation speed, blade configurations, such as type of blades, number of blades, and their orientation, can all be adjusted to control the mixing performance of the blender. While mixing through lifting and tumbling in the radial direction, as well as backward and forward in the radial direction by the blades inside the blender, the powder mixture is pushed outward with a net flow equal to the inlet feeding flow rate; as such, a steady-state blending with continuous flow is achieved.

As for a low-shear blender, e.g., a tumble mixer, when the outer shell rotates, the powder mixture inside is pulled up in the same direction by the friction between the particles and the wall and among particles, which is then pulled downward by the gravitational force (Velázquez et al. 2018). This results in various forms of powder flows including slipping, slumping, rolling, cascading, cataracting, and centrifuging (Aissa et al. 2012). Among them, the rolling is the one with highest interaction that forces the particles inside the mixer to change position between them. Under continuous operation mode, inlet powders enter axially from one tubular end, while the powder mixture leaves through the holes in the radial wall pushed out by the centrifugal force, as shown in Fig. 6.2. The rotation of the wall promotes the accumulation of powders, which forms a powder bed and undergoes the powder flows. Feeding flow rate, mixer speed and diameter, and exit area can be designed to obtain a desired flow regime, such as the rolling for specific materials.

There are also other classes of blenders that are under investigation or commercializing, for example, the GEA Buck Systems continuous dry blender for continuous convective powder mixing. Unlike the conventional horizontal design, the inlet powders are lifted upward through an inclined tubular section (Portillo et al. 2008).

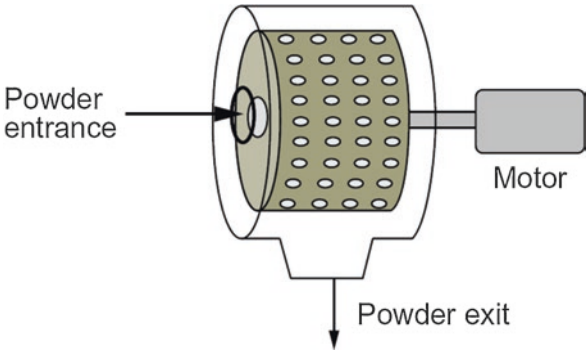


Fig. 6.2 Schematic of a continuous tumble blender (Reprint permission obtained from Velázquez et al. 2018. Copyright © 2018 Elsevier)

This is the first dedicated, purpose designed for the pharmaceutical industry, continuous blender (Warman 2011). Another example is a plant-scale Patterson Kelly Zig-Zag® blender that was investigated to understand its flow behavior and residence time distribution (RTD) and to examine the effect of operational variables on the blending performance (Pernenkil 2008). The blender consists of a drum section that houses an intensifier bar and V section which drives the powder from the inlet to the outlet. Interestingly, the external V section shell of the blender can rotate about its own axis. The angle of incline of the blender can also be changed to adjust the fill volume and mean residence time of the powder in the blender.

Overall, in all types of blender, the blending performance is influenced by the powder particle size and distribution, shape, morphology, and density of the active and excipient ingredients that are to be mixed. Not only the blender configuration parameters of type, size, and shape but also the operational parameters of mixing time, speed of rotation, flow rate, and volume of powder bed can affect the blending performance. All such parameters should be taken throughout the process design and operation to maintain consistently the required level of blended uniformity.

6.3.3 Process Design

The choice of continuous feeders and blenders for process design in continuous manufacturing is largely influenced by the material properties. Though more and more continuous feeder and blender modules are provided by equipment vendors, the integration of them to form a continuous manufacturing was also reported, such as the Portable, Continuous, Miniature, and Modular (PCMM) technology developed by Pfizer Inc. (Coupe 2015); there is no such simple guidance available that will determine the feeding accuracy and blend uniformity during the actual process. For example, as being said, a longer residence time in continuous blending does not simply mean a better blend uniformity; it may result in the segregation of certain ingredient from other materials. The residence time for mixing in blender should be adjusted accordingly to the formulation and blender characteristic, which again highlights the importance of a deeper understanding on the feeding-blending process design under the pharmaceutical QbD framework (FDA 2009).

First, variables that have a significant impact on process performance and product qualities should be identified at the process design step (Yu et al. 2014). For example, physical or chemical properties of raw material regarded as potential sources of change to the process performance and product quality goals should be identified and ensured to be within an appropriate limit, range, or distribution. These variables are known as the critical material attributes (CMAs), e.g., the ingredient particle size. Varying particle size distribution will result in the variation in the powder density thus the variation in the feeding flow rate or the settling of large particles in the powder bed within a blender. Analogously, the concept applies to identify the critical process parameters (CPPs). Characterizing the effect of potential CPPs on product or in-process material quality attributes throughout the process, or known as the critical quality attributes (CQAs, e.g., drug content uniformity as discussed in this chapter), is one of the goals of process design.

Second, the characterization of those quality critical variables will form the design space of the continuous manufacturing system or for the studied feeding and blending subsystem, within which consistent on-spec quality powder blend is achievable. In some cases, it may turn out that the design space is limited as subject to various constraints on those critical attributes or parameters, e.g., a feeding flow rate (CPP) for an active material that needs to be added in a small quantity but with high accuracy or the blended uniformity (CQA) of this small quantity is difficult to achieve. Then the process design should be re-evaluated by considering alternative approaches. For example, if an ingredient is a very minor component, there are three alternative approaches, viz., (1) by using a microfeeder, (2) by scaling up the design to a higher throughput, and (3) by pre-blending, discussed as follows.

There are some microfeeders that are available in the market. For instance, the Coperion K-Tron's family of twin- and single-screw microfeeders is designed to provide maximum accuracy at minimal feed rates, ensuring high-value ingredients to be fed accurately at flow rates as low as 32 g/hr with minimal residual material left in the feeder. Additionally, the 12 mm and 16 mm twin- and single-screw feeding modules are completely interchangeable and can feed a wide range of free-flowing to difficult powders (Coperion 2018). For example, a K-Tron MT12TM twin-screw microfeeder is capable of both batch and continuous operation.

Scaling up the overall process design to a higher throughput obviously will increase the feeding flow rate of a bottlenecked active ingredient; thus more feeder options are viable to the process design. However, this is not usually a desirable or practical solution. This is because in continuous manufacturing, in particular the end-to-end manufacturing, the overall throughput is notably constrained by the supply chain, e.g., in ingredient synthesis at upstream which involves complex reaction, crystallization, and filtration steps that are essentially more difficult to scale up. Additionally, the downstream tableting or capsuling process is also of a concern since the powder blend is processed on a basis of unit dosage, i.e., a few hundred milligrams.

While pre-blending a minor ingredient with a major ingredient into a well-mixed binary mixture via batch mixing prior to the continuous manufacturing may be a feasible solution (Oka et al. 2017), pre-blending may also be a pragmatic preprocessing step if the number of ingredients would require a large number of

individual feeders that would become cost prohibitive or overly complicated for implementation or to control the blend uniformity. Clustering more than six feeders around a single blender is difficult and also expensive. In addition, some ingredients, irrespective of the feed rates, are difficult to feed accurately using screw feeders. For example, silicon dioxide is observed to electrostatically adhere to the feeder parts leading to bearding. It is advisable for such ingredients to be introduced in the process by pre-blending them with major ingredients. Finally, if the continuous processing line is intended for multiple products, the proportions of major and minor ingredients will change from formulation to formulation; careful selection of feeders with combined pre-blending steps enables a fast and convenient product transition in feeding-blending system.

Accidently, it is also critical around the addition of some minor ingredients and what type of effect it is trying to achieve. For example, should the lubricant be distributed within the blend but remain as a discrete powder or should it be smeared over particles of the other components (Warman 2011)? More fundamental research and understanding are needed actually in feeding-blending system design in continuous manufacturing.

6.4 Process Monitoring

6.4.1 PAT Tools for Sensor Measurement

During continuous manufacturing, the previously identified quality critical variables and parameters should be closely monitored in a statistically univariate or multivariate way by control chart (Mason and Young 2002), which is already a mature and routine procedure in industrial applications. Herein, common process analytical technology (PAT) tools for real-time measurement of CQAs in feeding and blending system are briefly introduced, followed by highlighting the importance of data reconciliation for process monitoring in pharmaceutical continuous manufacturing.

Traditionally in batch manufacturing, destructive analytical methods, such as dissolution followed by HPLC or UV, to monitor and control the blending process can be used. However, this requires sampling and analyzing powder mixture, which can be delayed for real-time release in continuous manufacturing. On the other hand, though withdrawn, the FDA used to propose that collected samples from various identified locations within a lot or batch should all be within 10.0% of the mean to meet adequate level of blending. Analysis is only necessary when the active ingredient is 50 mg or greater and when the dosage contains more than 50% active ingredient, requiring the blend complies with the limit of 85–115% or 75–125% of the target strength of individual units (LFA 2018).

However, a new guideline adaptive to the emerging continuous manufacturing by the FDA or the ISPE is not yet deliverable. One of the reasons is that with various

innovative PAT tools such quality analysis or inspection is now more readily taken out in situ and in real time, whereas a common sense on the quality criterion is difficult due to the fact that each has different levels of scrutiny scale, accuracy, precision, and sampling interval. The specifications, such as accuracy, precision, response time, operating range, resolution, sensitivity, drift, etc., are the important decisions that need to be made on the selection of these PAT tools (Velázquez et al. 2018). For example, PAT tools for real-time monitoring of critical process and quality variables have been widely reported for the feeding-blending system. Among them, NIR spectrometers are mainly applied for API composition, content uniformity (Vanarase et al. 2010), and powder bulk density measurement, load cells (built within feeders) for ingredient feeding flow rates from each feeder, and an X-ray sensor for the total powder mixture flow rate (Ganesh et al. 2017). Besides, there are some recent innovations worth mentioning here.

For example, in previous studies, only a single NIR probe was implemented at a fixed location by the exit of a continuous blender. Uniformity of the consecutive unit dosage measurement is generated at a fixed time interval by calculating the relative standard deviation of the measured ingredient composition within a time window (Zhao et al. 2013). However, the spatial unit dose to unit dose variability can be missed out. This is even more important when the scrutiny scale of a unit dosage around 200–300 mg per second is compared to a production flow rate of 30 kg/hr, viz., 8.3 g/sec, meaning about 2.0% of the material is sampled (Vanarase et al. 2010). Though larger portion of material, as people may think, is sampled compared to a batch manufacturing, the principle randomness in spatial locations is not considered in fact, whose effect is not yet rigorously studied.

The recent innovation in multipoint NIR system has shown the potential in addressing this concern. For example, an Innopharma Multieye$_2$ is a multiple near-infrared spectrometer designed for real-time in-line process monitoring. A single sensor with up to four discrete channels allows measurements from four probes located within a process and can monitor the blend uniformity without the channel-to-channel variation commonly found with multiple single-point systems. An application of similar design with up to six different position of a batch blender has been reported previously (Scheibelhofer et al. 2013) to detect the end point; however limited work is reported in continuous blender.

A hyperspectral imaging (HSI) technology was also proposed to study the behavior of solid particles in various unit processing steps as well as during multistep continuous processes. Hyperspectral imaging, or chemical imaging (CI), is the combination of spectroscopy and digital imaging. A hyperspectral image contains many spectra, one for each individual point on the sample's surface. The image can contain information about the spatial distribution of the materials within the sample. Herein, a hyperspectral refers to the full 1000–2500 nm range as short wave infrared (SWIR). This nondestructive method generates thousands of spectra per second and is able to provide more compositional information than conventional methods for blender monitoring in real time to analyze the average composition and the distribution of ingredients (Kemeny and Stuessy 2012).

6.4.2 *Data Reconciliation for Measurement Correction*

The previous section mentions a few of the recent developments of PAT sensors in measuring the critical quality variables in feeding and blending system. On the other hand, although the management and integration of these PAT measurement data for process monitoring have been receiving extensive interests (Markl et al. 2013; Cao et al. 2018), the measurement data reliability and the transition from a data-rich to information- and knowledge-rich manufacturing have not attracted enough attention thus far (Ierapetritou et al. 2016). For example, these in situ PAT sensors in the powder-based pharmaceutical manufacturing often suffer from random errors due to process disturbance or gross errors due to changes in material properties and environmental factors. While the current practice of multivariate statistical process control (MSPC) monitors the variation in the measurement data by projecting them into a small number of principal components and testing them against statistical control limits prespecified under nominal operation, e.g., Hoteling's T^2 and square prediction error (SPE), in order to detect measurement gross error or process failure, this data-driven approach does not dig out the true variation of the product quality attributes under data imperfection, which are of great concerns to support active product quality control and real-time process decision-making in pharmaceutical continuous manufacturing under the FDA QbD guidance (Moreno et al. 2017; Yu et al. 2014). Furthermore, there are some quality critical variables that cannot be measured directly, e.g., the mixing residence time or the weight of powder bed in a blender, and thus cannot be captured by MSPC methods. Besides, an adequate database of nominal operation for MSPC development is not always available, especially not during the early stages of product or process development. Hence, research efforts in tackling with data impreciseness as well as uncertainty in process monitoring based on process knowledge are merited in pharmaceutical continuous manufacturing.

Data reconciliation (DR) is a mathematical tool to correct imperfect measurement data to fall within the process system engineering's best knowledge of the process, i.e., mass and energy balance, process variance and dynamics, or correlation between variables. The original idea of data reconciliation in correcting the measurement data with normally distributed random errors with zero mean in 1960s has since been rapidly extended to the measurement gross error identification and elimination, unmeasured state variable estimation, model parameter estimation, etc., demonstrated in Fig. 6.3, which have been successfully implemented in various industries in its long history (Câmara et al. 2017). Recent industrial applications have been found, e.g., in a natural gas processing plant, wherein the mass flows of output streams and raw feed were reconciled to satisfy mass balance equations with an aim to check the presence of gross errors in the measured data (Rafiee and Behrouzshad 2016). A corrected expansion curve with reasonable enthalpy-entropy relationships and better estimates of isentropic efficiencies was obtained with data reconciliation implemented in a 1000 MW steam turbine power plant (Guo et al. 2016). More recently, the data reconciliation was also employed to improve the reli-

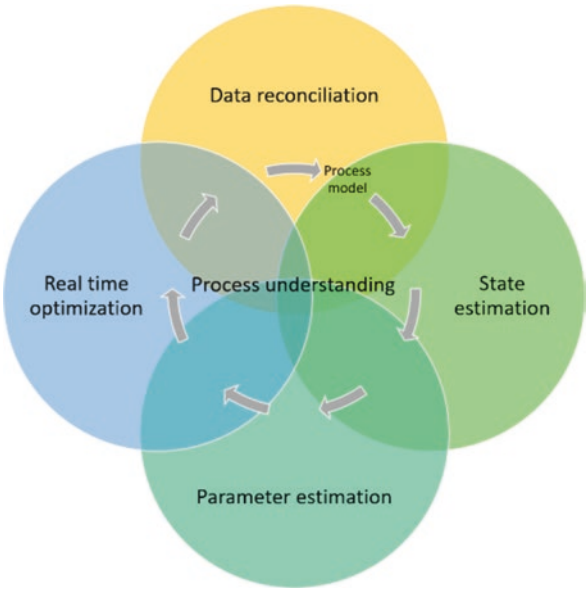


Fig. 6.3 Scope and connections of data reconciliation research. (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2019 Elsevier)

ability and accuracy of measured data for a direct air-cooling condenser of thermal power plant based on the relationship between back pressure of the steam turbine and the condenser-related variables (Li et al. 2018).

Hence, it is interesting to introduce data reconciliation to the emerging pharmaceutical continuous manufacturing in compliance with the QbD guidance. The principle of DR shares with the QbD guidelines in highlighting the importance of process understanding, while addressing the imperfection of PAT measurement data, for example, the design space in QbD can be described by the inequality constraints in DR (Guo et al. 2016), the process understanding developed in QbD can be mathematically modeled and cast into nonlinear equality constraints of DR (Valdetaro and Schirru 2011), or the variation in the process coefficients/material parameters can also be estimated in DR (Weiss et al. 1996). Furthermore, given the significant differences which normally occur between CQA and CPP measurement uncertainties, the implementation of DR within QbD is also of practical importance for active process control of CQAs by adjusting CPPs. For example, most CQA variables are now measured using in situ spectroscopy probes, e.g., API composition by a NIR probe. Such measurements are often subject to either inherent error in chemometric model calibration or the extra variation/drift due to material property change (particle size, bulk density, etc.), probe fouling, environmental humidity and temperature changes, etc. (Chen et al. 2011). On the other hand, the CPP variables are commonly and directly measured using reliable mechanical or electrical sensors, e.g., the dispensed powder mass by a load cell, which tend to have less measurement uncertainties. It is worth mentioning that the robust design of traditional manufacturing equipment (automatic feeder, rotation motor, etc.) which provides reliable measurement of CPPs (e.g., blender rotation speed) with minimum variations has

allowed pharmaceutical manufacturing to continue to operate in the past century in batch mode, while handling CQA variations via end-of-line statistical quality control (SQC) methods. The challenge in pharmaceutical continuous manufacturing nowadays is to effectively integrate those noisier and possibly biased CQA measurements into the process monitoring system so as to effectively supervise the control of CPPs, while minimizing the need for end-of-line SQC.

More recently, pioneering works on implementing data reconciliation to continuous feeding and blending system have been reported. For example, based on a steady-state mass balance model of a feeding-blending system, gross errors with an NIR probe (CDI Inc.) located at the exit of a Gericke GCM 250 continuous blender for API composition measurement were detected by a steady-state data reconciliation strategy (Moreno et al. 2017). The potential application of a real-time optimization strategy to implement a dynamic joint data reconciliation and state estimation to a feeding-blending system was also reported (Liu et al. 2018). The authors employed the idea of tank in series to model the system dynamics of a continuous blender, with which the time response of API composition can be reconciled with the uncertain NIR composition measurement. More importantly, the mass of powder accumulated in each simulated tank, equivalent to the mass distribution of powder bed along the horizontal blender axial, was estimated. In such a way, the mixing residence time, a CPP variable to avoid overblending, can be monitored. Another application of a joint state estimation with data reconciliation was also reported by combining a Kalman state estimator with a Welsch robust estimator for a feeding and blending system; the authors assessed the risk impact of gross errors in API composition measurement on the process control system (Su et al. 2017).

To sum up, process monitoring in pharmaceutical continuous manufacturing is not just about the development and application of PAT sensors to measure the CQA variables but also the integration of these data to provide a mean to gain a nonbiased understanding, monitoring, and controlling of the process with a goal to quality production with real-time release strategy, which will be further discussed in the following sections for continuous feeding and blending system.

6.5 Process Modeling

Mathematical modeling is an important process system engineering tool to acquire and verify a scientific process understanding on the design space of critical process and quality variables, specifically on variations transduced from upstream critical material attributes (CMAs) to downstream critical quality attributes (CQAs), upon which optimal process design and operation can be sought. Depending on application purposes, mathematical models can be presented with different levels of detail and at different scales. For example, a discrete element method (DEM), computing the motion and effect of a large number of small particles, such as granular flows and powder mechanics, is widely accepted as an effective method of addressing engineering problems in individual equipment design, e.g., feeder screw or blender

agitator design (Sen et al. 2013), while less detailed models based on chemical engineering principles of conservation laws in mass, energy, and momentum, or as simple as a transfer function model can be developed for each unit operation and be integrated together in a plant-wide flowsheet model to address the issues of process dynamics. The following sections will discuss their applications in continuous feeding and blending system, with a focus on flowsheet modeling tools for process control purpose.

6.5.1 Unit Operation Models

A detailed review on the computational tools and mathematical modeling approaches for unit operation in continuous manufacturing of solid-based pharmaceutical dosage form can be found in recent reviews and book sections (Rogers et al. 2013; Yoon et al. 2018; Ierapetritou et al. 2017), including DEM simulation, population balance models, and reduced order models for feeders and blenders.

Except those mechanical models with computationally demanding DEM simulation, unit operations of feeding and blending are more often modeled or characterized by their process parameters and material properties for process operation purpose, such as those CMAs and CPPs. For example, flow and discharge behavior of pharmaceutical powders from several different hopper geometries has been studied experimentally from which a quantitative flow index was shown to correlate with hopper flow behavior based on data from a gravitational displacement rheometer (Faqih et al. 2007). Feeding models with process parameters and material properties of holdup, feed factor, and screw speed was also proposed to calculate the mass flow rate, in which the powder bulk density was used to determine the feed factor (Engisch and Muzzio 2012). In terms of continuous blender, it was reported that the characterization and quantification of the stirring action can be empirically related to the flow rate and the rotational speed of a continuous blender. By doing so, it can systematically model the effects of operating conditions (such as rotational speed and processing angle) and design parameters (such as blade design) on the mixing efficiently (Marikh et al. 2005).

Just to mention a few, it is noted that many unit operation models with detailed characterization on the feeding and blending steps, individually, have been reported; however, the integration of them in a flowsheet model is rarely reported.

6.5.2 Flowsheet Models

A flowsheet model can contain a number of independent unit operation models that connect each other with material streams to represent a process flow diagram. Material properties, operating parameters, and control strategies defined in a process design can be simulated with a flowsheet model. In this way, the impact of

process start-up, shutdown, unexpected intermittent disturbances, and proposed control strategies can be investigated in silico, which informs decisions made in a process design or in a real plant. There are several commercial software platforms supporting the flowsheet modeling, mainly Simulink (MathWorks) (Su et al. 2017), Jacobian (RES) (Lakerveld et al. 2013), Aspen Plus (Aspen Technology), and gPROMS (Process Systems Enterprise, PSE Ltd) (Singh et al. 2014). PSE Ltd also offers gPROMS Formulated Products, a package designed specifically for pharmaceutical manufacturing processes (Yoon et al. 2018). Usually, these platforms come with a library that provides basic modular functional blocks.

The flowsheet modeling plays an important role in efficient plant-wide control strategy development by accelerating process design optimization and achieving desired control objectives. For example, global sensitivity analysis and system identification based on flowsheet simulation can help to identify and address the potential challenges or risks in process control design, i.e., decentralization, pairing, stability, resilience, etc. For example, a flowsheet model was developed for a feeding-blending system in Simulink as shown in Fig. 6.4, where two loss-in-weight feeders and a blender were modeled (Su et al. 2017). The feeder was modeled using a first-order plus time delay (FOPTD) transfer function and was controlled using proportional-integral-derivative (PID) controllers to adjust the screw rotation speed. The continuous blender was modeled with a two-dimensional compartmental model, with fluxes estimated for each component in forward, backward, and radial directions. The aim of the flowsheet modeling was to investigate the process control design and risk analysis in continuous feeding and blending system, by which system identification with state-space models; control design and analysis metrics;

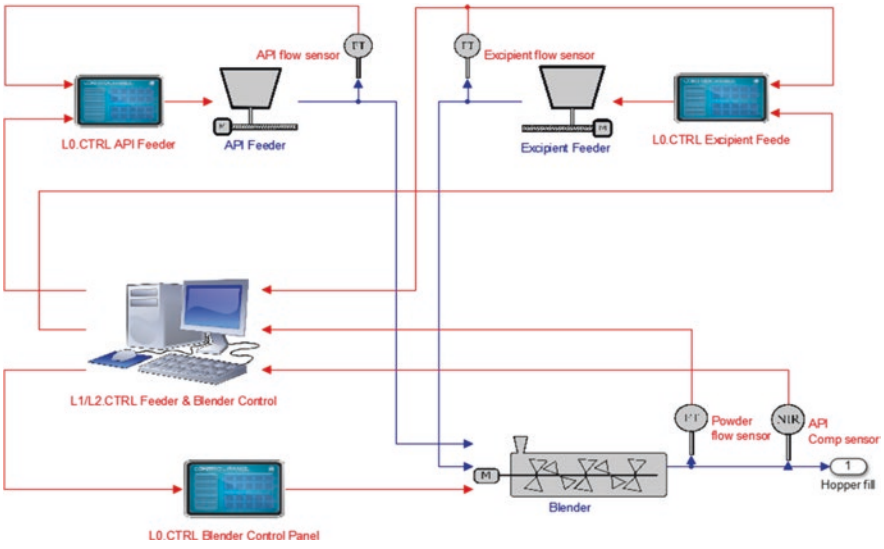


Fig. 6.4 A Simulink flowsheet model of the feeding-blending system. (Reprint permission obtained from Su et al. 2017. Copyright © 2017 Springer Nature)

hierarchical control structures; risk mapping, assessment, and planning (Risk MAP) strategies; and control performance indicators were studied.

A flowsheet model based on gPROMS platform for a continuous direct compaction process was also reported to help on-site design and implementation of control system (Singh et al. 2014). In this flowsheet model, the feeding unit operation was also considered as a simple transfer function with parameters of process gain, time constant, and time delay estimated from observed experimental flow rates, while a much more detailed multidimensional population balance model was constructed to model the blending process that accounts for several solid ingredients and two external coordinates (axial and transverse direction in the blender) and one internal coordinate (size distribution due to segregation) (Boukouvala et al. 2012). The flowsheet model can be integrated with DeltaV using the gORUN feature of gPROMS, MS Excel, and I/GEAR connecting software, as shown in Fig. 6.5. The signal of the control variable generated from the model simulation was sent to DeltaV and acts as the input to the controller that generated the actuator. The controller output (actuator) was sent to the process model and acts as the input to calculate the new signal of the control variable. Similar integration can be done with MATLAB OPC toolbox (Singh et al. 2013).

More recently, a flowsheet model for a pharmaceutical continuous direct compaction process based on gPROMS was also presented (García-Muñoz et al. 2018). The flowsheet model was used to produce residence time distributions (RTDs) at different process conditions and a graphical representation of the allowable range of variability or disturbances in the screw feeders that can be mitigated by the process. The dynamics of composition and mixing were thus main focus of this flowsheet model. Hence, a perfect feeding was assumed in this model, while a horizontal blender was described as a dispersed plug flow reactor. Note that the powder stream was treated as a fluid therein and the dispersion term was assumed to be the same

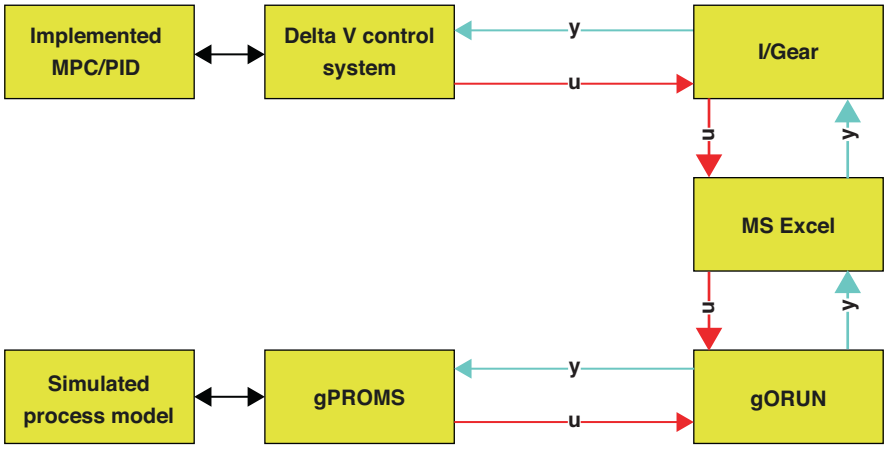


Fig. 6.5 Integration of flowsheet model with DeltaV control platform (Reprint permission obtained from Singh et al. 2014. Copyright © 2014 Elsevier)

for all components. The characterization of the RTDs with the process configuration, operational parameters, and feeding disturbances helped in evaluating the material traceability and proposing off-spec product diverting strategy (Oka et al. 2017).

Overall, due to the scale of flowsheet models in understanding the plant-wide process dynamics and entire residence time distribution, contributions from detailed unit models of feeders and blenders are usually neglected, for example, a perfect feeding is often assumed. This may be due to the assumptions that (1) the magnitude of dynamics or residence time in feeding-blending system is negligible compared to the entire system (2) or those dynamics can be represented as random variation or time delay. Though most often these assumptions hold for traditional plant-wide process control design purpose, it should be cautioned, before considerable experimental verification, to implement any control design or draw any conclusion based on an oversimplified flowsheet model.

Furthermore, for product quality control purpose in a highly regulated pharmaceutical manufacturing process, continuous improvement in individual unit models with more details in low-level control strategy, equipment parameters, and material properties is essential in gaining in-depth process knowledge within the design space and eventually achieving a quality-by-design paradigm. It is found that rare flowsheet models, to the best knowledge of the authors, have considered the variations in powder properties such as the CMAs of bulk density or particle size distribution and their effect on feeding accuracy or blending uniformity in a flowsheet modeling. Without such, the variation in the downstream operation cannot be well understood or diagnosed. For example, when a nonuniformly blended powder mixture is entering a downstream rotary tablet press, one would notice an inconsistence in tablet weight and punch pressure, which are both critical process parameters concerning the drug dosage and dissolution profile (LFA 2018; Su et al. 2018a).

To sum up, with tremendous advancement in capacities of high-performance computing, parallel computing, and cloud computing, more constituent modular with fundamental powder mechanics are necessary and possible to be integrated within the flowsheet modeling to understand not only the plant-scale dynamics but also the propagation of variations in critical quality attributes.

6.6 Process Control

There are two main objectives in process control development for a continuous feeding and blending system in pharmaceutical continuous manufacturing, viz., process automation and product quality control. Achieving process automation is important to start up a continuous manufacturing system to rapidly reach the steady-state set points for those CQA and CPP variables of ingredient mass fraction, mixing uniformity, total powder flow rate, etc., in a feeding-blending subsystem by automatically adjusting the feeding screw speed of each feeder and blender agitator rotation speed. In the aspect of product quality control, the control system monitors these CQAs and CPPs of the output powder mixture in real time. In case of any

offsets from the set point due to process disturbance or raw material property change, the control system will accordingly fine-tune actuators of feeding screw speed or blending agitator rotation speed to assure that CQAs consistently conform to the established criteria. This active role of process control in continuous manufacturing is the most adaptive and can enable real-time release testing and provides an increased level of quality assurance compared to traditional ways of end-product testing (Yu et al. 2014; Lee et al. 2015).

Process control system has been a routine in many other continuous manufacturing processes, and there are also many control system development frameworks that take on a step-by-step guidance in deploying process automation and control units. However, the reliance of product quality on control system has never been so difficult and paralleled to the pharmaceutical continuous manufacturing, for example, the challenges mentioned in Sect. 6.2 such as small quantities of active ingredients to be added with accuracy, limited buffer volume during processing, etc. Hence, only basic control modules provided by equipment vendors to adjust most CPPs of feed rate and rotation speed, termed as the level 0 control herein, are now known to be implemented in commercial continuous feeding and blending system, despite the encouraging progresses that are made in developing advanced higher-level control systems in most research or pilot facilities. More research work regarding the resilience, fault tolerance, or risk assessment is indeed important to make the active process control an important ingredient in the QbD framework.

The following sections will discuss the control algorithms and development frameworks that are available for the feeding-blending system in the literature.

6.6.1 PID Control

A proportional-integral-derivative (PID) controller continuously calculates an error value as the difference between a desired set point (SP) and a measured process variable (PV) and applies a correction based on proportional, integral, and derivative terms of the error. It supports the control of a single measured process variable by adjusting a single manipulated variable in a closed loop, viz., the single-input and single-output (SISO) control loop. Besides the basic level 0 control provided by equipment vendors of feeder and blender, PID loops are the most common control algorithm that are implemented to achieve process automation and product quality control in a feeding-blending system. These PID control loops can be designed to supervise the basic level 0 control; hence, their implementation is regarded as the level 1 control and is usually based on a distributed control system (DCS), such as the Emerson DeltaV control platform.

Generally speaking, there are two specific control strategies for the single-input and single-output PID control loop design, viz., (1) adjusting the flow rate of the API until the desired concentration is achieved or (2) adjusting the flow rate of the API as a ratio of the flow rate of a main excipient. The first one is known as the most popular feedback PID control, while the latter can be regarded as a complementary feedforward control by a ratio controller (Su et al. 2017).

For example, a PID controller was experimentally performed as an initial attempt to real-time control the homogeneity of the powder mixture produced in a pilot plant with two loss-in-weight feeders and a Gericke GCM 500 continuous blender (Zhao et al. 2013). The rotation speed of the stirrer, identified as an important deciding factor toward the mixer's efficiency, was used as the manipulated variable. Closed-loop control based on either the mean API concentration or its relative standard deviation was evaluated for continuous blending subjected to step changes in feed rates of the mixer. The PID tuning was found to be a critical step to obtain a good control performance, and it was mainly due to challenge in the process or measurement fluctuations.

In a recent work on the control of a similar feeding-blending system, where a tumble mixer was employed instead, three PID controls were reported to control the API concentration, API relative standard deviation, and powder flow rate, respectively (Velázquez et al. 2018). The difficulty in controlling the blending process at low API dosages or concentrations was pointed out; see the fluctuations of API (Naproxen) concentration in Fig. 6.6. Two major problems were the accuracy of dispensing low flow rates of the API which tend to have low flowability and the development of the NIR calibration model for low content (Vanarase et al. 2010).

Similar implementations of level 1 PID controls for feeding and blending system are also found in recent work based on continuous manufacturing pilot plants at Engineering Research Center for Structured Organic Particulate Systems (ERC-SOPS) of Rutgers University and Purdue University (Ierapetritou et al. 2016).

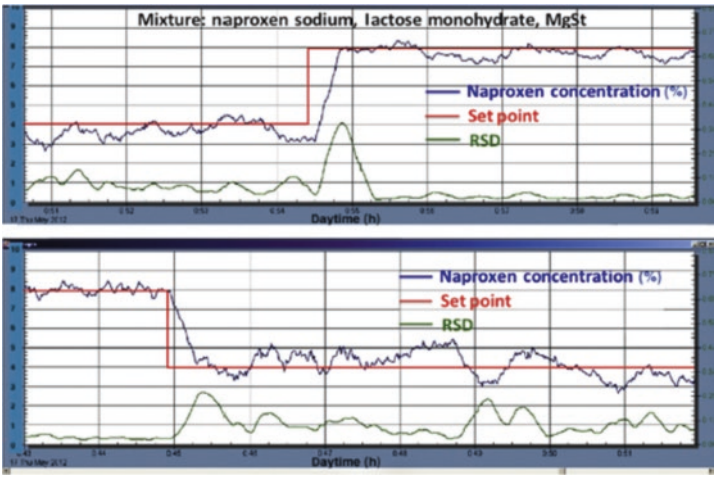


Fig. 6.6 Closed-loop dynamics of a continuous tumble mixer for API centration and RSD (Reprint permission obtained from Velázquez et al. 2018. Copyright © 2018 Elsevier)

6.6.2 Advanced Model-Based Control

The challenges in continuous feeding and blending system for pharmaceutical manufacturing also see the increasing demands for advanced model-based control strategies, for example, the model predictive control (MPC). MPC is used to control a process while satisfying a set of constraints and has been in wide applications in chemical plants and oil refineries since 1980s. MPC relies on dynamic models of the process, most often linear empirical models obtained by system identification, which can be conveniently done in most commercial DCS platforms, e.g., the Emerson DeltaV MPC toolbox. As such, MPC has the ability to anticipate future events by running the model and can take control actions accordingly, while this predictive ability is not seen in a PID controller. More than that, based on a process model, the MPC controller is more promising in controlling the process as a whole involving several CQAs and CPPs as in a feeding and blending system, viz., the multiple input and multiple output (MIMO) control, compared to the single input and single output of a PID controller. Due to the superiority of the model-based control algorithm and its reliance on process model, its implementation and maintenance are referred as the level 2 control.

The development and implementation of MPC controller to a feeding-blending system were reported in a continuous pharmaceutical tablet manufacturing pilot plant (Singh et al. 2014), as shown in Fig. 6.7. A filtered API composition signal measured by NIR was used as the input for a model predictive controller to manipulate the feeding ratio of API feed rate vs. total powder flow rate. Though the studied MPC control was demonstrated with better control performance than the PID control, e.g., shortening the transition time in reaching a set point, fluctuations even in the filtered API composition signal were still observed.

Relating to this, recent research work has seen the improvement in combining other model-based methodologies, such as data reconciliation and state estimation to improve measurement accuracy of API composition, with MPC control in feeding and blending system. Other than the linear empirical models, the state-space model and nonlinear dynamic process model are adopted in these model-based control algorithms.

For example, a state-space model describing the feeding-blending system at Purdue University was recently reported, which involves two loss-in-weight feeders and a Gericke GCM 250 continuous blender (Su et al. 2017). A robust Kalman state estimator was then designed using the identified state-space model of the feeding and blending system to estimate state variables ($\mathbf{x}$) from current process output measurement ($\mathbf{y}$) of API composition and total powder flow rate, as well as the process input ($\mathbf{u}$) of feeding flow rates, blender rotation speed, etc., as shown below:

$$\mathbf{x}[k \mid k] = \mathbf{x}[k \mid k-1] + \mathbf{M}\left(\mathbf{y}[k] - \mathbf{C}\mathbf{x}[k \mid k-1] - \mathbf{D}\mathbf{u}[k] - \mathbf{e}[k]\right)\mathbf{R} \tag{6.1}$$

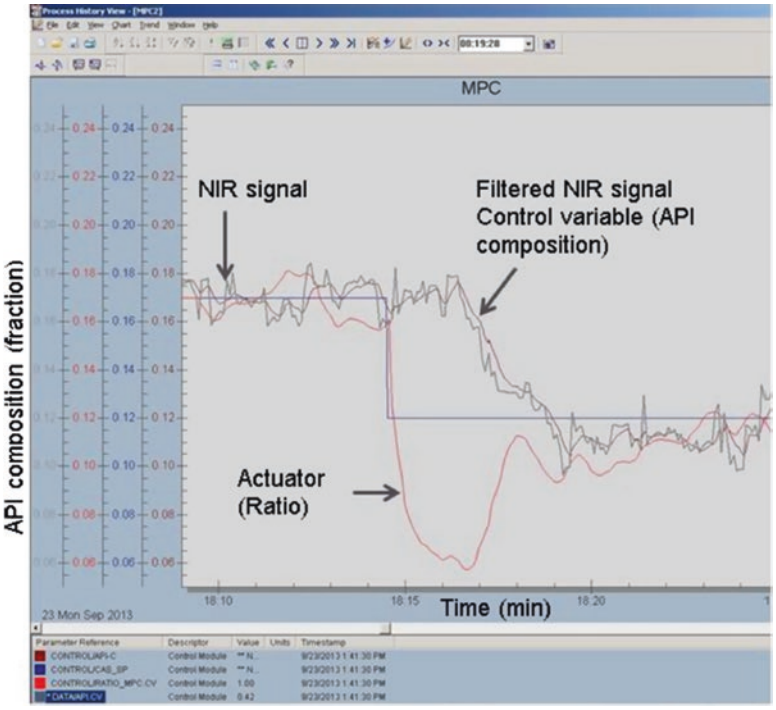


Fig. 6.7 Experimental closed-loop response for API composition control by a MPC controller (Reprint permission obtained from Singh et al. 2014. Copyright © 2014 Elsevier)

$$\mathbf{R} = \mathbf{diag}\left(\exp\left(-\left(\frac{e_i}{c\sigma_i}\right)^2\right)\right) \tag{6.2}$$

where $\mathbf{x}[k|\, k-1]$ is the state variable for kth sampling time that is estimated at k-1th sampling time; $\mathbf{y}[k]$, $\mathbf{e}[k]$**,** and $\mathbf{u}[k]$ are the current process output, bias, and input measurement at kth sampling time; and $\mathbf{M}$ is the Kalman innovation gain (Franklin et al. 1990). $\mathbf{R}$ is the data reconciliation matrix, using the Welsch robust estimator, to reject the gross error in measurements when model-plant mismatch $\mathbf{e}_i$ deviates from its nominal standard deviation $\boldsymbol{\sigma_i}$; note that $\boldsymbol{c}$ is a tuning parameter. The integration of robust Kalman state estimator with a state-space model-based MPC design at level 2 was found to be able to reject the measurement gross errors in the API composition measurement and thus avoid major fluctuations in the control system, as shown in Fig. 6.8. When measurement gross error in API composition occurred from 0–200 seconds, the measured variable was drifted from 10% to 8%. However, the robust Kalman estimator identified and eliminated this gross error and predicted a steady API composition measurement, by which the process variables can remain at stead state and under control.

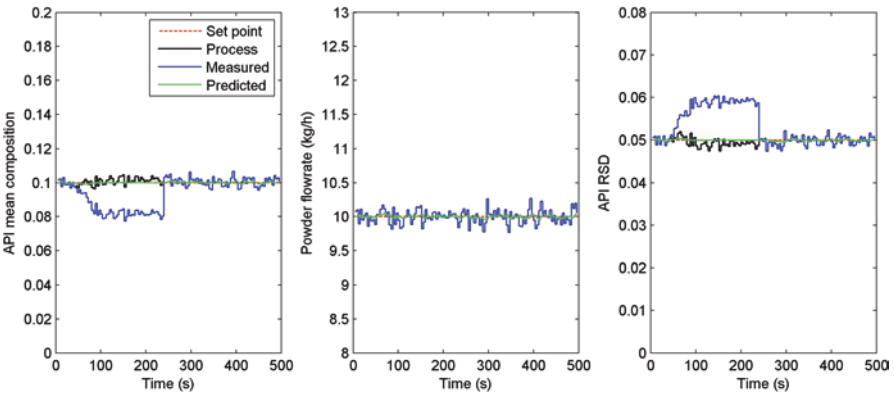


Fig. 6.8 Control performance of MPC with a robust Kalman state estimator. (Reprint permission obtained from Su et al. 2017. Copyright © 2017 Springer Nature)

More recently, an integrated moving horizon robust state estimation and nonlinear MPC control of the same feeding-blending system were also proposed (Liu et al. 2018). To improve the accuracy and robustness of state estimation, robust estimators within the standard moving horizon estimation (MHE) skeleton were proposed, leading to an extended MHE framework. Numerical simulations showed that the studied approach was robust to gross errors and can provide reliable state estimates when measurements were contaminated with outlines and drifts, as shown in Fig. 6.9 for the total powder mass flow rate at the exit of the Gericke GCM 250 continuous blender, where most gross errors were eliminated.

Hence, the promising potential of implementation of model-based control strategies in an integrated form of data reconciliation, state estimation, and process control demands more research work in systematic control design framework and experimental verification and validation in continuous feeding-blending system.

6.6.3 Framework for Control Design

Systematic approaches to the design and implementation of control systems for continuous pharmaceutical manufacturing processes have been proposed and progressively improved in the past few years by the Engineering Research Center for Structured Organic Particulate Systems (ERC-SOPS) (Singh et al. 2014; Singh et al. 2009). A range of control techniques under the frameworks, ranging from simple proportional-integral-derivative (PID) controllers to advanced model-based control and real-time optimization (Singh et al. 2015a, b; Ramachandran et al. 2011), have also been demonstrated. Performance in set point tracking and disturbance rejection was evaluated to finalize the control system design. The Novartis-MIT center proposed another plant-wide control strategy, e.g., involving a bottom-up and hierarchical top-down approach, which stabilizes the underlying

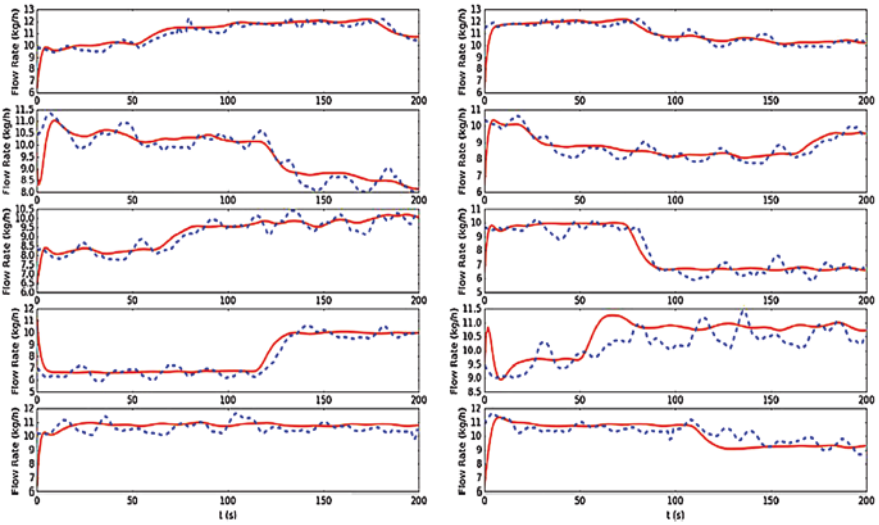


Fig. 6.9 Measured (dashed line) and predicted (solid line) total powder mass flow rates (Reprint permission obtained from Liu et al. 2018. Copyright © 2018 Elsevier)

process control layers upward and prioritizes the control objectives downward (Lakerveld et al. 2013, 2015). This work has also encouraged attempts to categorize control techniques and systematically evaluate their comparative performance with respect to aspects that are of concern to regulators.

In this chapter, a systematic framework for control design and risk analysis for continuous pharmaceutical solid-dosage manufacturing is briefly introduced to achieve efficient and robust active process control, as shown in Fig. 6.10. Its implementation to a feeding-blending system is also demonstrated here (Su et al. 2017). The framework consists of the following components: system identification with state-space models; control design and analysis metrics (see Fig. 6.11); a hierarchical three-level control structure; risk mapping, assessment, and planning (Risk MAP) strategies; and control performance indicators. Specifically, instead of relying on transfer function or empirical convolution models for system identification, state-space models are advocated. State-space models offer advantages due to their inherent flexibility to represent multivariate nonlinear processes and to model unmeasured disturbances and process time delays (Darby and Nikolaou 2012). In addition to the commonly used integral of time absolute error (ITAE) for control tuning or performance evaluation, three new performance indicators are also proposed therein to directly relate the control performance to consistent product quality.

The studied continuous feeding-blending system (Su et al. 2017) consists of two Schenck AccuRate PureFeed® AP-300 loss-in-weight feeders that are capable of achieving and maintaining specified feed rates through an imbedded level 0 control system. The feeders continuously feed the API, acetaminophen (APAP), and excipient, Avicel microcrystalline cellulose PH-200 (MCC 200), into a Gericke GCM 500

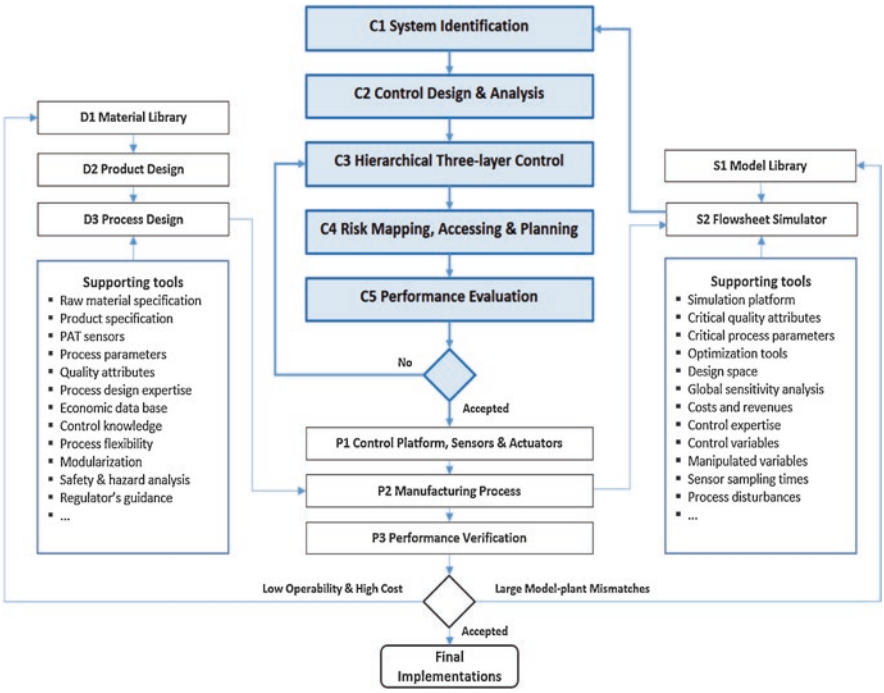


Fig. 6.10 A systematic framework for fault-tolerant process control design and risk analysis. (Reprint permission obtained from Su et al. 2017. Copyright © 2017 Springer Nature)

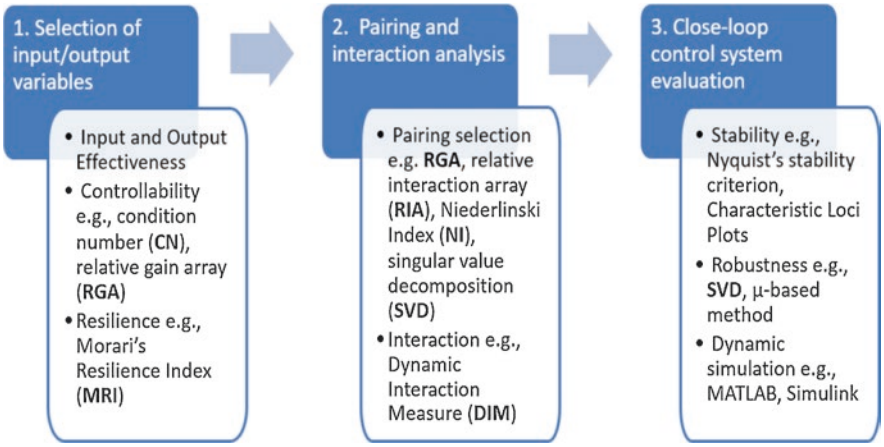


Fig. 6.11 A three-step procedure for control design and interaction analysis. (Reprint permission obtained from Su et al. 2017. Copyright © 2017 Springer Nature)

continuous blender, wherein the two components are mixed. The nominal operating conditions (NOC) consist of API flow of 1.0 kg/h, excipient flow of 9.0 kg/h, and blender rotation speed of 200 rpm. The API mass fraction was measured in situ using a near-infrared spectrometer (Control Development, Inc.) at the exit of the blender (Austin et al. 2014; Vanarase et al. 2010). The content uniformity was statistically estimated in the form of relative standard deviation (RSD) using mean and variance of the API mass fraction measurements within a time window (Zhao et al. 2013). The powder flow is measured using an X-ray-based mass flow meter (SETXvue XP-300, Enurga, Inc.) (Ganesh et al. 2017). The corresponding control schemes at different levels for the feeding-blending system are provided in Table 6.1.

For system identification, a system transfer function matrix **G** was identified using the MATLAB system identification toolbox in the form of a state-space model. With the system transfer function matrix **G**, control design and analysis metrics in Fig. 6.11 can be applied.

The condition number of 3.2945 (<25) and Morari's resilience index of 0.4372 show that the feeding-blending system was controllable and stable at the current nominal operating condition even with a decentralized SISO control loops at level 1. Furthermore, the Niederlinski index of 1.10 (>0), the relative gain array (RGA), and the relative interaction analysis (RIA) pairing metrics all indicate that a stable design was achieved using a diagonal pairing under which the API composition was controlled by manipulating the API feeding flow rate, the blender powder flow rate by the excipient feeding flow rate, and the API mixing RSD by the blender rotation speed, as shown below:

$$\mathrm{RGA} = \begin{bmatrix} \mathbf{0.9008} & 0.0999 & -0.0007 \\ 0.0944 & \mathbf{0.9058} & -0.0002 \\ 0.0048 & -0.0057 & \mathbf{1.0009} \end{bmatrix} \tag{6.3}$$

$$\mathrm{RIA} = \begin{bmatrix} \mathbf{0.1101} & 9.0131 & -1405.4 \\ 9.5942 & \mathbf{0.1040} & -5371.2 \\ 208.81 & -177.54 & \mathbf{-0.0009} \end{bmatrix} \tag{6.4}$$

Table 6.1 The feeding-blending control system. (Su et al. 2017)

Unit operation	Process output (y)	Process input (u)	Control level	Controller type
API feeder	API flow rate	Screw rotation speed	L0	PID
Exp. feeder	Exp. flow rate	Screw rotation speed	L0	PID
Blender	API composition	API flow rate	L1/2	PID, ratio, MPC
	Powder flow rate	Excipient flow rate	L1/2	PID, MPC
	API mixing RSD	Rotation speed	L1/2	PID, MPC
	Rotation speed	Motor current	L0	PID

For the chosen diagonal pairing at level 1 control, the dynamic interaction measure (DIM) of 40.50 > 15 and the diagonal DIM of 60.25 > 15 both suggested that compensation was needed to reduce the interaction, and the μ interaction measure of 2.667 > 1 also indicated that the decentralized controllers at level 1 may not be stable in closed loop. Furthermore, the performance interaction measure (PIM) of 2.00 suggested that there was considerable performance interaction in the system. To address these features, a level 1 design in which the decentralized SISO PID control loops were compensated by a feedforward ratio controller was implemented. A level 2 design was implemented with the identified system state-space model incorporated into a linear MPC algorithm. The three-level control designs are schematically demonstrated in Fig. 6.12.

The overall risk assessment for the three-level control designs can be found in Fig. 6.13, where the controllability of the three acceptable risk scenarios were assessed. An acceptable risk is a risk that is understood and tolerated usually because the cost or difficulty of implementing an effective countermeasure exceeds the expected impact of the risk event on process operations. For instance, reduced flowability of powders may occur due to increased humidity in the environment during the rainy seasons of the year, and this could have an adverse effect on the mixing uniformity of the API and excipient blend. However, it might be too costly to monitor moisture content in feed materials and to add a unit operation for reducing the water content to ensure specified flowability measures. The acceptable risk scenarios identified from the risk mapping are further classified into three categories according to their frequency and severity: R0 low risk, R1 medium risk, and R2 high risk. It is observed that level 0 control was capable of tackling the R0 detectable disturbance in the feeders. However, it failed to respond to risk scenarios at R1 median and R2 high-risk levels, whereas the level 2 control is most advanced in handling the R2 risks of measurement gross errors in PAT probe fouling or drifting. This is due to the fact that the level 2 model-based control adopted the robust Kalman filter as discussed in the previous section. Finally, it should be pointed out here that the PAT tools for CQA measurement play a vital role in the fault-tolerant control design for real-time release strategy in pharmaceutical continuous manufacturing (Yu et al. 2014).

Pharmaceutical continuous manufacturing processes depend on regulatory control systems to maintain consistent quality production. However, changes made to the PAT sensor system and the natural degradation over time of the equipment may cause some controllers and process performance to suffer (Hoo and Piovoso 2003). Although the pioneering work on control performance monitoring was first reported almost three decades ago (Qin 1998), there are also many commercial services and products available (Bauer 2016), e.g., the Emerson DeltaV Tune and the Honeywell Loop Scout®; specific metrics and analysis methods with process knowledge and experience in certain unit operations remain unresolved technical challenges (Jelali 2013), for example, the control performance monitoring and continuous improvement strategies in control systems for the continuous feeding-blending system or for the plant-wide pharmaceutical continuous manufacturing process. Hence, several research questions and regulatory concerns are to be explored as indicated in the proposed systematic framework.

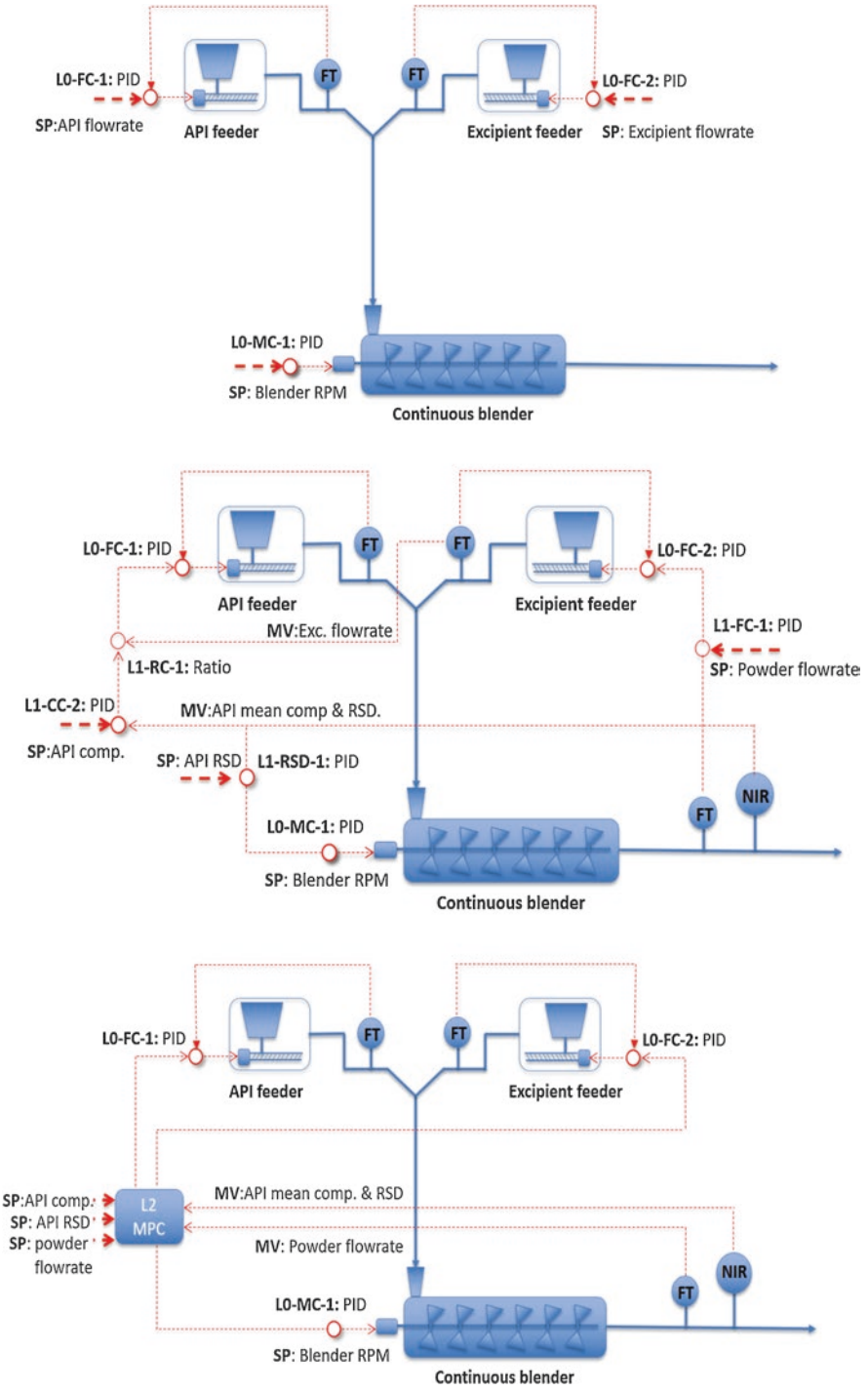


Fig. 6.12 Level 0 (top), level 1 (center), and level 2 (bottom) for the feeding-blending system. (Reprint permission obtained from Su et al. 2017. Copyright © 2017 Springer Nature)

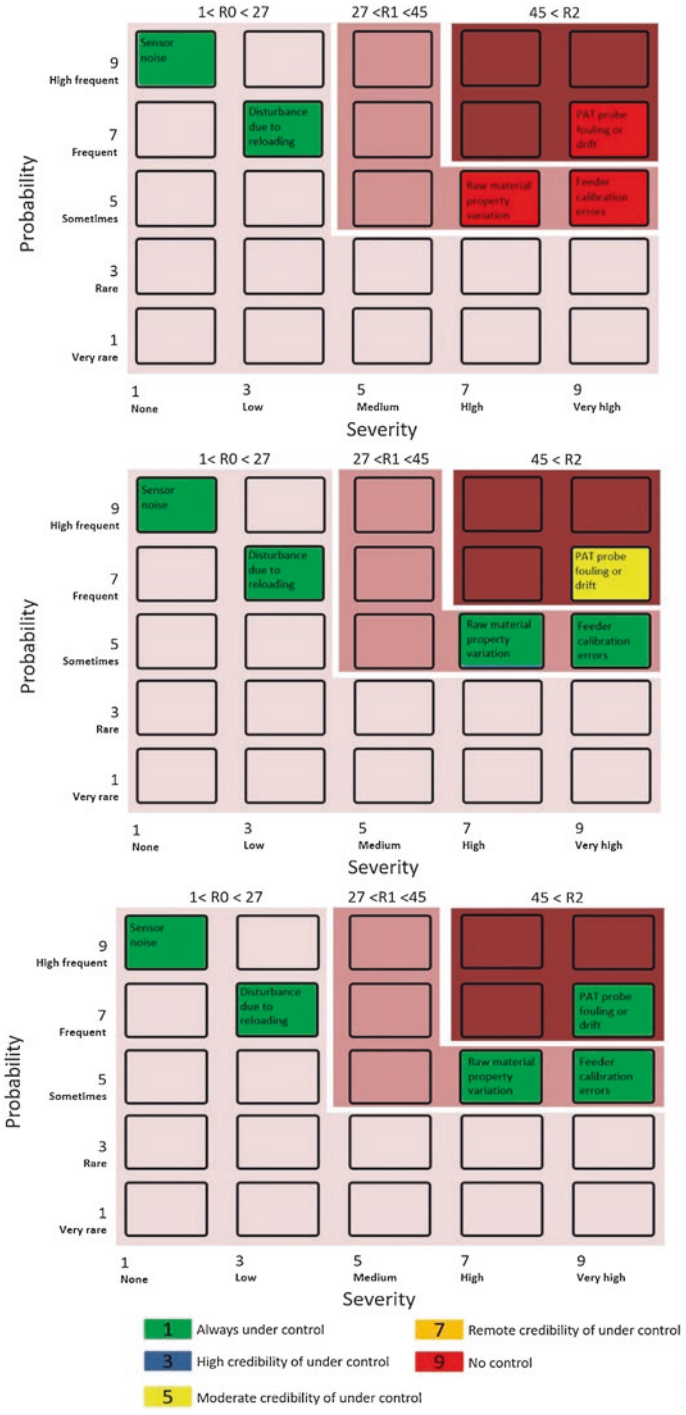


Fig. 6.13 Risk assessment for three-level control designs (top three, level 0 control; center three, level 1 control; bottom three, level 2 control). (Reprint permission obtained from Su et al. 2018a, b. Copyright © 2018 Elsevier)

To sum up, with the development of generic systematic frameworks for control design in pharmaceutical continuous manufacturing, the feeding-blending system will be receiving more research interests to address the challenges highlighted in Sect. 6.2 in this chapter to deliver a resilient and fault-tolerant control design for both process automation and quality control purposes.

6.7 Conclusions and Future Work

Continuous feeding-blending unit operation in pharmaceutical continuous manufacturing is the first processing step in most pharmaceutical solid-dosage manufacturing processes to provide a homogeneous powder mixture of active ingredients and excipients to achieve better processing capabilities in flowability, tabletability, etc., and targeted product attributes in drug potency, release profile, etc. The challenges in continuous feeding-blending process lie not only in the equipment design of feeders and blenders to deliver accurate feeding flow and mixing uniformity, respectively, but also in the integration of these pieces of equipment together to comply to the pharmaceutical quality-by-design guidance which requires the development of in-depth process knowledge in a design space during process design, process monitoring, process modeling, and process control of a feeding-blending system.

The recent developments in continuous feeding-blending system have seen the rapid progresses, specifically, in process analytical technologies in real-time process monitoring of ingredient mass fraction and relative standard deviation using near-infrared spectroscopy and process control strategies in pilot plant facilities.

However, the acute operating conditions of involving fine cohesive powders, especially in the continuous blending unit, have led the fouling problem as one of the challenging issues in many PAT sensor applications. Furthermore, besides the chemometric model calibration for most of the PAT sensors based on spectroscopy, common raw spectra data pretreatment methods, e.g., Savitzky-Golay filter and extended multiplicative scatter correction, are inefficient in handling environmental uncertainties, raw material property variations, or process disturbances, e.g., humidity, powder flow rate, particle size distribution, etc. For example, the NIR spectroscopy collects the complex diffusions that are sensitive to both chemical compositions and physical properties, making a reliable NIR sensor especially challenging. Specifically, though the NIR sensing has been acknowledged as a convenient spectroscopy method without the need of sample pretreatment, except the various in-house designs of NIR boxes or flow chutes in academic research centers, there is a limited number of reports in open literature in terms of reliable application of NIR sensors in the pharmaceutical industry. Hence, a complete process engineering solution to the secondary development of most commercially available spectroscopic probes is indeed necessary to produce reliable PAT sensors, for example, engineering solutions of sensor selection, sampling automation, sensor placement, chemometric calibration, model maintenance, sensor redundancy and network design, and

data reconciliation, are critically important for a highly reliable PAT sensor to keep a good track of CQAs in pharmaceutical manufacturing.

In the aspect of process control development, continuous improvement in feeding and blending process control should also be pursued as a consequence of a deeper understanding of the manufacturing system and its components which naturally develop as manufacturing experience where product and process are accumulated. Given the real and perceived regulatory burden of change approvals, continuous improvement has not been pursued as aggressively in pharmaceutical manufacturing as it might. Hence, the advent of continuous pharmaceutical manufacturing opens the door to continuous improvement at multiple levels, e.g., predictive maintenance, control performance monitoring, control structure reorganizing, etc., since such improvements can be targeted to achieve tighter tracking of CQA and more robust plant-wide control which will maintain the process within its design space. The direct impact is to allow longer continuous runs without forced interruption, reduced frequency and duration of periods during which nonconforming materials are generated, and reduced risk that a product lot released may actually include nonconforming material.

Last but not least, more investigation efforts are also encouraged to develop and evaluate the practice of pharmaceutical continuous feeding-blending design based on science- and risk-based process understanding and approaches. These include but are not limited to data reconciliation to address PAT measurement uncertainties, unmeasured material parameter, and process state estimation to characterize the variations in critical material attributes and critical process parameters, risk analysis and management, etc. Conclusively, process system engineering tools are the most prestigious engineering tools in manufacturing industry to deal with the challenges in pharmaceutical continuous feeding and blending unit operation.

References

Aissa AA, Duchesne C, Rodrigue D. Transverse mixing of polymer powders in a rotary cyliner part I: active layer charaterization. Powder Technol. 2012;219:193–201.

Austin J, Gupta A, McDonnell R, Reklaitis GV, Harris MT. A novel microwave sensor to determine particulate blend composition on-line. Anal Chim Acta. 2014;819:82–93.

Bauer BH. The current state of control loop performance monitoring-a survey of application in industry. J Process Control. 2016;38:1–10.

Blackshields CA, Crean AM. Continuous powder feeding for pharmaceutical solid dosage form manufacture: a short review. Pharm Dev Technol. 2018;23(6):554–60.

Boukouvala F, Niotis V, Ramachandran R, Muzzio FJ, Ierapetritou MG. An integrated approach for dynamic flowsheet modeling and sensitivity analysis of a continuous tablet manufacturing process. Comput Chem Eng. 2012;42:30–47.

Bridgwater J. Mixing of powders and granular materials by mechanical means - a perspective. Particuology. 2012;10(4):397–427.

Câmara MM, Soares RM, Feital T, Anzai TK, Diehl FC, Thompson PH, Pinto JC. Numerical aspects of data reconciliation in industrial applications. Processes. 2017;5(4):56.

Cao H, Mushnoori S, Higgins B, Kollipara C, Fermier A, Hausner D, et al. A systematic framework for data management and integration in a continuous pharmaceutical manufacturing processing line. Process. 2018; In press.
Chen Z, Lovett D, Morris J. Process analytical technologies and real time process control a review of some spectroscopic issues and challenges. J Process Control. 2011;21(10):1467–82.
Cholayudth P. Establishing blend uniformity acceptance criteria for oral solid-dosage forms. Pharm Technol. 2017;41(2):42–52.
Coperion. Micro screw feeders. 2018, June 14. Retrieved from Coperion: https://www.coperion.com/en/products-services/process-equipment/feeders/micro-screw-feeders/.
Coupe A. P2, or not P2: that is the question from development to design. CMAC annual open day, 2015, Glasgow.
Darby ML, Nikolaou M. MPC: current practice and challenges. Control Eng Pract. 2012;20:328–42.
Engisch W, Muzzio F. Method for characterization of loss-in-weight feeder equipment. Powder Technol. 2012;228:395–403.
Engisch W, Muzzio F. Feed rate deviations caused by hopper refill of loss-in-weight feeders. Powder Technol. 2015;283:389–400.
Faqih A, Alexander A, Muzzio F, Tomassone M. A method for predicting hopper flow characteristics of pharmaceutical powders. Chem Eng Sci. 2007;62:1536–42.
FDA, U. U.S. Guidance for industry: Q8(2) pharmaceutical development. Maryland: Food and Drug Administration; 2009.
Franklin GF, Powell JD, Workman ML. Digital control of dynamic systems. 2nd ed., World student series ed: Reading, Mass Addison-Wesley; 1990.
Ganesh S, Troscinski R, Schmall N, Lim J, Nagy Z, Reklaitis G. Application of x-ray sensors for in-line and non-invasive monitoring of mass flow rate in continuous tablet manufacturing. J Pharm Sci. 2017;106(12):3591–603.
García-Muñoz S, Butterbaugh A, Leavesley I, Manley L, Slade D, Bermingham S. A flowsheet model for the development of a continuous process for pharmaceutical tablets: an industrial perspective. AICHE J. 2018;64:511–25.
Guo S, Liu P, Li Z. Inequality constrained nonlinear data reconciliation of a steam turbine power plant for enhanced parameter estimation. Energy. 2016;103:215–30.
Hoo K, Piovoso MJ. Process and controller performance monitoring: overview with industrial applications. International Journal of Adaptive Control and Signal Processing. 2003;17:635–62.
Ierapetritou M, Muzzio F, Reklaitis G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AICHE J. 2016;62(6):1846–62.
Ierapetritou M, Sebastian E-EM, Singh R. Process simulation and control for continuous pharmaceutical manufacturing of solid drug products. In: Kleinebudde P, Khinast J, Rantanen J, editors. Continuous manufacturing of pharmaceuticals: Hoboken NJ, Wiley; 2017. p. 33–105.
Jelali M. Control performance management in industrial automation: assessment, diagnosis and improvement of control loop performance. London: Springer-Verlag; 2013.
Kemeny G, Stuessy G. Imaging the blending processes. Pharm Manuf. 2012, January 03:1–2.
Lakerveld R, Benyahia B, Braatz RD, Barton PI. Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. AICHE J. 2013;59(10):3671–85.
Lakerveld R, Benyahia B, Heider PL, Zhang H, Wolfe A, Testa CJ, et al. The application of an automated control strategy for an integrated continuous pharmaceutical pilot plant. Orgainc Process Research & Development. 2015;19:1088–100.
Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10(3):191–9.
LFA. What is the importance of blend uniformity in the pharmaceutical industry. Retrieved 20 June 2018 from LFA tablet presses: https://www.lfatabletpresses.com/articles/importance-blend-uniformity.
Li X, Wang N, Wang L, Kantor I, Robineau J-L, Yang Y, Marechal F. A data-driven model for the air-cooling condenser of thermal power plants based on data reconciliation and support vector regression. Appl Therm Eng. 2018;129:1496–507.

Liu J, Su Q, Moreno M, Laird C, Nagy Z, Reklaitis G. Robust state estimation of feeding-blending systems in continuous pharmaceutical manufacturing. Chemical Engineering Reserach and Design. 2018;134:140–53.
Marikh K, Berthiaux H, Mizonov V, Barantseva E. Experimental study of the stirring conditions taking place in a pilot plant continuous mixer of particulate solids. Powder Technol. 2005;157:138–43.
Markl D, Wahl PR, Menezes JC, Koller DM, Kavsek B, Francois K, et al. Supervisory control system for monitoring a pharmaceutical holt melt extrusion process. AAPS PharmSciTech. 2013;14(3):1034–44.
Mason RL, Young JC. Multivariate statistical process control with industrial applications. Philadelphia: The American Statistical Association and the Society for Industrial and Applied Mathematics; 2002.
Moreno M, Liu J, Ganesh S, Su Q, Yazdanpanah N, O'Connor T, et al. Steady-state data reconciliation of a direct compression tableting line. AIChE annual meeting. Minneapolis: AIChE Annual Meeting; 2017.
Nowak S. Improving feeder performance in continuous pharmaceutical operations. Pharm Technol. 2016;40(10):68–73.
Oka S, Muzzio F. Continuous powder blenders for pharmaceutical applications. Heidelberg: Springer; 2012.
Oka SS, Escotet-Espinoza MS, Singh R, Scicolone JV, Hausner DB, Ierapetritou M, Muzzio FJ. Design of an integrated continuous manufacturing system. In: Kleinebudde P, Khinast J, Rantanen J, editors. Continuous manufacturing of pharmaceuticals: Hoboken NJ, Wiley; 2017. p. 405–46.
Osorio J, Vanarase A, Romanach R, Muzzio F. Continuous powder mixing. In: Cullen P, Romanach R, Abatzoglou N, Rielly C, editors. Pharmaceutical blending and mixing. Chichester: Wiley; 2015.
Pernenkil L. Continuous blending of dry pharmaceutical powders. Boston: Massachusetts Institute of Technology; 2008.
Pernenkil L, Cooney CL. A review on the continuous blending of powders. Chem Eng Sci. 2006;61(2):720–42.
Portillo P, Ierapetritou M, Muzzio F. Characterization of continuous convective powder mixing processes. Powder Technol. 2008;182(3):368–78.
Previdi F, Belloli D, Cologni A, Savaresi SM, Cazzola D, Madaschi M. Control system design for a continuous gravimetric blender. Preprints of the 18th IFAC world congress, pp. 1025–1030, 2011, Milano, Italy.
Qin SJ. Control performance monitoring-a review and assessment. Comput Chem Eng. 1998;23:173–86.
Rafiee A, Behrouzshad F. Data reconciliation with application to a natural gas processing plant. J Nat Gas Sci Eng. 2016;31:538–45.
Ramachandran R, Arjunan J, Chaudhury A, Ierapetritou MG. Model-based control-loop performance of a continuous direct compaction process. J Pharm Innov. 2011;6:249–63.
Rehrl J, Kruisz J, Sacher S, Khinast J, Horn M. Optimized continuous pharmaceutical manufacturing via model-predictive control. Int J Pharm. 2016;510:100–15.
Rockoff JD. Drug making breaks away from its old ways. 2015, Feb 8. Retrieved October 6, 2016, from The Wall Street Journal: http://www.wsj.com/articles/drug-making-breaks-away-from-its-old-ways-1423444049.
Rogers A, Hashemi A, Ierapetritou M. Modeling of particulate processes for the continuous manufacturing of solid-based pharmaceutical dosage forms. Processes. 2013;1:67–127.
Scheibelhofer O, Balak N, Wahl PR, Koller DM, Glasser BJ, Khinast JG. Monitoring blending of pharmaceutical powders withi multipoint NIR spectroscopy. AAPS PharmSciTech. 2013;14(1):234–44.
Sen M, Dubey A, Singh R, Ramachandran R. Mathematical development and comparison of a hybrid PBM-DEM description of a continuous powder mixing process. J Powder Technol. 2013;2013:1–11.

Singh R, Gernaey KV, Gani R. Model-based computer-aided framework for design of process monitoring and analysis systems. Comput Chem Eng. 2009;33(1):22–42.
Singh R, Ierapetritou M, Ramachandran R. System-wide hybrid model predictive control of a continuous pharmaceutical tablet manufacturing process via direct compaction. Eur J Pharm Biopharm. 2013;85(3 Part B):1164–82.
Singh R, Sahay A, Muzzio F, Ierapetritou M, Rohit R. A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. Comput Chem Eng. 2014;66:186–200.
Singh R, Muzzio FJ, Ierapetritou M, Rohit R. A combined feed-forward/feed-back control system for a QbD-based continuous tablet manufacturing process. Processes. 2015a;3:339–56.
Singh R, Sen M, Ierapetritou M, Ramachandra R. Integrated moving horizon-based dynamic real-time optimization and hybrid MPC-PID control of a direct compaction continuous tablet manufacturing process. J Pharm Innov. 2015b;10:233–53.
Su Q, Moreno M, Giridhar A, Reklaitis GV, Nagy ZK. A systematic framework for process control design and risk analysis in continuous pharmaceutical solid-dosage manufacturing. J Pharm Innov. 2017;12:327–46.
Su Q, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK Variation and risk analysis in tablet press control for continuous manufacturing of solid dosage via direct compaction. The 13th international symposium on process systems engineering PSE 2018a, San Diego.
Su Q, Moreno M, Ganesh S, Reklaitis GV, Nagy ZK. Resilience and risk analysis of fault-tolerant process control design in continuous pharmaceutical manufacturing. J Loss Prev Process Ind. 2018b;55:411–22.
Su Q, Bommireddy Y, Shah Y, Ganesh S, Moreno M, Liu J, et al. Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. Int J Pharm. 2019a;563:259–72.
Su Q, Ganesh S, Moreno M, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. Comput Chem Eng. 2019b;125:216–31.
Valdetaro ED, Schirru R. Simultaneous model selection, robust data reconciliation and outlier detection with swarm intelligence in a thermal reactor power calculation. Ann Nucl Energy. 2011;38(9):1820–32.
Vanarase AU, Muzzio FJ. Effect of operating conditions and design parameters in a continuous powder mixer. Powder Technol. 2011;208(1):26–36.
Vanarase AU, Alcalà M, Jerez Rozo JI, Muzzio FJ, Romañach RJ. Real-time monitoring of drug concentration in a continuous powder mixing process using NIR spectroscopy. Chem Eng Sci. 2010;65:5728–33.
Velázquez C, Florían M, Quinones L. Monitoring and control of a continuous tumble mixer. Computer Aided Chemical Engineering. 2018;41:471–87.
Warman M. Continuous processing in secondary production. In: Ende DJAM, editor. Chemical engineering in the pharmaceutical industry - R&D to manufacturing. New Jersey: Wiley; 2011. p. 837–51.
Weinekotter R, Gericke H. Mixing of solids: Kluwer Academic Publishers, Dordrecht, the Netherlands; 2000.
Weiss GH, Romagnoli JA, Islam KA. Data reconciliation-an industrial case study. Computers Chemical Engineering. 1996;20(12):1441–9.
Yoon S, Galbraith S, Cha B, Liu H. Chapter 5 - Flowsheet modeling of a continuous direct compression process. Computer Aided Chemical Engineering. 2018;41:121–39.
Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, Woodcock J. Understanding pharmaceutical quality by design. AAPS J. 2014;16(4):771–83.
Zhao XJ, Gatumel C, Dirion JL, Berthiaux H, Cabassud M. Implementation of a control loop for a continuous powder mixing process. Proceeding of the 2013 AIChE annual meeting, 2013, San Francisco.

Chapter 7
Recent Progress in Roll Compaction Process Development for Pharmaceutical Solid Dosage Form Manufacture

Ariel R. Muliadi, Alamelu Banda, and Chen Mao

Abstract Roll compaction technologies have advanced substantially over the past decade and are increasingly adapted by formulation scientists as a preferred means of granulation. This chapter presents recent progress in pharmaceutical roll compaction with respect to the following three key aspects: formulation development, process modeling, and equipment. On the modeling front, we outline various continuum-based modeling efforts and their role in furthering the fundamental, mechanistic understanding of the process. Here, we discuss the significant body of work that has been derived from the rolling theory for granular solids, focusing on both the quantities that that model predicts accurately and the reasons it recurrently falls short in predicting various process outcomes, as well as the rise of high-fidelity computational simulations and their potential to overcome the shortcomings of the rolling theory. With regard to the formulation development aspects of the process, common pharmaceutical excipients and their impact on the resulting ribbon and granule qualities are discussed. Lastly, we present the current state of the art in equipment design and process control that enable the processing of a wide variety of pharmaceutical powders. Emphases are placed on different roll compactor options and their ability to precisely control the roll force and roll gap, perform efficient deaeration, and minimize powder leakage.

Keywords Pharmaceutical · Roll compaction · Dry granulation · Modeling · Johanson · Finite element analysis · Ribbon · Density · Solid fraction

A. R. Muliadi (✉) · A. Banda · C. Mao
Genentech Inc., Small Molecule Pharmaceutical Sciences, Research and Early Development (gRED), South San Francisco, CA, USA
e-mail: muliadi.ariel@gene.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_7

7.1 Introduction

Pharmaceutical powder handling and processing often necessitate the need for granulation. Through this process, fine powders are agglomerated into (sub)millimeter-sized granules via the use of an aqueous binding agent ("wet" granulation) or application of mechanical force ("dry" granulation). Owing to the granules' larger size and increased bulk density, reduction in dusting, improved flowability, and a more consistent die filling may be achieved. As an added benefit, components comprising a powder blend are also "locked in" within the individual granules, therefore preventing them from segregating when subjected to further handling/processing.

When granulating moisture-sensitive drugs or when integration of an additional processing step for post-wet granulation moisture removal is unfeasible, dry granulation is often performed in place of wet granulation via the use of a roll compactor.

A schematic of a roll compactor is shown in Fig. 7.1. The roll compaction process begins when powder, typically a pre-blended mix of solid active pharmaceutical ingredients (APIs) and excipients (fillers, disintegrants, lubricants, etc.), enters the compactor hopper. Immediately upon discharge from the hopper, the powder is conveyed by an auger system toward two counterrotating rolls that are mounted such that there exists a small gap ("roll gap") between them. Friction between the rolls and the powder allows the rolls to grip the powder and drive it toward the gap, compacting it into a thin strip ("ribbon") that is then broken apart into granules using a mill. In pharmaceutical solid dosage form manufacture, the granules may then be further processed, blended with additional excipients, and ultimately compacted to form tablets.

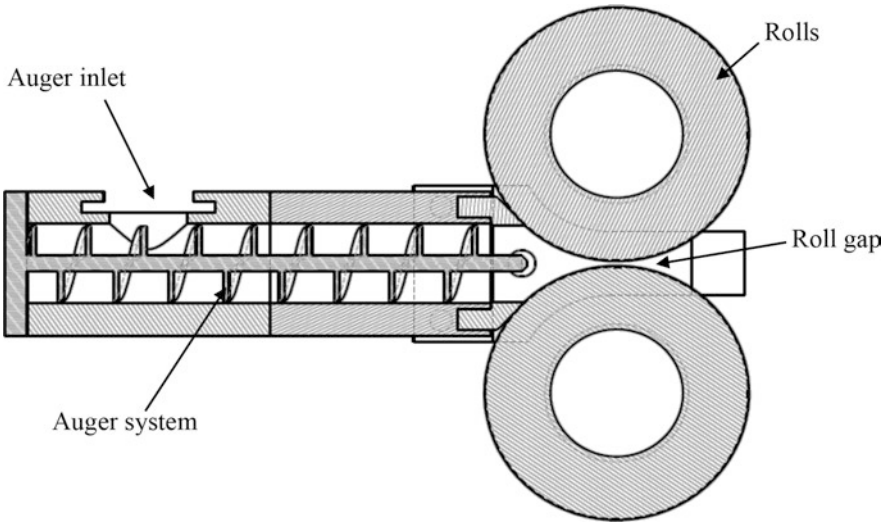


Fig. 7.1 Cross-sectional view of a roll compactor

The roll compaction process is not unique to the manufacturing of pharmaceutics; rather, it is used in many industries. What we know today as pharmaceutical roll compaction evolved from "cold rolling," a metalworking technique that dates back to the late 1500s. Fast-forward to the 1800s, increasing demands for coal as a fuel source prompted industrial revolutionists to adapt that process for compressing coal dusts into coal briquettes. Fast-forward to the late twentieth century, the rising popularity of continuous manufacturing incentivized a wide adaptation of that process in the pharmaceutical industry, owing in part to the fact that it is inherently continuous, plus its many benefits mentioned above.

As is often the case for many pharmaceutical unit operations and despite what may conceptually seem uncomplicated, the roll compaction process is not well understood. Consequently, up until a decade ago, our knowledge of the process is limited to empiricist and utilitarian "know-hows" obtained through extensive—but not necessarily scientific—factorial experiments.

Such factorial experiments are not the focus of this chapter. Instead, we provide a review of the studies that aim to impart mechanistic understanding of the roll compaction process. The discussions are provided in four main subchapters. The first subchapter focuses on a handful of analytical and computational analyses available for modeling the roll compaction process. The second and third subchapters focus on materials and equipment considerations for the roll compaction process. The last subchapter briefly discusses process development strategies and scale-up approaches. Experimental studies that highlight factors that are often difficult to account for in a mechanistic model will also be presented.

The goal is to introduce readers with factors and perspectives that twenty-first-century formulators must consider when designing a roll compaction process.

7.2 Roll Compaction Models

The characteristics of roll-compacted milled granules are set primarily by the ribbon density: studies (Kleinebudde 2004; von Eggelkraut-Gottanka et al. 2002; Inghelbrecht and Paul Remon 1998a, b) have reported that the average granule size is directly proportional to the ribbon density. In addition, as no further compression is applied during the milling of the ribbon, the density of the individual granules—which determines their "reworkability" for further compaction processes (Wu and Sun 2007)—is dependent on the ribbon density. The poured and tapped densities of the bulk granules have also been found to directly correlate with the ribbon density (Hancock et al. 2003).

Considering the importance of ribbon density, it is not surprising that most, if not all, currently available analytical and computational models that predict the roll compaction process put the most emphasis on estimating what parameter and factors directly influence it. Roll compaction models may be classified according to their dimensionality: one-dimensional models, e.g., the Johanson model (1965) and the "slab" method (Katashinskii 1986; Katashinskii and Shtern 1983a, b) plus its

derivatives, (Dec 1991; Cunningham 2005) two-dimensional numerical simulations utilizing the Galerkin (Gun et al. 1986) or finite element methods (Cunningham 2005; Zavaliangos et al. 2003; Muliadi et al. 2012) and upper-bound theorem-based models (Shima and Yamada 1984); and three-dimensional finite element models (Cunningham 2005; Cunningham et al. 2010; Muliadi et al. 2013; Mazor et al. 2016, 2017).

The above modeling approaches are similar to one another in that they are all continuum level and require user-specific boundary conditions to define the roll compactor geometry (e.g., roll diameter, the roll gap distance, etc.), the inlet condition (e.g., inlet streamwise normal stress, powder inflow velocity, etc.), the powder/tooling friction angle, and the material constitutive parameters. Depending on the model's simplifying assumption and solution method, additional user inputs may be required—these are highlighted in the paragraphs that follow. The bulk of the discussions will revolve around recent work that utilizes finite element method (FEM) and the Johanson model, the latter for its widespread adaptation in the pharmaceutical industry.

The slab method (Katashinskii 1986; Katashinskii and Shtern 1983a, b) considers the equilibrium force and mass balance equations for one-dimensional trapezoidal plane elements ("slabs") of differential length that represent the deforming powder downstream of the nip angle—the angle at which the transition from the powder slipping on the roll surface to it being gripped by the rolls and drawn toward the roll gap occurs. The mass and momentum conservations are related to one another via empirically determined stress-strain constitutive relations. The resulting systems of equations are then solved to yield roll normal stress distribution and ribbon density. The slab method requires specification of experimentally measured nip angle. Its other inputs include the aforementioned user-specified boundary conditions, as well as powder density at the nip angle.

Experimental validations for the slab method are presented in Katashinskii and Shtern (1983a), Dec (1991), and Dec et al. (2003). Katashinskii and Shtern (1983a) compared the slab model-predicted roll normal stress distributions and ribbon density with those measured experimentally. The boundary conditions were based on unexplained, assumed values. The authors found that the computed roll normal stress distributions qualitatively agreed with experimental trends, but their magnitudes were an order of magnitude smaller than their experimentally measured counterparts. They also found that the computed ribbon density was independent on the inlet stress—an observation that did not agree with the experimental data.

Further validations for the slab method's roll normal stress distribution results are presented in Dec (1991) and Dec et al. (2003) Here, the inlet stress and density inputs to the model were used as a fitting parameter, i.e., they were determined by means of iterations so the model-predicted ribbon density matched that measured experimentally. Experimentally measured powder stress-strain constitutive properties, powder/roll friction angle, and the nip angles were used as additional model inputs. Similar to Katashinskii and Shtern's findings, Dec et al.'s roll normal stress predictions qualitatively agreed with their experimental data. Quantitative comparisons, however, showed that the computed roll normal stress distributions were

significantly higher than the measured ones for five of the cases considered but were lower for the remaining cases (Dec et al. 2003).

Upper-bound theorem-based roll compaction models compute the nip angle, roll normal stress distribution, and ribbon density by fitting a kinematically admissible velocity field to the flow of powder downstream of the nip angle such that the energy dissipation rates due to powder densification and roll/powder friction are minimized. As well as roll diameter, roll gap distance, powder stress-strain constitutive relations, and roll/powder friction angle, inputs to the model include powder density at the nip angle and roll velocity. Contrasting the one-dimensional Johanson and slab models, upper-bound theorem-based models described in Shima and Yamada (1984) and Deshmukh et al. (1998) acknowledge a two-dimensional velocity field in the powder flow; in the former study, for example, the authors assumed that the powder velocity at the roll/powder interface was equal to the roll circumferential velocity but varies linearly in the direction perpendicular and parallel to the streamwise direction.

Shima and Yamada (1984) also performed experimental studies to validate ribbon density results from an upper-bound theorem-based model. The powder density at the nip angle and the powder/roll friction angle were based on assumed values. Good agreements between predicted and measured results were observed. Deshmukh et al. (1998) applied a similar upper theorem-based approach to compute the density of roll-compacted porous metal sheets. While the sheet density at the inlet was experimentally measured, the powder/roll friction angle constant was arbitrarily assumed. Predicted and measured results again showed good agreements.

Rapid advances in computational processing and proliferations of end-user finite element software packages in recent years have led to widespread adaptation of two- and three-dimensional finite element method for modeling the roll compaction process. This high-fidelity approach also provides additional depths that are otherwise absent from the models discussed above. For example, using a three-dimensional finite element method, it is possible to simulate the density variations along a ribbon width ("spanwise") and length ("streamwise"). The ribbon spanwise density distribution is experimentally reported as parabolic in profile, with the ribbon edges having a lower density than its center for rolls equipped with stationary cheekplates (i.e., the platen seals on either sides of the rolls) (Cunningham 2005; Cunningham et al. 2010; Muliadi et al. 2013; Michrafy et al. 2011; Miguelez-Moran et al. 2009) or vice versa for roll compactors whose cheekplates rotate with the rolls (Mazor et al. 2016; Wang et al. 1998). The ribbon streamwise density distribution, on the other hand, consists of strips of dense sections that oscillate along the ribbon centerline in a periodic manner (Michrafy et al. 2011; Guigon and Simon 2003; Simon and Guigon 2003), a behavior hypothesized to result from the periodic, helical flight of the auger conveyor (Michrafy et al. 2011; Guigon and Simon 2003; Simon and Guigon 2003). An immediate consequence of these density distributions is that granules milled from the edges of the ribbon can potentially differ in characteristics from those milled from the ribbon center portions, leading to more fine productions and/or tablet capping (Kleinebudde 2004) due to variations in the reworkability of the granules.

Two- and three-dimensional finite element simulations are performed by first dividing the deforming powder region into a number of continuously distributed elements. The kinematics concerning the motion and deformation (strain) of each of the elements are then computed numerically, the latter through stress-strain constitutive relations, to yield solutions that satisfy mass, momentum, and energy conservations. For simple cases, these kinematical solutions may also be acquired without the use of the mesh elements via numerical schemes such as the Galerkin ("element-free") method (Gun et al. 1986).

When applied to modeling the roll compaction process, the deformation of the elements is a consequence of their mechanical properties and interactions with the roll, the inlet boundary condition (velocity or stress), and the roll compactor geometry. The nature of the finite element method also allows modeling to be performed with fewer assumptions than the abovementioned models. The powder "nipping" process, for instance, occurs naturally in an FEM simulation, therefore rendering a priori assumptions or experimentally measured input unnecessary. Typical finite element simulation outputs include roll normal and shear stress distributions, powder velocity and stress fields, plastic strain measures—which can be correlated to powder density (Gurson 1977)—and reaction force and torque at the roll pin.

Examples of powder roll compaction FEM models are those presented in Dec et al. (2003), Wang et al. (1998), Muliadi et al. (2012, 2013), Liu and Wassgren (2016), Zavaliangos et al. (2003), Cunningham (2005), Cunningham et al. (2010), and Michrafy et al. (2011) Interestingly, comparison of FEM results with experimental data is absent for many of these studies; experimental validations were only presented in Michrafy et al. (2011) and Mazor et al. (2016), albeit only qualitatively, in Muliadi et al. (2013), and indirectly in Cunningham (2005).

Three-dimensional FEM models providing ribbon density distributions were presented in Michrafy et al. (2011) Here, the Drucker-Prager/Cap (DPC) model with empirically determined density-dependent constants was used to define the powder stress-strain behaviors. Roll/powder friction coefficient was experimentally measured. Two different inlet boundary conditions were used: a uniform velocity and a uniform stress, both with constant, assumed magnitudes. When using the uniform inlet velocity boundary condition, the predicted ribbon densities were found to be the densest at their center and the least dense at their edges. In contrast, a uniform ribbon density distribution resulted when the uniform inlet stress boundary condition was used; this contradicts similar studies from the literature (e.g., Muliadi et al. (2013)). The former observation qualitatively agreed with experimentally measured density data from a mercury porosimeter. The latter observation may result from improperly set up FEM simulations. As noted above, no quantitative comparisons were given.

Cunningham (2005) used two-dimensional FEM models to compute nip angle and ribbon density. A uniform stress was used as the inlet boundary condition to describe the auger feeding action; its magnitude was experimentally measured via a load cell mounted on one of the cheekplates, perpendicular to the powder flow direction. The powder stress-strain relations were defined using the DPC model with density-independent constants (taken as some average of experimentally

measured data). The roll/powder friction angle was experimentally measured. Model validation was provided by comparing the computed ribbon density at different levels of roll force with the corresponding experimental data. It was found that the two quantities agreed to within 20%, with the predicted results being higher than the experimentally measured ones. Such discrepancies were hypothesized to result from inaccurate inlet boundary conditions and post-gap material expansion that was not captured by the FEM simulations. It is also worth noting that the plane strain assumption used in Cunningham's model will tend to overpredict the roll forces as it does not account for the spanwise ribbon density distribution.

Muliadi et al. (2013) performed three-dimensional FEM simulations to predict the spanwise ribbon density distribution resulting from various roll compaction conditions. This study is unique in that deliberate care was taken to ensure that the experimental and simulated conditions matched. For example, the experiments were performed with an air-powered piston feeder configuration retrofitted to an Alexanderwerk WP-120 roll compactor so that a known, uniform stress could be applied on the powder. The roll gap for each experimental run was tracked using a laser distance sensor. All model boundary conditions were based on experimentally measured values (e.g., powder/tooling friction angles were measured by a ring-shear tester using "wall coupons" manufactured to the same specifications and with the same material as the tooling surfaces). In addition, to further improve model accuracy, the simulations also utilized density-dependent DPC parameters, whose values were taken from Cunningham (2005), to describe the powder mechanical behaviors.

Results show that FEM-predicted density distributions agree well with the corresponding experimental data. Important experimental trends were also successfully captured by the simulations. They include the spanwise ribbon density distribution—consistent with other findings (Cunningham 2005; Cunningham et al. 2010; Michrafy et al. 2011; Miguelez-Moran et al. 2009)—that arose even when the streamwise inlet stress was applied uniformly. The authors postulated that such trends are the result of the powder/tooling interactions that occur in the process: the powder/roll friction drives the powder toward the gap and facilitates nipping, while the powder/cheekplate friction counteracts it. As a result of these two opposing effects, the powder will move at different velocities along the spanwise direction. The powder closest to the cheekplates is subjected to the largest shear due to powder/cheekplate friction forces; therefore, it moves the slowest thereat. In contrast, the powder closest to the roll half-width moves the fastest as the influence of the powder/cheekplate friction is the least significant there. Accordingly, the powder at the center will be nipped further upstream from the gap than at the edges. Since the powder mass available for compaction decreases along the streamwise direction, the ribbon near the cheekplates is the least dense, while at the roll half-width, the ribbon is the densest, leading to the observed density distributions. Muliadi et al.'s hypothesis is supported by their FEM-predicted material velocity distributions (Muliadi et al. 2013).

Another important finding is the observation that the average ribbon density was independent of the inlet stress when the system was operated at conditions that

allowed the roll gap to "float." The authors also noted that the experimentally measured spanwise ribbon density distribution was slightly asymmetrical against the roll half-width—a trend hypothesized to result from cantilevered roll shaft support that caused the rolls to be not perfectly parallel. Incorporating nonparallel roll configurations in the FEM simulations resulted in an asymmetrical ribbon density distribution similar to the experimental observations (Muliadi et al. 2013).

Mazor et al. (2016) used three-dimensional FEM models to predict ribbon density distributions as produced from two different roll compactor setups: one where the seals on either side of the rolls are stationary ("cheekplates") and one where the seals rotate with the rolls ("rimmed rolls"). The FEM model presented therein utilized the DPC model with experimentally measured density-dependent parameters. A constant, uniform stress was defined at the inlet. The inlet stress magnitude was adjusted such that the FEM-predicted roll force matched that set for the experiments. The powder/roll friction angle was assumed to be identical to the powder/seal one, and the value was taken from the experimental data in Muliadi et al. (2013) It is worth noting that Muliadi et al.'s data are specific to the friction angle between microcrystalline cellulose powder (Avicel PH-102) and stainless steel whose surface finish is identical to that of the rolls equipped on an Alexanderwerk WP-120 roll compactor. Mazor et al. used a Gerteis Mini-Pactor, which had different knurling pattern on the rolls and acrylic cheekplates.

Mazor et al.'s results are summarized as follows. First, in the case of stationary cheekplates, the FEM-predicted ribbon density distribution is in agreement with the trends reported in other studies (Cunningham 2005; Cunningham et al. 2010; Michrafy et al. 2011; Miguelez-Moran et al. 2009). The authors also quantitatively compared their FEM results to the experimental data of Muliadi et al. (2013) They noted that, while the FEM-predicted ribbon density near the center of the ribbon matched the corresponding experimental data, the ribbon edges did not. The authors attributed such a trend to material leaking through the cheekplates during the experiments, which they argued would lower the experimentally measured ribbon density along its peripheries. However, note once again that the roll compactor used in Mazor et al. studies differed significantly from that used in Muliadi et al. Muliadi et al.'s Alexanderwerk WP-120 was equipped with a piston feeder and had PTFE cheekplates and smooth rolls that were 120 mm in diameter and 40 mm wide. Mazor et al.'s Gerteis Mini-Pactor had an auger feeder, acrylic cheekplates, and knurled rolls that were 250 mm in diameter and 25 mm wide. As such, it is not surprising that discrepancies between Mazor et al. FEM results and Muliadi et al. experimental data exist.

The FEM-predicted ribbon density corresponding to the rimmed roll setup, on the other hand, appeared largely uniform across the ribbon width. This is in contrast to the authors' experimental data, which showed the ribbon edges having a higher density than its center. No explanation was provided to address that discrepancy.

In a separate study, Mazor et al. (2017) used the discrete element method (DEM) to study the influences of the auger feed helical motion on the distribution of streamwise stress at the roll compactor inlet and how it in turn affected the density distribution of the ribbon. Because DEM is generally not suitable for simulating powder

compression, the authors' simulation considered only regions upstream of the nip angle. The model results showed that the helicity of granular flow as a result of the auger motion imposed a nonuniform stress distribution at the inlet and that ascribing such stress distribution as the inlet boundary condition of an FEM simulation that models the process downstream of the inlet region resulted in a heterogeneous ribbon density pattern similar to the previously discussed experimental observations. While the authors noted that their combined DEM/FEM approach to simulating the roll compaction process could not quantitatively predict the ribbon density distribution, their work confirmed the prevailing hypothesis that its nonuniformity is partly caused by the feed auger action.

Despite promising recent progresses, FEM models are not without any limitations. For example, the arbitrary Lagrangian-Eulerian explicit scheme employed in the above-discussed studies is valid only when forces resulting from the compaction process are much larger than the inertial forces (i.e., when the process is quasi-static). The elastic "springback" that the ribbons may exhibit upon being extruded from the rolls, however, can no longer be approximated as quasi-static as inertial effects become important for this process. This, combined with the unstable mesh oscillations accompanying the springback response (Muliadi et al. 2013) and excessive dilation resulting from an unloading stress path that touches the shear failure surface of the DPC material model employed in those studies (Cunningham 2005; Muliadi et al. 2013), ultimately causes FEM simulations to inaccurately predict the material behavior post the roll gap. While mesh oscillations could be rectified by applying artificial damping on mesh elements located downstream of the roll gap and by specifying small time steps, erroneous results may still stem from the aforementioned limitation of the material constitutive model (Muliadi et al. 2013).

In addition, the computation time of an FEM simulation is considerably longer than that of the phenomenological/analytical models discussed in the beginning of this section—the latter can generate data in a fraction of a second, while the former computation times can range from hours to weeks depending on the model complexity. In addition, the implementation of an FEM model is also more challenging as it involves the use of a numerical solver to compute the equilibrium equations describing the stress and strain states of the deformable body (i.e., the powder). Although the user-friendliness of such software has improved significantly in recent years, the effort involved in using FEM software effectively is much greater than that required for solving algebraic closed form equations from typical analytical models. For these reasons, many recent roll compaction studies still employ the use of analytical models. One model in particular is favored more than others: the rolling theory of granular solids, more commonly referred to as the Johanson model (1965).

The Johanson model provides predictions of the roll normal stress distribution, roll torque and force, nip angle—i.e., the angular location along the roll surface at which powder begins to move at the velocity of the rolls—and ribbon density at the roll gap. For clarity, the Johanson model equations and accompanying assumptions are presented below. More detailed derivations can be found in Johanson (1965).

A schematic of the roll compaction process as per Johanson's analyses is shown in Fig. 7.2. In the slip region (i.e., upstream of x_α), Johanson assumes that the powder follows the Jenike-Shield yield criterion. Under this assumption, the average stress gradient on the roll surface, $d\sigma/dx$, is defined as a function of the angle of friction between the material and the rolls φ, the material effective frictional angle δ, and process geometry (Fig. 7.2):

$$\left.\frac{d\sigma}{dx}\right|_{\text{slip}} = \frac{4\sigma\left[(\pi/2)-\theta-\nu\right]\tan\delta}{\frac{D}{2}\left(1+\frac{s}{D}-\cos\theta\right)\left[\cot(A-\mu)-\cot(A+\mu)\right]}. \tag{7.1}$$

Here, x, θ, s, and D are the distance upstream from the roll gap, the roll angle measured from the roll gap, the roll gap, and the roll diameter, respectively. The parameters A, ν, and μ are constants, calculated using the following relations:

$$A = \frac{1}{2}\left(\theta+\nu+\frac{\pi}{2}\right), \tag{7.2}$$

$$\nu = \frac{1}{2}\left[\pi-\sin^{-1}\left(\frac{\sin\varphi}{\sin\delta}\right)-\varphi\right], \tag{7.3}$$

$$\mu = \frac{\pi}{4}-\frac{\delta}{2}. \tag{7.4}$$

Note that ν may be interpreted as the angle between the major principal stress and the tangent to the roll if drawn on Mohr's circle.

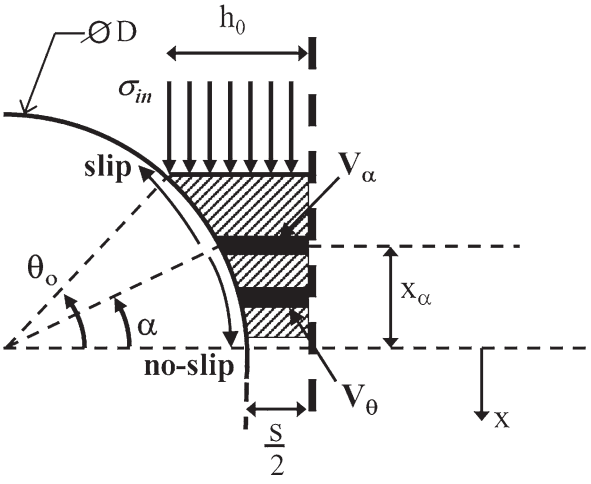


Fig. 7.2 Roll compaction geometry assumed in Johanson's model

For angles smaller than the nip angle, i.e., the no-slip region, Johanson argues that the Jenike-Shield yield behavior no longer applies. Instead, the stress gradient here is derived from the continuity equation:

$$\frac{Dm}{Dt} = \frac{\partial m}{\partial t} + \mathrm{v}.\left(\nabla m\right) = 0, \tag{7.5}$$

where m and $\mathbf{v}$ are the material mass and velocity, respectively. Assuming steady, one-dimensional flow and no slip against the rolls, Eq. (7.5) reduces to:

$$\gamma_\alpha V_\alpha = \gamma_\theta V_\theta, \tag{7.6}$$

where V and γ are, respectively, the volume and relative density of a thin element of material as shown in Fig. 7.2.

Experimental data show that the bulk relative density of a particulate material may be modeled as being proportional to an applied normal stress when plotted on a log-log scale, i.e.:

$$\frac{\sigma_\alpha}{\sigma_\theta} = \left(\frac{\gamma_\alpha}{\gamma_\theta}\right)^K. \tag{7.7}$$

Here, K is the slope of the empirical stress-density plot; it represents the compressibility of the material, with smaller K values corresponding to more compressible materials.

Substituting Eqs. (7.6) to (7.7) gives a normal stress-volume relationship:

$$\frac{\sigma_\theta}{\sigma_\alpha} = \left(\frac{V_\alpha}{V_\theta}\right)^K. \tag{7.8}$$

From Eq. (7.7), the material relative density at the roll gap ($\theta = 0$) may be found using:

$$\gamma_{\theta=0} = \gamma_\alpha \left(\frac{\sigma_{\theta=0}}{\sigma_\alpha}\right)^{\frac{1}{K}}, \tag{7.9}$$

assuming that the material relative density at the nip angle ($\theta = \alpha$) is known. Furthermore, by calculating the volume as a function of s, D, and θ via trigonometric relations, the normal stress in the no-slip region can be derived from Eq. (7.8) using:

$$\sigma_\theta = \sigma_\alpha \left[\frac{\left(1 + \frac{s}{D} - \cos\alpha\right)\cos\alpha}{\left(1 + \frac{s}{D} - \cos\theta\right)\cos\theta} \right]^K . \tag{7.10}$$

Finally, noting that:

$$x = \frac{D}{2}\sin\theta, \tag{7.11}$$

the stress gradient in the no-slip region can be written as:

$$\left.\frac{d\sigma}{dx}\right|_{no\ \ slip} = \frac{K\sigma_\theta\left(2\cos\theta - 1 - \frac{s}{D}\right)\tan\theta}{\frac{D}{2}\left[\left(1 + \frac{s}{D} - \cos\theta\right)\cos\theta\right]}. \tag{7.12}$$

At the nip angle (i.e., at the transition between the slip and no-slip regions, $\theta = \alpha$), Johanson assumes that the gradient of the roll normal stress remains continuous. Therefore, the following holds:

$$\left.\frac{d\sigma}{dx}\right|_{slip} = \left.\frac{d\sigma}{dx}\right|_{no\ \ slip} . \tag{7.13}$$

The nip angle can be determined by substituting Eqs. (7.1) and (7.12) into Eq. (7.13) for $\theta = \alpha$:

$$\frac{4\left(\frac{\pi}{2} - \alpha - \nu\right)\tan\delta}{\left[\cot\left(A - \mu\right) - \cot\left(A + \mu\right)\right]} = \frac{K\left(2\cos\alpha - 1 - \frac{s}{D}\right)\tan\alpha}{\cos\alpha}. \tag{7.14}$$

Equation (7.14) may be solved iteratively or through the use of root-finding numerical schemes such as the Newton-Raphson method.

Solving for the roll normal stress distribution via Eq. (7.10) requires σ_α to be known a priori. The lack of a closed form equation to compute σ_α in Johanson (1965) leads many previous studies to rely on experimentally measured data for estimating σ_α. The normal stress at the nip angle, σ_α, can in fact be written in terms of the inlet conditions θ_o and σ_o by numerically integrating Eq. (7.2), rewritten below as a derivative of θ, to the nip angle values α and σ_α:

$$\frac{d\sigma}{d\theta} = \frac{4\sigma\cos\theta\left(\frac{\pi}{2} - \theta - \nu\right)\tan\delta}{\left[1 + \frac{s}{D} - \cos\theta\right]\left[\cot\left(A - \mu\right) - \cot\left(A + \mu\right)\right]}. \tag{7.15}$$

Johanson defines θ_o as the entry angle, i.e., the angle at which the streamwise inlet normal stress, σ_{in}, is applied (refer to Fig. 7.2). The original rolling theory assumes this angle to be a function of the material properties:

$$\theta_0 = \frac{\pi}{2} - \nu. \tag{7.16}$$

Commercially available roll compactors such as the Gerteis Mini-Pactor and the Alexanderwerk WP series, however, have their θ_o fixed by the geometry of the machine through the bore size of the auger housing (h_0 in Fig. 7.2). Therefore, θ_o may be written as:

$$\theta_0 = \cos^{-1}\left(\frac{D + s - 2h_0}{D}\right). \tag{7.17}$$

In order to solve for σ_o, Johanson assumes that there is no shear stress at the entry region θ_o. Under this assumption, the streamwise normal stress acting on the powder at the inlet, σ_{in}, becomes the major principal stress. The roll normal stress at the inlet, σ_o, can therefore be calculated from the Jenike-Shield yield criterion as:

$$\sigma_o = \frac{\sigma_{in}}{1 - \sin\delta}. \tag{7.18}$$

Assuming that the roll normal stress magnitude at the slip region is significantly smaller than that at the no-slip region, the roll force per unit roll width can then be computed via:

$$F_{roll} = \frac{WD}{2}\int_{\theta=0}^{\theta=\alpha} \sigma_\theta \cos\theta d\theta. \tag{7.19}$$

The above equation may be expressed in terms of the maximum roll pressure, σ_{max}, and the rolling angle, θ, as:

$$F_{roll} = \frac{WD\sigma_{max}}{2}\int_{\theta=0}^{\theta=\alpha}\left[\frac{s/D}{\left(1 + s/D - \cos\theta\right)\cos\theta}\right]^K \cos\theta d\theta. \tag{7.20}$$

While the Johanson model is perhaps the most widely used roll compaction model, experimental attempts to validate it remain limited. Examples include the work of Bindhumadhavan et al. (2005) and Yusof et al. (2005)

In the former investigation, the roll normal stress distributions and nip angles computed from the Johanson model were compared with experimentally measured ones. The Johanson model-predicted nip angles were found to agree to within 15% of the experimentally measured ones. Comparisons of predicted and measured roll normal stress data, however, proved inconclusive as the former quantities were noted to be highly dependent on the normal stress at the nip angle—a quantity that the authors claimed could neither be computed from the Johanson analysis nor accurately measured from experimental data. As noted above (Eq. (7.15)) and in Muliadi et al. (2012), however, the normal stress at the nip angle may actually be determined provided that the inlet geometry and streamwise stress are known.

In the latter investigation, the Johanson model-computed roll forces and torques were compared to experimentally measured ones for a gravity-fed roll compactor. Here, the stress-strain constitutive parameters were experimentally measured, and the powder/roll friction angle was determined via the method of differential slices; the inlet stress was assumed to be proportional to the powder bulk tapped density, gravitational acceleration, and powder bed height prior to roll compaction. Reasonable agreement between the measured and computed quantities was observed but only for gap values smaller than 0.15 mm. In agreement with Bindhumadhavan et al.'s findings, Yusof et al. noted that the Johanson model predictions were sensitive to the model input parameters; for example, the authors' results showed that a 10% increase in inlet stress, which was within the uncertainty of their tapped density measurements, tripled the predicted roll force values.

Other studies highlighted the shortcomings of the model's simplifying assumptions. Sommer and Hauser (2003), for example, argued that Johanson's approach of determining the nip angle by assuming a continuous stress gradient across the slip and no-slip region is inaccurate as such condition is not likely to be satisfied given the two different constitutive laws to describe the material flow therein.

Muliadi et al. (2012) performed two-dimensional finite element simulation of the roll compaction process and compared the respective results with those calculated using the Johanson model. Care was taken to ensure that the two modeling approaches used the same material properties and boundary conditions. The authors' data showed that, while the Johanson model-predicted nip angle agreed well with the FEM result, its ribbon density and stress predictions were grossly larger than the FEM model counterparts. The authors also noted that, in almost all cases, the ribbon densities predicted by the Johanson model were greater than one. They explained such an observation as follows.

First, the authors hypothesized, based upon the FEM-predicted powder velocity, that the nip process occurs gradually and is two dimensional: the powder closest to

the roll surface gets nipped first and that farther from it gets nipped further downstream. This is in agreement with previously observed experimental observations (e.g. see Ref. Muliadi et al. (2013) in Bi et al. (2014) and some of the particle image velocimetry results presented in Krok et al. (2014)) but is in contrast with Johanson's treatment of that process: the Johanson model, being one dimensional, assumes that the nip process is an abrupt change in powder velocity along the streamwise direction at the nip angle. That is, if using Fig. 7.2 notations, powder at $x > x_\alpha$ moves slower than the rolls whereas that at $x \leq x_\alpha$ moves at the same speed as the rolls. If accounting for the two-dimensional nip process, Eq. (7.5) does not reduce to Eq. (7.6), and subsequent treatments of the conservation equation overestimate the amount of powder that gets nipped. Ultimately, evaluation of the ribbon density using Eq. (7.9) will result in an overprediction.

Building on the two-dimensional nip process hypothesis, recent work proposed modifying the Johanson model by introducing a multiplication factor to the right-hand side of Eq. (7.8) to correct the amount of powder that gets nipped in the no-slip region. Bi et al. (2014) used a constant mass correction factor and showed that that quantity could be calibrated using at-scale experimental data and that it was material specific (i.e., different correction factors were derived for different materials). Liu and Wassgren (2016) performed a similar study, but unlike Bi et al. (2014), these authors assumed that the mass correction factor varied with the rolling angle. They then calibrated it using results from a two-dimensional finite element model in Muliadi et al. (2012) Good agreements between the modified Johanson model in Liu and Wassgren (2016) and the finite element model results were reported.

Other studies took a more experimental approach to improving the Johanson model. Reynolds et al. (2010) developed a novel series of equations to include feed auger parameters in the Johanson model and derived a closed form solution for ribbon solid fraction as a function of the powder properties (e.g., true density, internal effective friction angle, roll/powder frictional angle, compressibility constant), roll dimensions, roll force, auger rotational speed, minimum gap width, and powder density at the inlet (γ_α in Eq. (7.9), referred to as powder pre-consolidation density in Reynolds et al. (2010)). The authors then demonstrated that accurate ribbon density predictions could be made by treating the powder compressibility—K in Eq. (7.7)—and inlet powder density as fitting parameters calibrated via at-scale experimental data. Interestingly, they also noted that the compressibility constant derived from uniaxial slugging experiments was different than that derived through at-scale experiments.

Nesarikar et al. (2012a) fitted three load cells inside the top roll of an Alexanderwerk WP-120 roll compactor, placing them side-by-side along the roll width at equidistant spacing, and used the spatially resolved stress data to improve the Johanson model. Specifically, the authors used such data to determine the nip angle and roll normal stress distribution along the ribbon width, both of which were used for a more accurate determination of the roll force via the Johanson model. The former quantity was evaluated by first fitting a straight line to the loading and

unloading curves of the stress vs rolling angle plot. The angle between the two lines were then taken as the nip angle. The latter quantity was borne out from the authors' observation that the load cell placed in the middle of the roll measured higher stress than that measured by the other two cells. The stress data from the three cells were then used to assume a stress profile that resembled the top half of a stop sign (i.e., an isosceles trapezoid). Subsequently, the roll force was calculated by adding second integral to Eq. (7.19) to account for the stress distribution along the roll width. The nip angle term to solve Eq. (7.19) was taken as that derived from the load cell data. The stress profile was then optimized such that the predicted roll force agreed with the experimentally measured counterpart for a roll-compacted placebo formulation. Lastly, the predicted force was correlated to ribbon density using statistical analysis. The resulting optimized stress profile and statistical correlation for predicting ribbon density from the roll force were used to predict the roll compaction of API-containing blends. Good agreement between model result and experimental data was reported.

Particularly worth noting is that, despite its shortcomings, the success of the optimized Johanson models described above in predicting experimental trends suggests that the model fundamentals largely capture the physics of the roll compaction process. Hence, by examining the parameters involved in Johanson model, one can appreciate that the roll compaction process and the quality of the resulting ribbon are profoundly affected by a full spectrum of factors covering machine, material, and processes. Specifically, the following aspects are considered important for achieving a desirable roll compaction run:

1. Material considerations (i.e., powder properties): powder density; powder flow properties, including internal friction angle and wall friction angle; and powder compressibility.
2. Equipment considerations: roll diameter and width, surface finish, and surface pattern of rolls.
3. Process parameters: roll force, feed auger settings, and minimum roll gap width.

The above list is by no means a thorough collection of aspects leading to successful roll compaction. It is only the result of an attempt to put Johanson's rolling theory into practice. Many other factors, such as the feeding system, the side seal, or the milling system, should also be considered when designing a roll compaction process.

The technology in roll compaction has advanced significantly over the past decade. Many modern roll compactors are equipped with adequate instrumentations and proportional integral derivative (PID) control loop feedback systems that enable precise control of the roll compaction process. As a consequence, what used to be the process variables lacking full control, such as roll force or minimum roll gap width, can now be considered as well-controlled variables and display minimum variability throughout the roll compaction run. These advancements made it possible for formulators to move beyond empirical development and implement a more mechanistic-based approach for designing a roll compaction process.

7.3 Materials Considerations for Roll Compaction

As mentioned previously, the ultimate goal of roll compaction is the densification and particle size enlargement of a powder formulation containing the drug substance. However, often, the drug substance properties relevant to roll compaction (such as powder flowability and compaction properties) are variables largely outside the control of formulators. As a workaround, pharmaceutical excipients are blended with the drug substance to impart the formulation with desirable attributes for roll compaction.

Similar to tablet compression, the excipients for roll compaction usually encompass fillers and lubricants. Disintegrants are added to facilitate drug product disintegration for immediate release formulations. In some cases, binders or glidants are also included to improve the performance for the roll compaction process or the quality of the resulting ribbon and granules. In general, development efforts are needed to optimize the type and quantity of the excipients, so that for a given drug substance, a successful roll compaction, pertaining to the following aspects, can be achieved:

1. The powder blend exhibits acceptable density and flowability to allow for unobstructed and consistent supply of powder into the slip region and nip region.
2. Effective particle size enlargement and densification can take place under the pressure range imparted by the roll compaction.
3. The compaction pressure does not adversely impact the performance of final drug product, such as dissolution slow down or major decrease in tablet mechanical strength.

7.3.1 Filler/Excipients

The choice of fillers for roll compaction is not significantly different from the tablet compression process. The industrial norm is a combination of “ductile” fillers with “brittle” ones. It was shown, in practice, that a balanced plasticity/brittleness ratio of the blend is appropriate for roll compaction. Microcrystalline cellulose (MCC), a ductile filler, is an indispensable part of nearly all roll compaction formulations, thanks to its superior compaction characteristics, particularly the high tensile strength of the compact achievable under moderate compression pressures. Common brittle excipients include lactose and dibasic calcium phosphate (DCP). In some instances, mannitol is also employed as a replacement for lactose due to the susceptibility of lactose (a reducing sugar) to Millard reaction with drug molecules containing amine groups.

Although the choice of “brittle” fillers may vary depending on the preference of formulators, MCC is almost always the most advantageous “ductile” filler for roll

compaction, due to its ease to undergo plastic deformation at relatively low compression forces and the strong, hydrogen-bond-assisted affinity between plastically deformed particles. Higher proportion of MCC usually translates to ribbons with higher tensile strength at a given solid fraction, as well as the lower fine fractions of the resulting granules (Iyer et al. 2014; Kushner et al. 2011). Since MCC is present in a vast majority of the roll compaction formulations, understanding its limitations is important. Generally, MCC possesses typical disadvantages for materials subject to plastic deformation under pressure, which include:

1. Lubrication sensitivity
   MCC is known to be lubrication sensitive (Otsuka et al. 2004). Over-lubrication could take place when excessive amount of lubricant is used or the blend is mixed with lubricant for a prolonged period of time. Over-lubrication of MCC could lead to the loss of ribbon tensile strength or the increase of fine fraction of the resulting granules.
2. Particle size effect
   MCC particles are not known to fracture under compressive stress and thereby do not generate new surfaces through fragmentation. Therefore, coarser MCC particles, having smaller specific surface area, usually lead to softer ribbons. Nevertheless, when selecting the MCC particle size, one needs to find a balance between the powder flow and compaction properties. The fine grades of MCC, such as Avicel® PH 105 (nominal D_{50} of 20 μm), are not frequently used for roll compaction due to poor flowability. The superior particle-particle affinity of MCC enables adequate densification and size enlargement even with relatively coarse grades of MCC. In practice, MCC grades with medium particle size (D_{50} between 50 to 100 μm) are most common choices for roll compaction.
3. Work hardening
   MCC is characterized by plastic deformation, which, from the supramolecular perspective, results from the dislocation and movement of molecules along slip planes. The pressure endured during the roll compaction can reduce the ability of the material to deform following these mechanisms upon the second compression during tableting, i.e., the material is work hardened. Therefore, when a large portion of MCC is present, the tablet tensile strength is reduced following the roll compaction, as compared with the non-roll-compacted formulation having the same composition and particle size (Mosig and Kleinebudde 2015; Sun and Kleinebudde 2016). The extent of work hardening is particularly pronounced in MCC, which predominantly undergoes plastic deformation. Due to this limitation, studies are usually recommended for roll compaction to find the optimal roll force that strikes a balance between the particle size and density of granules and the mechanical strength of the final tablets.

The most popular “brittle” fillers are lactose and DCP, both of which are known to undergo fragmentation under stress. In fact, there are appreciable differences between the two excipients, with respect to both compressibility and compactabil-

ity. DCP, being an inorganic excipient, exhibits substantially stronger resistance to densification and can hardly be densified to a solid fraction above 0.70 under a typical compression pressure for tableting (LaMarche et al. 2014), whereas lactose can easily be consolidated to a solid fraction above 0.85. On the other hand, DCP possesses higher compactability (higher tensile strength at a given solid fraction) than lactose, which offsets its lower compressibility. DCP also has higher bulk density than lactose. Both materials are frequently used to blend with MCC to impart the desired powder properties for roll compaction.

The "brittle" excipients undergo particle fracture upon compression, thereby generating new surfaces. For this reason, lactose and DCP do not sustain the same weaknesses as "ductile" materials such as MCC. They are less sensitive to overlubrication and work hardening effects (Souihi et al. 2013a). For lactose and DCP, the trend that small particle size leads to stronger ribbon and coarser granule size remains, but to a less extent.

While DCP is largely a crystalline solid regardless of the manufacturing and processing method, for lactose, significant surface amorphization can be created through the spray-drying process (Vromans et al. 1987). Studies showed that the spray-dried lactose gives rise to ribbon with higher tensile strength and granules with lower fine fraction. The ability to render partial amorphization to lactose enables the use of coarse particles because the presence of surface amorphous portion compensates for the low specific surface area, as compared with finer, fully crystalline lactose (Omar et al. 2015). There is little incentive to use fine grades of DCP or lactose for roll compaction because the benefit of mechanical strength improvement does not outweigh the detrimental effect on flowability. The un-milled form of DCP and agglomerated or spray-dried form of lactose are the most common types of "brittle" fillers for the roll compaction process.

7.3.2 Lubricants

As an empirical practice, lubricant is frequently added to the formulation to facilitate a smooth roll compaction process. Lubrication generally enhances flow and transmission of powder inside the roll compactor and reduces the tendency of powder to stick to the surface of the roll. The types of lubricants used for roll compaction are identical to those for the tablet compression process. They generally belong to boundary lubricants, i.e., amphiphilic materials with strong tendency to form film on the surface of the particles to reduce powder-wall friction (Miller and York 1988). There are three main types of lubricants for roll compaction: (1) metallic salts of fatty acid (such as magnesium stearate), (2) fatty acids (such as stearic acid), and (3) fatty acid esters (such as sodium stearyl fumarate) (Wang et al. 2010). Among them, magnesium stearate is the most common due to its superior lubrication efficiency compared with other alternative candidates.

The main effect of lubricants is to reduce the powder-wall friction. It was also shown to lower the internal friction for certain pharmaceutical powders. Mechanistically, lubrication was reported to alter the roll compaction process by the following means: (Dawes et al. 2012; Miguelez-Moran et al. 2008, 2009; Yu et al. 2013).

1. Reducing powder-wall friction in the feeding zone and enabling more efficient and homogeneous powder feeding into the slip zone.
2. Reducing powder-wall friction on the roll surface, thereby decreasing the nip angle.
3. Altering maximum pressure, σ_{max}, which is achieved at $\theta = 0$ in Eq. (7.10), as a result of smaller nip region and/or improvement of powder flowability through the roll gap.

These studies demonstrate that the extent of lubrication not only affects the properties of the bulk powder but also influences the stress gradient of the powder during roll compaction. Specifically, an identical formulation with different level of lubrication could potentially be subject to different compression profiles, even when the roll compaction parameters (roll force, gap, roll speed, etc.) are set equal.

It is also worth noting that many of existing lubrication studies were conducted using roll compactors with a fixed gap width (i.e., roll force or hydraulic pressure was not set at a constant value). For modern commercial roll compactors enabling both force and gap control, the results may be different. However, no report has been identified to date to assess and validate the effect of lubrication on such equipment.

Similar to tablet compression, over-lubrication during roll compaction can have deleterious effect on the resulting ribbon and granules. A lubricated ribbon can exhibit lower tensile strength and higher fine fractions upon milling, as compared to the non-lubricated ribbon with similar solid fraction. Over-lubrication during roll compaction can also lead to decrease of mechanical strength of the resulting tablets (Yu et al. 2013; He et al. 2007). Therefore, excessive lubrication is not recommended if the powder has consistent flow in the roll compactor with minimal powder sticking to roll surface.

7.4 Equipment Considerations for Roll Compaction

Relying on mechanistic understanding of the roll compaction process through the use of available models alone may result in an overly simplistic treatment of it; there are inevitably factors that are absent from a model, either through its simplifying assumptions or simply because they cannot be accurately modeled with the current state of the art. Examples include issues associated with feeding powders using the augers, powder leakage through the roll compactor side seals, roll speed effects (note that FEM models often assume a quasi-static process and the Johanson model

in its original form does not have velocity terms in any of its equations), the presence of interstitial air in the powder during compression, and the orientation of the feed augers and the rolls (vertical, horizontal, inclined). In the paragraphs that follow, we discuss how these factors affect the outcome of the process and highlight how different roll compactor designs invariably affect them.

7.4.1 Feed Design and Mechanism

The feed system plays a critical role in ensuring a steady stream of powder is made available into the rolls. Intermittent feeding could lead to generation of excess fines, increased loss through uncompacted powder, and poor-quality granules (Miller 1997). In roll compactors with PID-controlled gap width, it may also lead to difficulties meeting the user set gap width, which in turn could lead to inadvertent process shutdown.

Two types of feeders are used for conveying powder blends in roll compactors: gravity driven and force feeding through the use of auger(s). In the case of dense, free-flowing powders, gravity may provide sufficient continuous feeding. However, in pharmaceutical applications, the intragranular blends often possess low density and exhibit poor flow properties. Here, the use of feed augers and arch breakers at the hopper outlet is often required. Feed augers also help predensify the powder prior to it entering the roll compactor inlet. This further ensures a more consistent nip process and more uniform ribbon density as well as increased throughput (Miller 1997).

The orientation of the feed auger is also an important factor. Vertically oriented augers (i.e., augers that feed powder vertically—see Fig. 7.3a), for instance, are often installed directly above the roll assembly. Such a design, while taking

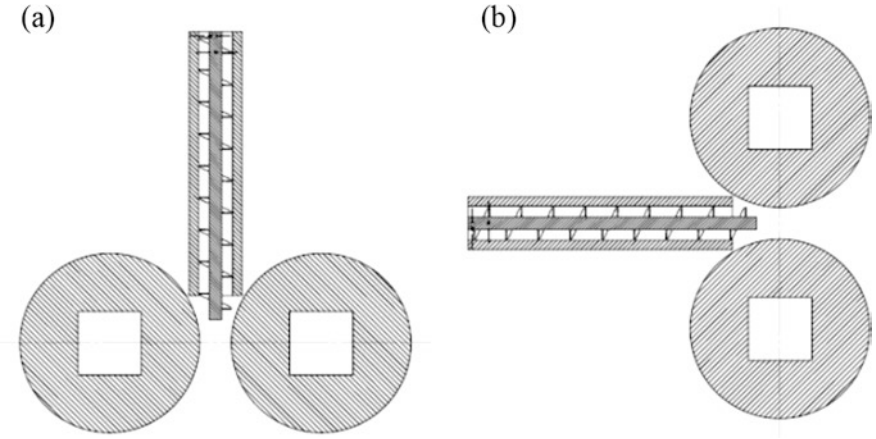


Fig. 7.3 Schematic of (**a**) vertical and (**b**) horizontal feed auger orientations

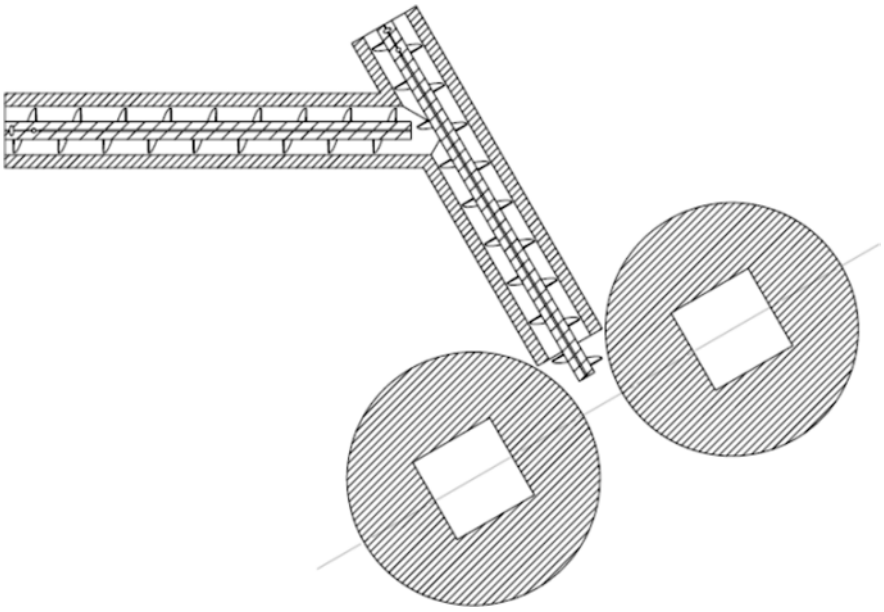


Fig. 7.4 Schematic of dual-feed auger orientations similar to that used in a Gerteis Mini-Pactor or Macro-Pactor roll compactor

advantage of the gravity to aid with powder flow from the hopper and in between auger flights to ensure a more continuous powder feed to the rolls, may be prone to the risk of powder passing through the roll gap width uncompressed. Orienting the augers horizontally (see Fig. 7.3b) below the hopper obviously will abate such risks. Miller (1997) noted that the use of two feed augers, one oriented horizontally and the other vertically, combined the advantages of the two designs and provided more consistent feed of the powder into the rolls over a single-screw configuration.

It is worth noting that the Gerteis Mini-Pactor and Macro-Pactor are examples of roll compactors that employ two differently oriented augers. Its horizontal auger ("feed auger") conveys powder from the hopper to an inclined, second auger ("tamp auger"), which deaerates and pre-compresses the powder as it is force fed into the roll inlet zone (see Fig. 7.4).

7.4.2 Deaeration Mechanisms

The presence of interstitial air in the powder is hypothesized to contribute to uneven feed, increased leakage of uncompressed powder, decreased throughput, and heterogeneous ribbon density distribution. Miller (1997) proposed that this is the result of powder fluidization that occurs as air pockets within the powder escape and fluidize the powder as the powder is compressed.

Different roll compactors achieve powder deaeration differently. The Gerteis Mini-Pactor, for example, uses two feed augers to maximize powder pre-compression and thus deaeration, prior to it entering the roll inlet zone. Alexanderwerk roll compactor models, on the other hand, do so through their Combi-Vent feed design. Here, a sintered stainless steel filter attached to a self-contained vacuum pump is installed just upstream of the roll inlet zone. Running the vacuum pump then forcibly pulls air pockets within the powder as it passes through the filter. Another component of the Combi-Vent design is a dedicated chamber in the powder hopper. This chamber acts as an outlet for air to escape from and thus not disrupt the flow of powder from the hopper, when the powder is predensified by the feed auger action.

The performance of a roll compactor equipped with a vacuum deaerator was experimentally investigated by Miller (1997). He compared the process throughput and quantified the amount of uncompressed powder leakage for processes run on a roll compactor equipped with a single-feed auger with and without vacuum deaeration. It was found that running the process with vacuum deaeration resulted in up to 30 kg/hr. increased throughput and reduced uncompressed powder leakage by up to 28% than running it without. The benefits of vacuum deaeration were not observed to the same extent from a roll compactor system equipped with two feed augers, wherein vacuum deaerator was installed throughout the length of the second feed auger (the one just upstream of the rolls). The author noted that this was likely because significant powder predensification resulted from the feed augers actions and that the vacuum deaerator could no longer effectively pull interstitial air from the already largely deaerated powder.

It is important to note that the analytical and computational approaches described in this chapter deliberately ignore the presence of interstitial air in the powder. As such, these models will fail to predict the effects of vacuum deaerator on the outcomes of the process.

7.4.3 Roll Orientation Effects

Major differences in roll orientation can be observed between different roll compactor models. Roll orientations can be categorized according to the angle the roll pair axes make relative to the horizontal plane. Fitzpatrick and Bohle roll compactors have their roll pair axes aligned with the horizontal plane (see Fig. 7.3a). For these compactors, the feed auger is oriented perpendicular to the roll pair axes and feeds powder vertically toward the rolls. As mentioned before, such a design may suffer from an increased risk of powder passing through the roll gap width uncompressed due to the effects of gravitational forces.

Alexanderwerk and Gerteis roll compactors provide a solution to such risk by orienting the roll axes differently. The former compactor roll axes are aligned to the vertical plane (see Fig. 7.3b), whereas the latter compactor roll axes sit at a 30° angle relative to the horizontal plane (Fig. 7.4). Such roll orientations tend to minimize powder leakage as the powder conveyance toward the rolls is predominantly

governed only by the screw feeder (Kleinebudde 2004); they have been shown to reduce uncompressed powder pass-through the roll gap width by 10% and thus are also the preferred configurations for low-dose products (am Ende et al. 2007). Regardless, M.T. am Ende et al. (2007) have shown that the pass-through effect could be minimized through optimization of processing parameters, for example, by favoring roll compactor settings that lead to significant powder predensification in the feed auger or by using vacuum deaerator.

7.4.4 Roll Control and Surface Types

Modern roll compactors normally have one of their rolls fixed and the other movable. The movable roll is driven by a hydraulic system. These compactors operate in its PID roll gap width control set to either "on" or "off." When it is set to "on," the user sets a value for the roll gap width—typically between 1 mm to 6 mm—and the roll force (e.g., Gerteis) or roll pressure (e.g., Alexanderwerk). The machine PID controller then varies the amount of powder fed to the rolls until the user set point values for roll gap width and force or pressure are met. When set to "off," the user sets a value for the roll force or roll pressure and the feed auger speed. The roll gap is then let "floating" and may take any value at steady state. With such a design, there exist at least two independent variables controlling the process at any given time: the roll force or roll pressure, which sets the ribbon density (Muliadi et al. 2013), and the throughput, which is set by a combination of roll speed and either the roll gap width or the feed auger speed.

A variety of roll surface finishes ranging from smooth to textured are typically available. Different roll surface types correspond to different the roll/powder friction angle, which directly affects the nip angle—see Eq. (7.14). A rougher pair of rolls will increase the nip angle; its effect on the process, however, is complex and dependent upon the mode of operation.

When the roll gap control is set to "on," the feed auger will need to reduce the amount of material fed into the roll inlet region in order to keep the same roll gap, force, and rotational speed as the roll surface becomes increasingly rougher. Such a trend may be predicted through the Johanson model. Figure 7.5a for instance, shows the effects of increasing the powder/roll friction coefficient on the nip angle, whereas Fig. 7.5b shows how the streamwise inlet stress, σ_{in} in Eq. (7.18), must vary in order to keep the roll gap and roll force used unchanged. The decreasing σ_{in} following an increase in the nip angle in Fig. 7.4 may be achieved through a decrease in the amount of fed material and/or the degree of material predensification at the roll inlet region, both of which may result from a reduction in the feed auger rotational speed.

On the other hand, when operating with the roll gap control set to "off," increasing the roll surface roughness while keeping the same roll force, roll speed, and

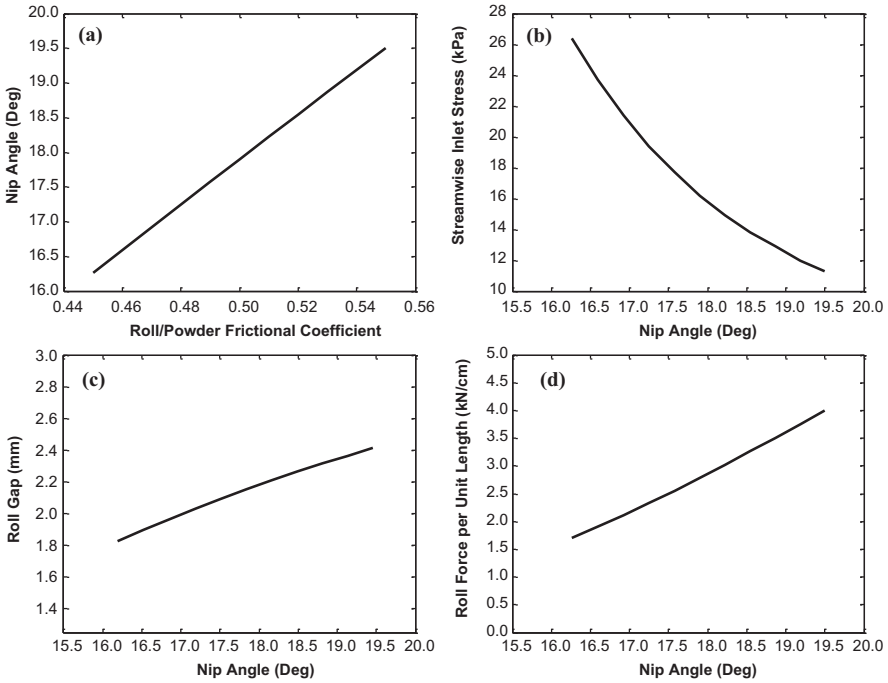


Fig. 7.5 The Johanson model-predicted (**a**) nip angles for different roll/powder friction angles with gap control set to "on," (**b**) streamwise inlet stresses required to maintain the same roll gap and roll force for different nip angles with gap control set to "on," (**c**) roll gap required to maintain the same roll force at a fixed streamwise inlet stress for different nip angles with gap control set to "off," and (**d**) roll force required to keep a constant roll gap at a fixed streamwise inlet stress for different nip angles with gap control set to "off." For all cases, the material compressibility constant and effective internal frictional angle are 5° and 50°, respectively. Roll compactor dimensions are based on those of an Alexanderwerk WP-120, and roll/powder frictional coefficient varies from 0.45 to 0.55

feeding rate will increase roll gap. Once again the Johanson model may be used to predict how roll gap must vary to compensate the outcome of making the roll surface rougher (i.e., the increase in nip angle—see Fig. 7.5c). Accordingly, in order to maintain the same roll gap under a fixed feeding rate and roll speed but increasing roll surface roughness, the roll force must increase (Fig. 7.5d). Of the three hypothetical scenarios outlined above, only the last one will result in an increase in ribbon density. The first two scenarios likely will have little impact on ribbon density.

Lastly, in practice, a smooth roll surface is preferred when roll compacting powders that have a tendency to stick. On the contrary, surfaces with increased roughness (e.g., corrugated, serrated, or knurled) and deeper pockets are suitable for powders with low bulk density.

7.4.5 Side Seal Design

Most roll compactors are equipped with some form of seal assembly to prevent uncompressed powder from leaking through the sides of the compactor. There are generally two types of side seals: stationary ones ("cheekplates") and ones that are a part of one of the rolls and rotate with it ("rimmed roll").

Alexanderwerk compactors typically come with stationary cheekplates. These cheekplates are made of Teflon™ to minimize friction thereat. Regardless and as explained previously, the frictional drag between the powder and the cheekplates minimizes accumulation of powder at the edges and as a result yields lower ribbon density thereat (Muliadi et al. 2013; Miguelez-Moran et al. 2009).

The rimmed roll seals are available for Gerteis and Bohle compactors. Here, the rimless roll fits into the rimmed roll. As explained previously, rimmed roll compactors produce ribbons that are denser at their edges than in the middle (Mazor et al. 2016).

Perez-Gandarillas et al. (2016) investigated the effects of side seal designs on fine fraction and the tensile strength of the tablets made from the milled ribbons. The authors' results showed that irrespective of the seal designs and the corresponding density distribution in the ribbon, approximately the same average ribbon density could be achieved with both the cheekplate-equipped compactor and the rimmed roll-equipped one. However, they noted that the use of rimmed roll translated to less fine fraction and slightly stronger tablets, especially at high compression forces.

Other studies highlight the additional flexibility that the rimmed roll seal design could provide in minimizing the ribbon density distribution. Miller (1997), for example, studied the effect of the rim angle—the angle that the inner rim surface makes with the roll surface—on the amount of uncompressed powder leakages. The authors found that leakage was minimized when the rim angle was 65°. They also claimed that the rimmed roll setup generally provided better seals than the cheekplate equivalent, noting 3% total leakage vs 15% when, respectively, the rimmed roll and the cheekplate were used.

7.4.6 Roll Compactor Selection

There are several commercially available roll compactor designs developed and marketed for pharmaceutical applications. Csordas and Kleinebudde (2017) compared the performance of three most commonly used lab-scale roll compactors: Alexanderwerk BT-120, Bohle BRC-25, and Gerteis Mini-Pactor. The purpose of the study was to examine the deviation from the set point values for roll force and gap width for each of the compactors. Different grades of MCC, mannitol, and a 1:1 mix of MCC and mannitol were roll compacted. The differences between the three compaction units and their operational parameters are summarized in Table 7.1.

Table 7.1 Roll compactor settings used in Csordas and Kleinebudde (2017)

	Alexanderwerk BT-120	Bohle BRC-25	Gerteis Mini-Pactor
Roll surface finish used	Knurled and smooth	Knurled, coarse grooved, and fine grooved	Knurled
Roll dimensions	120 mm diameter, 40 mm wide	250 mm diameter, 25 mm wide	250 mm diameter, 25 mm wide
Feed auger	One horizontal feed auger	One horizontal feed auger and one vertical tamp auger	One horizontal feed auger and one inclined tamp auger
Feed to tamp auger speed ratio	N/A	1:4	1:2
Compaction force adjustments	Hydraulic system	Spindle motor, electromechanical roll drive	Hydraulic system
Deaeration mechanism	Vacuum deaeration at the roll inlet zone	N/A	N/A
Roll axes orientation	Vertical	Horizontal	30° from the horizontal plane
Side seals	Cheekplates assembly	Rimmed roll assembly	Cheekplate assembly
Automation	N/A (“floating” roll gap)	PID gap control	PID gap control
Roll force or roll pressure	18, 24, 36, 48, and 60 bars	2 kN/cm, 17 kN/cm	2 kN/cm, 20 kN/cm
Roll speed	1–15 rpm	2 RPM	2 RPM
Roll gap	1.5, 1.9, and 2.3 mm when using smooth rolls; 1.5, 2.3, and 3.0 mm when using knurled rolls	1.50, 2.25, and 3.00 mm	1.50, 2.25, and 3.00 mm

The authors found that the Gerteis Mini-Pactor offered the best overall performance, with its roll force and gap width fluctuating the least from the corresponding set point values. In addition, the Gerteis Mini-Pactor was also able to adapt quickly to parameter changes during a given run—that is, it was able to meet new roll force and gap width set points the quickest. The second best overall roll compactor was the Bohle BRC-25. Large deviations between set point process parameters and actual values were observed from roll compaction runs performed using the Alexanderwerk BT-120. The authors attributed such observation from its lack of PID scheme that accurately controls the gap.

7.5 Roll Compaction Process Parameters

In the paragraphs that follow, we separately review the user-adjustable parameters commonly available for modern roll compactors possessing a variable gap feature. Specifically, the effects of roll force or roll pressure, roll speed, feed auger speed, and mill parameters on the attributes of the intermediates (i.e., ribbons and granules) and final tablets are discussed.

7.5.1 Effects of Roll Force or Pressure

For modern roll compactors with a variable gap feature, the pressure or force acting on the movable roll is one of the adjustable process parameters. An increase in roll force or pressure invariably increases the average ribbon density if all other process parameters are kept the same (Souihi et al. 2013b). Specifically, for floating gap roll compactors, when operating with the roll gap control "on," an increase in roll force or pressure while keeping the roll gap unchanged will cause the feed auger to feed more powder toward the gap, ultimately resulting an increase in the average ribbon density. When operating with the roll gap control "off," increasing roll force or pressure while maintaining the same feed auger speed will result in the roll gap becoming smaller, thereby increasing the ribbon density.

For fixed gap roll compactors, factors leading to an increase in roll force or roll pressure, assuming a feeding rate that ensures the same amount of material is delivered to the roll inlet region and/or that the same level of powder predensification is achieved thereat, include an increase in the powder internal frictional angle and roll/powder frictional angle—both of which increased nip angle—and roll compacting less compressible materials.

Experimentally, the effects of roll force on the characteristics of roll-compacted ribbons, the subsequently milled granules, and the tablets that they make were studied by Souihi et al. (2013b), Perez-Gandarillas et al. (2016), Inghelbrecht and Paul Remon (1998a), and Sakwanichol et al. (2012). Souihi et al. (2013b) noted that an increase in roll force led to an increase in the average size (D_{50}) of the granules and a narrower granule size distribution, as well as an increase in the granule bulk density. The authors did not report a reduction in tablet tensile strength ("loss in tabletability") following an increase in roll force and attributed such an observation to the fact that their blend comprised 55% mannitol, a brittle filler.

Perez-Gandarillas et al. (2016) reported that an increase in roll force generally resulted in a reduction in fine fractions (higher D_{10}) and denser granules. These observations are in agreement with those reported by Sakwanichol et al. (2012). Perez-Gandarillas et al. (2016) also compared the tensile strength of the tablets made from the milled granules and that made from the non-roll-compacted powder. They found that for both mannitol and MCC, the tensile strength of the tablets made from the milled granules was less than that of the non-roll-compacted powder.

However, the tensile strength of the MCC tablets were found to decrease with increasing roll force, whereas the mannitol tablets were not. These results are consistent with other work in the literature (Wu and Sun 2007; Sun and Himmelspach 2006).

Inghelbrecht and Paul Remon (1998a) compared the effects of roll force, roll speed, and feed auger speed on the milled granule size and friability. They found that the roll force was the most dominant factor in influencing these characteristics, followed by the roll speed and the feed auger speed. They further noted that the largest and strongest granules were produced by maximizing the roll force and minimizing the feed auger speed.

It is important to note that, since roll compaction is an intermediate granulation step, the target density for the ribbon is typically lower than that of the final tablet density. Overcompression of the ribbons due to overly high roll force or pressure may translate to excessive work hardening, especially if the intragranular blend comprises primarily a ductile filler such as MCC (see "fillers" section above), and is thus detrimental to the final tablet tensile strength and friability. When an intragranular blend contains both ductile and brittle fillers, some development work may be required to determine the roll force or pressure that optimizes both the favorable granule characteristics and the final tablet properties.

7.5.2 Effects of Roll Speed

For variable gap roll compactors, the choice of roll gap and roll speed determines the overall throughput of the process. The practical outcomes of changes in roll speed are as follows. When operating with the PID gap control set to "on," increasing roll speed while maintaining the same roll gap and roll force will result in an automated increase in feed auger speed. When operating in this mode, excessively high roll speeds and/or large roll gaps may lead to feeding issues. For example, if the powder is poor flowing, it may "plug" in the feed auger housing as a result of the feeding action (Gupta et al. 2005). The risk of uncompressed powder leakage from the roll compactor side seals typically rises when the powder is overfed as well.

When the PID gap control is "off," doing so while keeping the feed auger speed and roll force unchanged will lead to a decrease in roll gap. When the ratio of feed auger speed to roll speed is too small, insufficient feeding may occur, resulting in overly thin and friable ribbons (Gupta et al. 2005; Teng et al. 2009). It has been reported that when the feed auger speed to roll speed ratio is kept constant, the ribbon properties become independent on either roll speed or feed auger speed alone (Teng et al. 2009).

In addition to throughput, roll speed also controls the compression dwell time. The impact of dwell time on the properties of the roll-compacted products can be pronounced for plastically deforming ("ductile") materials such as MCC. For these materials, longer dwell times, which correspond to low roll speeds, may result in higher friability and more severe loss in tabletability of the resulting granules

(Inghelbrecht and Paul Remon 1998a). Materials known to have high elastic recovery such as pregelatinized starch are also sensitive to the dwell time effect. For example, as dwell time decreases following an increase in roll speed, the post-compaction elastic recovery may occur to a greater extent, an effect that Teng et al. (2009) have attributed to production of more friable ribbons.

For brittle materials, the effects of dwell time/roll speed generally do not affect the ribbon strength significantly (Wu and Sun 2007). This is likely because material fracture that is prerequisite to densification occurs at a much faster rate than typical dwell time for a roll compaction process. For blends containing predominantly brittle material, Souihi et al. (2013b) also reported an increase in the bulk density of the milled granules and improved tensile strength of tablets made from the granules following a decrease in roll speed. Similar experimental trends were reported by Inghelbrecht and Paul Remon (1998a).

In continuous manufacturing applications, the roll compactor throughput is generally matched to the overall production "line rate." As such, the roll speed may not be as easily adjustable as it would have been for a stand-alone process, wherein it is possible to optimize the roll speed based on the material attributes and select the maximum roll speed that provides the desired throughput without compromising the quality of roll-compacted products.

Lastly, it is important to note that the effects of roll speed are not captured by the Johanson model in its original form, as indicated by the absence of any velocity terms in any of the model equations (see Eqs. (7.1) through (7.20)). Modified Johanson model such as that proposed by Reynolds et al. (2010) does, however, account for roll speed as the roll and feed auger speeds appear as variables in the proposed equations. Additionally, Reynolds et al.'s approach of using roll-compacted ribbon density produced at different pressures for deriving the powder compressibility constant, K, may be used to indirectly include roll speed effects in the model as well, provided that the data capture post-compression elastic recovery that may result and how it then varies with roll speed.

Finite element models presented above treat the process as quasi-static and thus do not capture roll speed effects. Any velocity boundary conditions in the model are used as a mass-scaling means—a method for increasing simulation time steps and consequently achieving convergence faster, by artificially increasing the mass of the elements in the computational domain.

7.5.3 Effects of Feed Auger Speed

For modern roll compactors with variable roll gap, the feed auger speed is only user-adjustable when operating with the PID gap control mode set to "off," wherein a combination of the feed auger speed, roll force or pressure, and roll speed will determine the value of the gap width. Muliadi et al. (2013) result suggests that increasing the feed auger speed while keeping the roll force and roll speed unchanged increases ribbon thickness but not necessarily average ribbon density. A summary

of additional experimental observations related to the effects of feed screw speed on the roll compaction process has been provided in the preceding sections.

For roll compactors that are equipped with two feeders (e.g., the Gerteis Mini-Pactor, Bohle BRC-25), the process throughput is generally determined by the feed auger directly downstream of the hopper (Teng et al. 2009). For these compactors, the ratio between the first and second feed auger speeds is typically also a user-adjustable setting (Teng et al. 2009).

Feed auger speed effects appear indirectly in the original Johanson (1965) model equations as the inlet streamwise normal stress and the inlet density. It is, however, important to note that those quantities cannot be reliably measured or predicted at present. As mentioned previously, feed auger speed terms indeed appear in the equations for Reynolds et al.'s modified Johanson model (2010).

7.5.4 Effects of Milling Parameters

Milling parameters generally include the mill speed, rotational direction (clockwise, counterclockwise, or both), oscillating angle, as well as the screen size and the rotor types. For the Gerteis Mini-Pactor, for example, different rotor types such as an open star rotor and a pocket mold-grooved rotor are available. The former rotor type performs milling predominantly by abrasive fracture (i.e., shearing of ribbon pieces against the screen), whereas the latter type does so by a combination of abrasive fracture and impact fracture (i.e., shattering of ribbon pieces due to impact with the rotor blades) (Yu 2012).

The effects of mill types and milling parameters were studied by Samanta et al. (2012), Hancock and Vendola (2008), Perez-Gandarillas et al. (2016), and Sakwanichol et al. (2012) These studies are generally in agreement with one another, and their results can be summarized as follows:

- A decrease in average granule size, a broader granule size distribution, greater fine fraction, and a reduction in granule flowability were observed following a decrease in screen opening size (Sakwanichol et al. 2012).
- An increase in oscillating angle and mill speed results in a decrease in average granule size (Sakwanichol et al. 2012).
- Open star rotor typically yields smaller average granules and higher fine fractions than the pocket mold-grooved rotor (Sakwanichol et al. 2012). However, Perez-Gandarillas et al. (2016) reported that the opposite was true when milling MCC ribbons produced at high compression force (8 kN/cm).
- Higher throughput was achieved when milling was performed with the open star rotor than when the pocket mold-grooved rotor was used, when greater oscillating angles resulted were used, and when bidirectional rotation (clockwise and counterclockwise) of the rotor was implemented (Perez-Gandarillas et al. 2016).

- Mill parameters and rotor types do not significantly affect the tensile strength of the tablets made from the milled granules (Perez-Gandarillas et al. 2016; Hancock and Vendola 2008).

7.6 Process Development Strategies

In practice, roll compaction process development strategies often involve either selection of parameters based on prior knowledge and experience and trial-and-error studies or establishing a design space through comprehensive factorial experiments (i.e., design of experiment (DoE) studies) (Souihi et al. 2013b; Tezyk et al. 2017). These approaches vary in the amount of resources required and are often selected based on the clinical stage of the product.

Due to limited amount of API, early-stage studies tend to rely heavily on experience and opt for empirical approaches to generate a general understanding of how each of the process parameters affects the performance of the end product. As the development stage advances, more API becomes available, thus allowing for more systematic and comprehensive studies such as factorial experimentations to optimize the roll compaction process.

The three most commonly studied parameters in the roll compaction process are roll pressure, roll gap, and roll speed (Souihi et al. 2013b). A two-factor three-level DoE studying the effect of roll pressure and roll gap or roll speed on key measures of product performance such as tensile strength, disintegration, and friability is often sufficient as a first pass assessment of the process. For subsequent DoE studies, the optimal operational setting determined from the first pass can serve as the design "center point"; the lower- and higher-end ranges designed at the edges of failure of the process then define the "corners" of the optimal operating space.

A more systemic approach is through the use of QbD principles consistent with regulatory framework outlined in ICH guidelines: Q8(R2), pharmaceutical development; Q9, quality risk management; and Q10, pharmaceutical quality systems. Here, the critical quality attributes (CQAs) for the drug product (e.g., tablet hardness, tablet weight, friability, and dissolution rates) are first identified. Ribbon density and porosity are also often included as CQAs because they invariably affect downstream processing and the attributes of the final product. Next, the roll compaction process parameters that may affect the CQAs are identified.

Fahmy et al. (2012) provided an excellent summary of the process of combining failure mode and effects analysis (FMEA) and Plackett-Burman approach to design factorial experiments for a drug product requiring roll compaction granulation. The authors detailed that prior experience, literature survey, and pre-formulation data could be used in conjunction with FMEA principles to determine the drug product CQAs and the respective roll compactor parameters that affect them. They then applied Plackett-Burman analysis to identify the significant "main effects" among

the CQAs and design factorial experiments around these factors as a means to minimize the size of the DoE. Through this exercise, the authors identified roll pressure as the singular roll compaction parameter that significantly affects the mean granule size, granule bulk density, and Carr index. In addition, while the ratio of roll speed to feed auger speed and mill speed was also identified as potentially significant factors affecting the CQAs, the DoE results indicated that they did not significantly affect the quality target product profile (QTPP) and thus need not be scrutinized to the same degree as the roll pressure in later development stages (Fahmy et al. 2012).

7.7 Scale-Up Approaches

The purpose of the roll compaction scale-up activities is to identify the key process parameters of a production-scale roll compactor running at a higher throughput, so that the resulting ribbon and granules exhibit quality attributes similar to those obtained from a laboratory- or pilot-scale roll compactor. Unlike tablet compression process, where the stress path of the powder throughout the compaction cycle is known, such information is usually not readily available for roll compaction unless equipment with an instrument roll is used. Therefore, transfer of the roll compaction process to larger-scale equipment is not always straightforward.

7.7.1 Empirical or Semiempirical Approaches

Historically, some empirical rule of thumb was followed for roll compaction scale-up. The approach called for maintaining the roll force (expressed as force per unit roll width) and the ratio of feed screw/roll speed when transferring the process to larger equipment. A scale-up study following these practices was performed using a powder blend containing 50% drug substance. By employing laboratory-, pilot-, and production-scale roll compactors, all from the Freund-Vector line, the authors showed that the properties of the granules and the resulting tablets were similar among all equipment when the roll force and feed screw/roll speed ratio were kept constant (Sheskey et al. 2000). This is by far the most straightforward method because no additional experiments are required. However, due to the lack of full theoretical rationale, this method may be considered as a starting point for scale-up activities on equipment without adequate force and gap control.

Several scale-up approaches based upon statistical models were also proposed. For example, Shi and Sprockel suggested that two linear equations, in which the ribbon thickness and density are expressed as the function of the roll hydraulic pressure and roll gap, can be built through a design of experiment (DoE) study conducted on a pilot-scale roll compactor (Shi et al. 2013). They further argued that the same linear relationships will largely hold true regardless of the equipment size

except the y-intercepts, and only one run on the production-scale equipment is needed to correct the models. Such models can then be used to direct roll compaction at production scale. Additionally, efforts to build scale-up models using multivariate analysis were also reported. The concept is to build a model where the independent variables contain multiple material properties and process parameters and dependent variables contain ribbon or granule properties. The model is calibrated through a joint training set comprising run data obtained using both small- and large-scale roll compaction equipment. The calibrated model can then be inverted to determine the key process parameters of the large-scale equipment by entering the desired values of dependent variables (i.e., ribbon and granule properties) (Liu et al. 2011). These statistical model-based methods usually require multiple roll compaction runs, usually in a DoE style, at both small and large scale, and can become cumbersome experimentally. However, it is a step forward in reliability because these models were generated from data obtained using the same equipment intended for the scale-up production.

Another plausible approach to facilitate scale-up efforts is to use dimensionless numbers. To this end, the dimensionless quantities can either be derived under certain theoretical basis or through dimensional analysis. Rowe et al. proposed a modified Bingham number, originally derived in fluid mechanics, based on the hypothesis that the ribbon properties are dictated by the ratio of the yield stress versus the viscous stress (Rowe et al. 2013). The proposed dimensionless number entails equipment geometry (roll diameter, width, and surface area), process parameters (roll gap, feed screw and roll speed, and screw speed constant), and material properties (powder true density, and pre-consolidation relative density). Scale-up experiments using Alexanderwerk roll compactors showed that the ribbon density is scale-independent when the modified Bingham number was kept constant. Although the modified Bingham number was acquired with theoretical consideration, certain variables, such as the screw speed constant, require careful calibration experiments. The widespread use of this dimensionless number could be precluded as the result. Another, more empirical dimensionless number (*DV*) was proposed by Boersen et al. through dimensional analysis (Boersen et al. 2016). This dimensionless quantity is less intricate than the modified Bingham number and takes the following form:

$$DV = \frac{RP}{RS \cdot HFS \cdot \rho \cdot D^2}, \tag{7.21}$$

where *RP* and *RS* are roll pressure and roll speed, respectively, *HFS* is the horizontal feed screw speed, ρ is the blend true density, and *D* is the roll diameter. *DV* was shown to be successful in transferring several formulations between two roll compactors with very different design. Because all the parameters in the *DV* number are measurable, it could be easily adopted in actual scale-up practice. However, similar to other empirical methods, the *DV* value was obtained through dimensional analysis and lacks rigorous theoretical basis, and additional testing on other formulations and equipment is recommended to fully validate this approach.

7.7.2 Mechanistic Approaches

Conceptually, the roll compaction scale-up can be achieved by simply following Johanson's rolling theory. Unfortunately, scale-up endeavors using the Johanson model in its original form were rarely successful in practice due to its fundamental shortcomings discussed above. Regardless, optimizing the Johanson model by calibrating it with experimental data using the methods previously outlined has proved effective. Reynolds et al.'s method of backfitting the powder compressibility K using the experimental roll compaction data from a DoE study (Reynolds et al. 2010), for example, appeared to work well in a number of scale-up studies (Reynolds et al. 2010; Gago et al. 2018; Souihi et al. 2015). Nesarikar et al.'s (2012a) approach of using an instrumented roll to determine the nip angle and roll width-wise stress distribution to optimize the Johanson model was also used to guide process scale-up with great success (Nesarikar et al. 2012b).

In what appears to be a more plausible advancement, the concept of Johanson's rolling theory and its utility in scale-up are increasingly incorporated into the equipment design. Several roll compactor manufacturers now offer equipment line in which the pilot-scale and production-scale roll compactors exhibit almost identical geometry in the powder feed region. In particular, the roll diameter and roll surface are kept unchanged; the increase of throughput is purely achieved through an increase of roll width. This design concept ensures that the nip region is identical. As such, the parameters of a calibrated Johanson model are then fully translatable across equipment scales. Additionally, the PIDs attached to the equipment allow precise control of both the roll force and roll gap throughout the entire roll compaction run. Transferring a formulation to a production-scale equipment then becomes as straightforward as ensuring that key process parameters are kept the same as those from the pilot-scale experiments. The validity of these equipment design in scale-up studies has been demonstrated in a number of studies in the literature (Gago et al. 2018; Allesø et al. 2016). Enabling roll compaction scalability through equipment design has proven to be highly promising, and the pharmaceutical industry is increasingly resorting to this approach for its roll compaction scale-up endeavors.

7.8 Future Recommendations

In 1996, R.W. Miller contributed a chapter to the *Handbook of Pharmaceutical Granulation Technology*. The chapter is titled "Roller Compaction Technology." Much of the materials he wrote therein are invaluable information and, fittingly, cited many times throughout this chapter. In the closing of that chapter, he lamented the lack of use of roll compactor in the pharmaceutical industry, noting that "roll compaction was seldom used by the industry innovators and generic companies" and referring to a survey conducted around the time of his writing that concluded

that "… roller compaction does not fulfill the current granulation technology needs of the US pharmaceutical industry."

Today, just over two decades after that book was published, much has changed. Roll compaction has become the preferred granulation method by many in this industry. Transpiring switch from batch to continuous manufacturing helps incentivize the adaptation of roll compaction technology and all the more readily as major equipment manufacturers offer turnkey continuous manufacturing plants that come equipped with a roll compactor. An example is the ConsiGma™ manufacturing lines offered by GEA.

While what we fundamentally know about roll compaction has evolved significantly in a decade, much needed innovation and gaps in the science of the process still yet remain. Among them are the following:

1. Integration of process analytical technology for characterization of the roll compaction process.

   Today's roll compactors are capable of accurately measuring, controlling, and recording key process parameters (e.g., roll force, roll speed, etc.). However, they do not come standard with the technologies capable of measuring the process outcomes (e.g., ribbon density, granules size distribution, etc.) in real time. Studies have shown that ribbon densities could be accurately measured using near-infrared (NIR) technology. Acevedo et al. (2012) showed different methods to calibrate NIR spectra for real-time measurements of ribbon densities. McCann et al. (2010) used four NIR sensors positioned along the roll width at equal spacing to successfully measure the density distribution of the ribbon. Other authors have used similar approaches with great success (Crowley et al. 2017). As outlined above, load cells embedded within the rolls could serve as a highly valuable tool for building a predictive model for ribbon density, as well as for process scale-up applications. Unfortunately, at the time of this writing, commercial roll compactors do not have such sensors built-in, and thus their use in production is prohibitive.
2. Better and more accurate predictive roll compaction models.

   Promising recent work that utilized finite element simulations of the roll compaction process has shown that quantitative prediction of ribbon density distribution could be made when the streamwise stress at the roll inlet zone is known (Muliadi et al. 2013). However, there is currently no means for predicting and measuring that quantity. More importantly, finite element results have elucidated the reasons why the Johanson model, in its original form (1965), is inaccurate. The modifications to the Johanson models proposed by Bi et al. (2014) and Liu and Wassgren (2016) showed that reasonably accurate predictions for ribbon density are possible in the framework of a one-dimensional analytical solution, even if it must be combined with empirical calibrations or finite element modeling. The challenge, then, is to derive a fully analytical solution that addresses the shortcomings of the Johanson model, therefore allowing for accurate prediction of the ribbon density without the aid of at-scale empirical data and/or finite element results.

3. Material-sparing approaches to roll compaction developments.

   The use of DoE in roll compaction process developments requires performing extensive at-scale experimentation and therefore consumes a lot of APIs. Metzler et al. (2016) and Reimer and Kleinebudde (2019) proposed replacing these at-scale roll compaction experiments with ones performed on a highly instrumented compaction simulator that mimics the dwell time of the rolls. A special tooling was used to produce rectangular compacts having the same width as the ribbons. These rectangular compacts were then milled using a smaller replica of a roll compactor mill (Metzler et al. 2016). The authors showed that the roll force required to compress the ribbon to a given density could be predicted from the compaction simulator force via the use of an empirical correlation. They further noted that the characteristics of the granules made from the mill replica generally match those produced by the actual roll compactor mill. Such an approach, while still largely empirical, consumes only a fraction of the material needed to perform a similar study at scale and thus is worth further scrutiny in the development process.

Looking forward to the next decade, research focusing on the above three areas will without a doubt allow us to further understand the roll compaction process and more effectively use its technology in pharmaceutical granulation developments.

References

Acevedo D, Muliadi A, Giridhar A, Litster JD, Romanach RJ. Evaluation of three approaches for real-time monitoring of roller compaction with near-infrared spectroscopy. AAPS PharmSciTech. 2012;13:1005–12. https://doi.org/10.1208/s12249-012-9825-0.

Allesø M, Holm R, Holm P. Roller compaction scale-up using roll width as scale factor and laser-based determined ribbon porosity as critical material attribute. Eur J Pharm Sci. 2016;87:69–78.

am Ende MT, et al. Improving the content uniformity of a low-dose tablet formulation through roller compaction optimization. Pharm Dev Technol. 2007;12:391–404. https://doi.org/10.1080/10837450701369253.

Bi M, Alvarez-Nunez F, Alvarez F. Evaluating and modifying Johanson's rolling model to improve its predictability. J Pharm Sci. 2014;103:2062–71.

Bindhumadhavan G, Seville JPK, Adams MJ, Greenwood RW, Fitzpatrick S. Roll compaction of a pharmaceutical excipient: experimental validation of rolling theory for granular solids. Chem Eng Sci. 2005;60:3891–7.

Boersen N, Belair D, Peck GE, Pinal R. A dimensionless variable for the scale up and transfer of a roller compaction formulation. Drug Dev Ind Pharm. 2016;42:60–9.

Crowley ME, et al. Near-infrared monitoring of roller compacted ribbon density: investigating sources of variation contributing to noisy spectral data. Eur J Pharm Sci. 2017;102:103–14. https://doi.org/10.1016/j.ejps.2017.02.024.

Csordas K, Kleinebudde P. Evaluation of the performance of different types of roll compactors. Powder Technol. 2017; https://doi.org/10.1016/j.powtec.2017.04.009.

Cunningham JC. Experimental studies and modeling of the roller compaction of pharmaceutical powders. PhD thesis, Drexel University; 2005.

Cunningham JC, Winstead D, Zavaliangos A. Understanding variation in roller compaction through finite element-based process modeling. Comput Chem Eng. 2010;34:1058–71.

Dawes J, Gamble JF, Greenwood R, Robbins P, Tobyn M. An investigation into the impact of magnesium stearate on powder feeding during roller compaction. Drug Dev Ind Pharm. 2012;38:111–22.

Dec RT. Study of compaction process in roll press. Proc Inst Briquet Agglomerat. 1991;22:207–18.

Dec RT, Zavaliangos A, Cunningham JC. Comparison of various modeling methods for analysis of powder compaction in roller press. Powder Technol. 2003;130:265–71.

Deshmukh AR, Sundararajan T, Dube RK, Bhargava S. Analysis of cold densification rolling of a sintered porous metal strip. J Mater Process Technol. 1998;84:56–72.

Fahmy R, et al. Quality by design I: application of failure mode effect analysis (FMEA) and Plackett-Burman design of experiments in the identification of "main factors" in the formulation and process design space for roller-compacted ciprofloxacin hydrochloride immediate-release tablets. AAPS PharmSciTech. 2012;13:1243–54. https://doi.org/10.1208/s12249-012-9844-x.

Gago AP, Reynolds G, Kleinebudde P. Impact of roll compactor scale on ribbon density. Powder Technol. 2018; 337:92–103.

Guigon P, Simon O. Roll press design—influence of force feed systems on compaction. Powder Technol. 2003;130:41–8. https://doi.org/10.1016/S0032-5910(02)00223-1.

Gun GY, Stebunov SA, Katashinskii VP. Computer modeling and investigation of the metal powder rolling processes. Powder Metall Metal Ceram. 1986;25:8–11.

Gupta A, Peck GE, Miller RW, Morris KR. Effect of the variation in the ambient moisture on the compaction behavior of powder undergoing roller-compaction and on the characteristics of tablets produced from the post-milled granules. J Pharm Sci. 2005;94:2314–26. https://doi.org/10.1002/jps.20414.

Gurson AL. Continuum theory of ductile rupture by void nucleation and growth: Part 1–yield criteria and flow rules for porous ductile media. ASME J Eng Mater Technol. 1977;99:2–15.

Hancock BC, Vendola TA. The effect of mill type on two dry-granulated placebo formulations. AAPS PharmSciTech. 2008;32

Hancock B, Colvin JT, Mullarney MP, Zinchuk AV. The relative densities of pharmaceutical powders, blends, dry granulations, and immediate-release tablets. Pharm Technol. 2003;27:64–80.

He X, Secreast PJ, Amidon GE. Mechanistic study of the effect of roller compaction and lubricant on tablet mechanical strength. J Pharm Sci. 2007;96:1342–55.

Inghelbrecht S, Paul Remon J. The roller compaction of different types of lactose. Int J Pharm. 1998a;166:135–44. https://doi.org/10.1016/S0378-5173(98)00022-2.

Inghelbrecht S, Paul Remon J. Roller compaction and tabletting of microcrystalline cellulose/drug mixtures. Int J Pharm. 1998b;161:215–24.

Iyer RM, Hegde S, DiNunzio J, Singhal D, Malick W. The impact of roller compaction and tablet compression on physicomechanical properties of pharmaceutical excipients. Pharm Dev Technol. 2014;19:583–92.

Johanson JR. A rolling theory for granular solids. J Appl Mech. 1965;32:842–8.

Katashinskii VP. Analytical determination of specific pressure during the rolling of metal powders. Powder Metall Metal Ceram. 1986;5:765–72.

Katashinskii VP, Shtern MB. Stress-strain state of powder being rolled in the densification zone–II. Distribution of density, longitudinal strain, and contact stress in the densification zone. Powder Metall Metal Ceram. 1983a;22:972–6.

Katashinskii VP, Shtern MB. Stress-strain state of powder being rolled in the densification zone–I. mathematical model of rolling in the densification zone. Powder Metall Metal Ceram. 1983b;22:882–5.

Kleinebudde P. Roll compaction/dry granulation: pharmaceutical applications. Eur J Pharm Biopharm. 2004;58:317–26. https://doi.org/10.1016/j.ejpb.2004.04.014.

Krok A, Peciar M, Fekete R. Using the DPIV optical technique to measure the velocity of powder material in the space between the rollers in a roll compactor. Powder Technol. 2014;262:131–41. https://doi.org/10.1016/j.powtec.2014.04.067.

Kushner J, Langdon BA, Hiller JI, Carlson GT. Examining the impact of excipient material property variation on drug product quality attributes: a quality-by-design study for a roller compacted, immediate release tablet. J Pharm Sci. 2011;100:2222–39.
LaMarche K, Buckley D, Hartley R, Qian F, Badawy S. Assessing materials' tablet compaction properties using the Drucker–Prager Cap model. Powder Technol. 2014;267:208–20.
Liu Y, Wassgren C. Modifications to Johanson's roll compaction model for improved relative density predictions. Powder Technol. 2016;297:294–302. https://doi.org/10.1016/j.powtec.2016.04.017.
Liu Z, et al. Scale-up of a pharmaceutical roller compaction process using a joint-Y partial least squares model. Ind Eng Chem Res. 2011;50:10696–706.
Mazor A, Perez-Gandarillas L, de Ryck A, Michrafy A. Effect of roll compactor sealing system designs: a finite element analysis. Powder Technol. 2016;289:21–30. https://doi.org/10.1016/j.powtec.2015.11.039.
Mazor A, Orefice L, Michrafy A, de Ryck A, Khinast JG. A combined DEM & FEM approach for modelling roll compaction process. Powder Technol. 2017; https://doi.org/10.1016/j.powtec.2017.04.053.
McCann R, Muliadi A, Paaso J, Litster JD, Pinal R. On-line monitoring of roller compaction ribbon density using NIR spectroscopy. Paper presented at the AIChE annual meeting, Salt Lake City, UT, 7–12 November 2010.
Metzler C, et al. A rapid, materials-sparing approach to the exploration of a roller compaction process design space. Paper presented at the AIChE annual meeting, San Francisco, CA, 13–18 November 2016.
Michrafy A, Diarra H, Dodds JA, Michrafy M. Experimental and numerical analyses of homogeneity over strip width in roll compaction. Powder Technol. 2011;206:154–60. https://doi.org/10.1016/j.powtec.2010.04.030.
Miguelez-Moran AM, Wu C-Y, Seville JP. The effect of lubrication on density distributions of roller compacted ribbons. Int J Pharm. 2008;362:52–9.
Miguelez-Moran AM, Wu CY, Dong H, Seville JP. Characterisation of density distributions in roller-compacted ribbons using micro-indentation and X-ray micro-computed tomography. Eur J Pharm Biopharm. 2009;72:173–82. https://doi.org/10.1016/j.ejpb.2008.12.005.
Miller RW. Roller compaction technology. In Parikh DM, editor. In Handbook of pharmaceutical granulation technology. Marcel Dekker; New York, NY. 1997.
Miller T, York P. Pharmaceutical tablet lubrication. Int J Pharm. 1988;41:1–19.
Mosig J, Kleinebudde P. Critical evaluation of root causes of the reduced compactability after roll compaction/dry granulation. J Pharm Sci. 2015;104:1108–18.
Muliadi AR, Litster JD, Wassgren CR. Modeling the powder roll compaction process: comparison of 2-D finite element method and the rolling theory for granular solids (Johanson's model). Powder Technol. 2012;221:90–100.
Muliadi AR, Litster JD, Wassgren CR. Validation of 3-D finite element analysis for predicting the density distribution of roll compacted pharmaceutical powder. Powder Technol. 2013;237:386–99. https://doi.org/10.1016/j.powtec.2012.12.023.
Nesarikar VV, et al. Roller compaction process development and scale up using Johanson model calibrated with instrumented roll data. Int J Pharm. 2012a;436:486–507.
Nesarikar VV, et al. Instrumented roll technology for the design space development of roller compaction process. Int J Pharm. 2012b;426:116–31. https://doi.org/10.1016/j.ijpharm.2012.01.032.
Omar CS, et al. Roller compaction: effect of morphology and amorphous content of lactose powder on product quality. Int J Pharm. 2015;496:63–74.
Otsuka M, Yamane I, Matsuda Y. Effects of lubricant mixing on compression properties of various kinds of direct compression excipients and physical properties of the tablets. Adv Powder Technol. 2004;15:477–93.

Perez-Gandarillas L, et al. Effect of roll-compaction and milling conditions on granules and tablet properties. Europ J Pharm Biopharm. 2016;106:38–49. https://doi.org/10.1016/j.ejpb.2016.05.020.

Reimer HL, Kleinebudde P. Hybrid modeling of roll compaction process with the Styl'One evolution. Powder Technol. 2019;341:66–74.

Reynolds G, Ingale R, Roberts R, Kothari S, Gururajan B. Practical application of roller compaction process modeling. Comput Chem Eng. 2010;34:1049–57.

Rowe JM, et al. Mechanistic insights into the scale-up of the roller compaction process: a practical and dimensionless approach. J Pharm Sci. 2013;102:3586–95.

Sakwanichol J, Puttipipatkhachorn S, Ingenerf G, Kleinebudde P. Roll compaction/dry granulation: comparison between roll mill and oscillating granulator in dry granulation. Pharm Dev Technol. 2012;17:30–9. https://doi.org/10.3109/10837450.2010.508078.

Samanta AK, Ng KY, Heng PW. Cone milling of compacted flakes: process parameter selection by adopting the minimal fines approach. Int J Pharm. 2012;422:17–23. https://doi.org/10.1016/j.ijpharm.2011.10.015.

Sheskey P, Pacholke K, Sackett G, Maher L, Polli J. Effect of process scale-up on robustness of tablets, tablet stability, and predicted in vivo performance. Pharm Technol. 2000;24:30–52.

Shi W, Hajedemos D, Sprockel O. An energetic approach of powder mixing by manipulating cohesive interaction in a magnetic field. Powder Technol. 2013;235:400–4.

Shima S, Yamada M. Compaction of metal powder by rolling. Powder Metall. 1984;27:39–44.

Simon O, Guigon P. Correlation between powder-packing properties and roll-press compact heterogeneity. Powder Technol. 2003;130:257–64.

Sommer K, Hauser G. Flow and compression properties of feed solids for roll-type presses and extrusion presses. Powder Technol. 2003;130:272–6. https://doi.org/10.1016/s0032-5910(02)00204-8.

Souihi N, et al. A quality by design approach to investigate the effect of mannitol and dicalcium phosphate qualities on roll compaction. Int J Pharm. 2013a;447:47–61.

Souihi N, Josefson M, Tajarobi P, Gururajan B, Trygg J. Design space estimation of the roller compaction process. Ind Eng Chem Res. 2013b;52:12408–19. https://doi.org/10.1021/ie303580y.

Souihi N, et al. Roll compaction process modeling: transfer between equipment and impact of process parameters. Int J Pharm. 2015;484:192–206. https://doi.org/10.1016/j.ijpharm.2015.02.042.

Sun CC, Himmelspach MW. Reduced tabletability of roller compacted granules as a result of granule size enlargement. J Pharm Sci. 2006;95:200–6. https://doi.org/10.1002/jps.20531.

Sun CC, Kleinebudde P. Mini review: mechanisms to the loss of tabletability by dry granulation. Eur J Pharm Biopharm. 2016;106:9–14.

Teng Y, Qiu Z, Wen H. Systematical approach of formulation and process development using roller compaction. Eur J Pharm Biopharm. 2009;73:219–29. https://doi.org/10.1016/j.ejpb.2009.04.008.

Tezyk M, Jakubowska E, Milanowski B, Lulek J. Implementation of quality by design approach in manufacturing process optimization of dry granulated, immediate release, coated tablets – a case study. Drug Dev Ind Pharm. 2017;43:1626–36. https://doi.org/10.1080/03639045.2017.1328431.

von Eggelkraut-Gottanka SG, Abed SA, Muller W, Schmidt PC. Roller compaction and tabletting of St. John's wort plant dry extract using a gap width and force controlled roller compactor. II. Study of roller compaction variables on granule and tablet properties by a 3×3 factorial design. Pharm Dev Technol. 2002;7:447–55.

Vromans H, Bolhuis G, Lerk C, Van de Biggelaar H, Bosch H. Studies on tableting properties of lactose. VII. The effect of variations in primary particle size and percentage of amorphous lactose in spray dried lactose products. Int J Pharm. 1987;35:29–37.

Wang W, Cunningham JC, Zavaliangos A. The effect of side constraint in rolling compaction of powders. Paper presented at the international conference on powder metallurgy and particulate materials, Las Vegas, NV, 8-12 June 1998.

Wang J, Wen H, Desai D. Lubrication in tablet formulations. Eur J Pharm Biopharm. 2010;75:1–15.

Wu SJ, Sun C. Insensitivity of compaction properties of brittle granules to size enlargement by roller compaction. J Pharm Sci. 2007;96:1445–50. https://doi.org/10.1002/jps.20929.

Yu S. Roll compaction of pharmaceutical excipients. Ph.D. thesis, University of Birmingham; 2012.

Yu S, et al. The effects of lubrication on roll compaction, ribbon milling and tabletting. Chem Eng Sci. 2013;86:9–18.

Yusof YA, Smith AC, Briscoe BJ. Roll compaction of maize powder. Chem Eng Sci. 2005;60:3919–31.

Zavaliangos A, Dec RT, Komarek RK. Analysis of powder processing in the roller press using finite element modeling. Paper presented at the XXII international mineral processing congress, Cape Town, South Africa, 29 September – 3 October 2003.

Chapter 8
Continuous Wet Granulation

Arwa El Hagrasy, Li Ge Wang, and Jim Litster

Abstract Wet granulation is a critical unit operation in solid dosage form manufacturing in the pharmaceutical industry. Traditionally, wet granulation has been a batch process. Recently, there has been a move toward more advanced manufacturing approaches such as continuous processing to allow more rigorous process control, consistent quality assurance, and reduced capital costs. This chapter discusses continuous wet granulation process with emphasis on twin screw granulation as the most commonly used continuous granulation approach in the pharmaceutical industry. The key design features of a twin screw granulator (TSG) are described, and comparisons with batch high shear granulation equipment are made. The effects of formulation and process variables on granule attributes are discussed. Mechanistic studies of screw elements with the proposed granulation rate processes are presented. Real-time process monitoring tools, including spectroscopic and imaging techniques, are described. Implementation of dimensional analysis as a tool for scaling up/scaling out of continuous twin screw granulation is presented. Two numerical methods, that is, population balance model (PBM) and discrete element method (DEM), are chosen for the multiscale modeling of twin screw granulation, and their coupling mechanisms of data exchange between DEM and PBM are discussed.

Keywords Continuous wet granulation · Process design and process control · Twin screw granulation · Population balance model · Discrete element method · Granulation rate processes · Residence time distribution

The opinions and conclusions expressed in this chapter are solely the views of the authors and do not necessarily reflect those of the Food and Drug Administration.

A. El Hagrasy
Office of Pharmaceutical Manufacturing Assessment, Office of Pharmaceutical Quality, CDER, FDA, Silver Spring, MD, USA

L. G. Wang · J. Litster (✉)
Chemical and Biological Engineering, University of Sheffield, Sheffield, UK
e-mail: James.Litster@sheffield.ac.uk

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_8

Nomenclature

a	Particle radius, mm
b_M	Breakage function
$B(v, t)$	Birth rate, m^{-3} s^{-1}
B_{agg}	Birth rate of agglomeration, m^{-3} s^{-1}
B_{break}	Birth rate of breakage, m^{-3} s^{-1}
B_{nuc}	Birth rate of nucleation, m^{-3} s^{-1}
C_{impact}	Impact frequency, s^{-1}
d_p	Particle diameter, mm
D	Barrel diameter, mm
$D(v, t)$	Death rate, m^{-3} s^{-1}
D_{break}	Death rate of breakage, m^{-3} s^{-1}
$D_{m_p,\ nuc}$	Death rate of powder particles, kg m^{-3} s^{-1}
d_{50}	Median size of feed particle, mm
F_n	Series of geometric ratios
F_{powder}^{in}	Flow rate of additional powder stream, kgs^{-1}
G	Growth rate, s^{-1}
G_m	Maximum growth rate, s^{-1}
G_{shear}	Shear rate, s^{-1}
h	Interparticle gap, mm
k	Compaction rate constant, s^{-1}
l	Liquid volume, m^3
L	Barrel length, mm
$\dot{L}_{in,powder}(x,t)$	Rate of liquid addition to the fine powder, kgh^{-1}
m	Mass of particle, kg
M_{powder}	Mass of fine powder, kg
$M_{granule}$	Mass of granule, kg
$\dot{m}_p$	Mass flow rate of powder, kg/h
$\dot{m}_l$	Mass flow rate of liquid, kg/h
n	Population density of length function, m^{-3}
P_1	Rate coefficient
P_2	Size-dependent exponent
St_v	Stokes deformation number
$V_{droplet}$	Volume of a single liquid droplet, m^3
V_{nuc}	Volume of particle in saturated granule, m^3
$V_{L,\ nuc}$	Volume of liquid in saturated granule, m^3
V_p	Per-particle pore volume, m^3
S_1	Volume of solid component 1, m^3
S_2	Volume of solid component 2, m^3
$S_M(w, v)$	Specific breakage rate, s^{-1}
x_w	Moisture content of powder
x_{wc}	Critical moisture content of powder
Δx	Reduction of particle size

σ	Standard deviation
ω	Angular velocity of the shaft, rads^{-1}
ε	Granule porosity
ε_{bed}	Powder bed porosity
ε_{min}	Minimum granule porosity
ρ_b	Particle density, kg m^{-3}
u_0	Collision velocity, ms^{-1}
μ	Binder viscosity
LSR	Liquid/solid ratio
PFN	Powder feed number
Fr	Froude number

8.1 Introduction

In contrast to some other unit operations such as roll compaction and tableting, wet granulation in the pharmaceutical industry has traditionally been a truly batch process, undertaken in mixing equipment borrowed from the food processing industry. By far the most common batch granulator designs are vertical shaft high shear mixers, with a smaller number of horizontal shaft mixers also used (Hapgood et al. 2007).

In other industry sectors, continuous wet granulation is standard and most often used in high-throughput applications in mineral processing, fertilizers and industrial chemicals, and detergents. Tumbling granulators (drums and pans), fluidized bed granulators, and horizontal axis high shear mixers have all been used in continuous mode (Litster and Ennis 2004). The main driver in these industries for continuous processes is economies of scale. Therefore, standard equipment designs for these industries have throughputs from 1 to 100 tonne hr.$^{-1}$ and are much too large for pharmaceutical applications. In moving from batch to continuous processing, the pharmaceutical industry had to look for continuous processing options with much lower capacities in the range 5–50 kg hr.$^{-1}$. Once again, the industry borrowed equipment from another application – twin screw extruders used for polymer extrusion.

Figure 8.1 gives an illustration of a twin screw granulator (TSG). Powder and liquid are added separately to the twin screw barrel. Multiple addition points are possible. The wet granulation mass is conveying forward down the barrel. In contrast to operation as an extruder, there is no die plate, and the powder state is more akin to that in a screw feeder. In most of the barrel, the powder fill is low and there is no buildup of pressure along the barrel. Granules flow freely out of the end of the barrel and usually drop directly into a fluid bed drier. There is great flexibility in the choice of the length and configuration of the corotating screws with a wide variety of screw elements available for use including standard conveying elements (CE), kneading elements (KE), and distributive missing elements (DME).

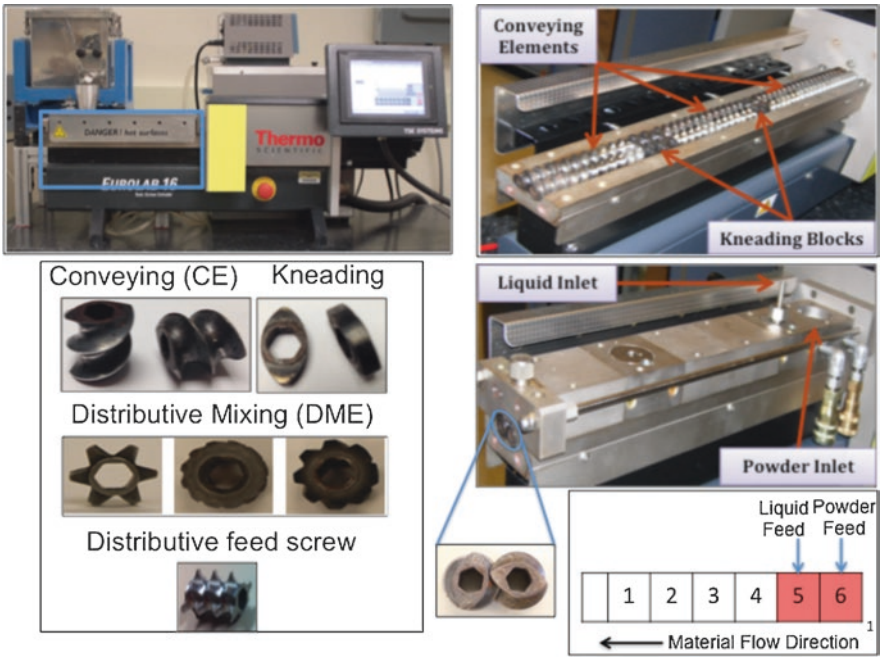


Fig. 8.1 Illustration of the key features of a twin screw granulator

The TSG has a number of intriguing features that contrast markedly from batch high shear wet granulation. The residence time in the granulator is short (5–20 s), and the pressure applied to the granules is low so that granules are high porosity, relatively weak, and not spherical. The confined nature of the barrel means good liquid distribution via a spray nozzle is not possible. Liquid is typically added via a jet or drips from a simple orifice so that bimodal granule distributions are common. Alternative methods of binder addition have been reported including foam granulation (Tan and Hapgood 2012) and melt binder (Mu and Thompson 2012). However, addition of liquid through a simple orifice is by far the most common in practice. The granules in the barrel are conveyed forward with limited back mixing. When combined with the ability to customize screw elements along the barrel, this gives the opportunity to develop a regime-separated granulator (Mu and Thompson 2012). The balance of rate processes in a TSG is very different to that in a HSWG, and TSG performance is much less sensitive to changes in formulation properties. Table 8.1 highlights the characteristics of continuous twin screw and batch high shear granulators and the very different products they produce.

In the following sections, we describe the controlling rate processes for granulation in a TSG, the impact of formulation, and process parameters on performance. The state of the art for process modeling of TSG is described and scale-up issues are discussed.

Table 8.1 Contrasting characteristics of twin screw granulation and high shear mixer granulation

HSWG characteristics	TSG characteristics
Batch process	Continuous process
Can handle a wide variety of powder and liquid feeds	Can handle a variety of powder and liquid feeds, but capacity may be limited by powder flow properties
Batch size up to 200 kg for 300 l granulator	Throughput up to 25 kg $hr.^{-1}$ for 25 mm barrel size
Batch times of 2–20 minutes	Residence time 5–20 seconds
Produces medium to high density granules	Produces low to medium density granules
Performance is sensitive to formulation properties	Performance is not sensitive to formulation properties
Coupled densification and growth are key rate processes	Breakage and layering are key rate processes
Changes to equipment are difficult	Screw configuration is easy to change leading to regime-separated granulation designs
Scale up through 3 or 4 scales from development to manufacturing. Granule properties difficult to maintain during scale-up	Development and manufacturing can be done at the same scale by increasing the continuous campaign time while keeping granule properties constant

8.2 Process Design and Characterization

8.2.1 Formulation and Process Variables

Effect of Binder Delivery Method/Binder Content and Delivery Method

The granulating binder in a twin screw granulation (TSG) process can be delivered in the powder phase, liquid phase, or a mixture of the two. The preferred method of binder delivery in TSG will depend on formulation properties as well as screw configuration. Formulation properties will define the rate of wetting, penetration, and spreading of the binder. Screw configuration determines the extent of shear imparted on the granulated material inside the TSG, which provides the energy needed to spread the granulating liquid and form liquid bridges between the particles.

El Hagrasy et al. observed that the use of solubilized binder reduced the level of fines and resulted in a narrower granule size distribution (GSD) compared to the dry binder delivery when using conveying elements (CE) and kneading elements (KE) in the screw configuration (El Hagrasy et al. 2013a). The short residence time in the TSG may prevent complete solubilization of the dry binder, thereby reducing its binding capacity compared to the dry form. Recent reports examined the impact of binder distribution and granulating liquid viscosity on granule properties and process attributes (Saleh et al. 2015; Dhenge et al. 2012a). The presence of binder in the liquid phase increases the viscosity of the granulating liquid, therefore increasing the liquid penetration time in a static bed. In the absence of adequate shear forces inside the granulator as in the case of CE, the less viscous granulating

liquid, that is, binder in the dry phase, produces narrower GSD due to faster spread and penetration of the droplets. The addition of KE increases the shear during granulation, resulting in a relatively more even distribution of the viscous binder solution and less difference in granule properties between the different binder delivery methods.

Dhenge et al. also examined the effect of incorporating increasing concentrations of binder in the liquid phase on granule attributes (Dhenge et al. 2012a). Higher binder concentrations in the liquid phase increase the viscosity of the granulating liquid. As a result, stronger liquid bridges bind the particles forming a larger median size of the granules. X-ray tomography images of the internal structure revealed less voids at higher binder concentrations. The residence time distribution (RTD) also increased when using higher liquid binder concentration, further contributing to the densification of the formed granules.

Considering that the mean residence time inside TSG is in the order of seconds, careful consideration of both the binder delivery method and the screw configuration used is important to ensure adequate distribution of the binder during granulation and desired granule attributes for downstream processing.

Effect of L/S Ratio

The variation in L/S ratio during granulation impacts the GSD obtained during TSG. El Hagrasy et al. reported that increasing L/S ratio initially shows a gradual increase in the ratio of coarse granules to ungranulated fines while maintaining the bimodal distribution characteristic of the TSG process (El Hagrasy et al. 2013a). Figure 8.2 shows a typical frequency distribution as a function of L/S ratio. The screw configuration was a combination of CE and KE. The broad granule size distribution obtained with TSG has been reported by other researchers as well, using a similar screw configuration (Dhenge et al. 2010; Djuric and Kleinebudde 2008, 2010; Djuric et al. 2009; Keleb et al. 2004; Van Melkebeke et al. 2008).

Per the nucleation regime map reported by Hapgood et al. for other types of granulation equipment, the operating regimes are defined by two dimensionless groups, namely, the dimensionless spray flux and the dimensionless drop penetration time (Hapgood et al. 2003, 2007). Due to the confined space inside a TSG, the granulating liquid is dripped directly into the powder inside the barrel. The relatively large droplets obtained compared to a spraying addition mode place the TSG nucleation process in the mechanical dispersion regime, which relies on shear forces for adequately distributing the granulating liquid. The broad bimodal size distribution of the granules at moderate L/S ratios points to inefficient liquid distribution by the standard screw elements typically used in TSG such as CE and KE.

Further increase in L/S ratio transfers the process into a rapid growth regime, where a rapid increase in d_{50} and a narrower monomodal GSD are observed. Nevertheless, the granule attributes, in terms of size, shape, and porosity, obtained at such high levels of L/S ratio will likely obviate their use in downstream pro-

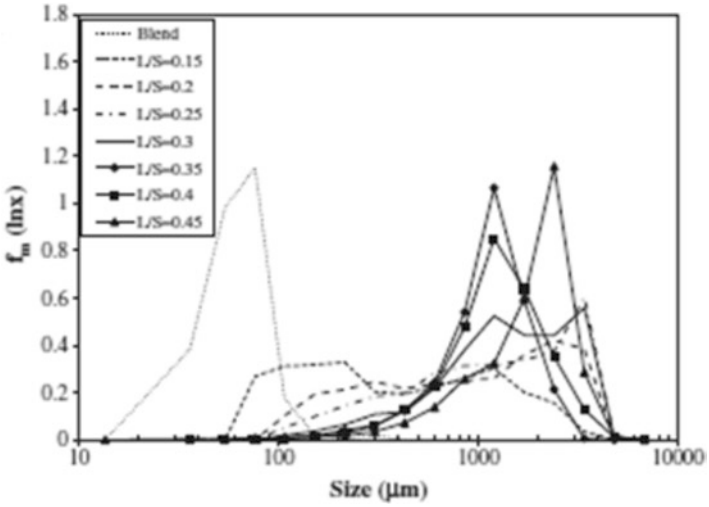


Fig. 8.2 The change in size distribution of granules obtained from a TSG process at different L/S ratios. (Adapted from El Hagrasy et al. (2013a))

cesses. The higher liquid content increases the deformability of the granules, yielding large elongated granules with dense structures.

Effect of Operating Channel Fill Level

The channel fill level inside a TSG depends on the material feed rate, screw speed, and the specific geometry of the screws and barrel. For a given screw configuration, variation in material feed rate and screw speed have opposite effects on the material holdup inside the granulator and thus the degree of channel fill (Kumar et al. 2014; Seem et al. 2015). Higher material input rate into TSG lowers the degree of channel fill, whereas higher screw speed decreases the degree of channel fill. The screw configuration affects the degree of channel fill because of the difference in the free volume available in the barrel, depending on the screw geometry used. CE generally offer more free volume than other types of mixing elements, thereby allowing for less restrictive flow of material through the granulator. With mixing elements, such as KE or distributive mixing elements (DME), the free volume will depend on the number of elements as well as the offset angle used. Thus, the fill level will also vary in the axial direction of the granulator, depending on the screw configuration along the shaft (Lee et al. 2012).

Recent studies characterizing the mixing dynamics inside the TSG using positron emission particle tracking (PEPT) and near-infrared (NIR) chemical imaging techniques allowed a faster sampling rate and thus more in-depth understanding of material transport inside the granulator and how it is impacted by operating parameters such as screw speed, material input rate, and screw configuration (Kumar

et al. 2014; Lee et al. 2012). Within the material flow rate and screw speed suitable for PEPT (10–20 g/min and 150–300 rpm, respectively), Lee et al. concluded that the RTD decreases with the increase in screw speed and material feed rate. The increase in offset angle was generally shown to increase the RT because of more restricted flow. However, normalized RTD curves and Peclet numbers revealed no difference in the degree of axial mixing at the studied variables of screw speed, material feed rate, and screw configuration. Kumar et al. quantified their RTD results in terms of the mean RT (t_m), the Peclet number (Pe), and the normalized variance (σ^2_θ) as a descriptor of the extent of axial mixing. The impact of screw speed and material input rate on RT observed using NIR chemical imaging by Kumar et al. was consistent with the PEPT study, where higher screw speed and material input rate reduced the (t_m). However, the analysis of (σ^2_θ) and Pe showed that varying material feed rate, screw speed, number of mixing elements, and their offset angles resulted in substantial variation in the mixing behavior. Plug flow was enhanced by low screw speed, high material feed rate, number of KE, and their offset angle. Although a higher number of KE and/or offset angle increase RTD, the material flow was hindered due to the restrictive flow imparted by the large number of KE combined with the low screw speed. Kumar et al. further studied the mixing dynamics inside the TSG and its impact on solid-liquid mixing and GSD as a function of screw speed, material throughput, L/S ratio, number, and offset angle of KE (Kumar et al. 2016). The study demonstrated that despite high axial mixing observed in certain experiments, yet the granule attributes were not optimal due to increase in the level of fines and coarse granules, an indication of insufficient liquid distribution characteristic of KE (El Hagrasy et al. 2013a; El Hagrasy and Litster 2013).

While the specific results in the literature were not always consistent, the overall conclusion is that the parameters impacting the degree of channel fill have a substantial effect on mixing behavior, material flow, and the resulting granule attributes. It is thus important to optimize the granulation operating parameters in an integrated CM process to allow sufficient degree of dispersive mixing and adequate liquid distribution in the granulator to achieve desired granule attributes. Mechanistic understanding of granulation rate processes, RTD characterization, and imaging techniques becomes important tools in this regard because they allow in-depth understanding of the impact of those parameters. Characterization of granulation process dynamics through RTD studies is particularly valuable in defining the optimum process and equipment design parameters that avoid plug flow and allow sufficient operational time to yield adequate liquid distribution.

8.2.2 Mechanistic Studies and Screw Characterization

Screw configuration has a strong impact on granule attributes in a TSG process. There are different screw designs available for use with a TSG. The early literature in this area has mostly described the use of CE and KE combinations, which are

originally adopted from the extrusion literature (Thiele 2003). The use of mixing elements has also been described in the TSG literature (Sayin et al. 2015a). The modular design of a TSG allows innumerable number of possible screw configurations, making screw configuration one of the most complex equipment design parameters in a TSG.

The output granule properties from a TSG are the result of granule evolution along the length of the TSG. Therefore, mechanistic understanding of granule formation, using different screw element designs and their combinations, is essential to optimizing the TSG process. Characterizing the impact of individual screw elements on different granule attributes provides a framework for the development of multidimensional PBM that can be implemented for future process optimization and control.

Conveying Elements (CE)

CE are helical-shaped screw elements known for their excellent transport properties and so are typically positioned at the inlet of liquid/solid material into the TSG and to transport material to or from different mixing zones. The powder in CE is subjected to minimal shear compared to other mixing zones in the granulator, which plays an important role in defining granule properties in this section as described later. CE design variables that determine channel volume and extent of drag flow include:

1. Flight pitch: axial distance between the tip of two adjacent flights.
2. Flight lead: axial displacement corresponding to one turn of the flight.

The effect of design variables of CE on granule properties was studied by different research groups (Djuric and Kleinebudde 2008; Thompson and Sun 2010). In general, larger element pitch provided smoother and more even feeding to subsequent elements at a given flow rate. However, GSD was essentially similar (within experimental error) across the length of the CE, indicating minimal influence of CE on granule attributes.

Granulating liquid is dripped into the conveying section during operation, leading to granule formation by immersion since the liquid drop is substantially larger than the powder primary particles (Dhenge et al. 2012b). El Hagrasy and Litster demonstrated that a bimodal size distribution of ungranulated fines and wet agglomerates is generally produced (El Hagrasy and Litster 2013). The wet agglomerates represent the nuclei formed during the initial wetting step, and the higher L/S ratio leads to droplet coalescence and formation of a higher fraction of the large agglomerates. Figure 8.3 shows that the fraction of coarse granules to fines varied depending on the L/S ratio, with a higher fraction of coarse granules obtained at higher L/S ratios. However, minimal liquid distribution and shape transformation occurred in this section regardless of the L/S ratio used.

Granule growth in CE was also found to be dependent on viscosity of the granulating liquid (Dhenge et al. 2013). The lower viscosity of the granulating liquid

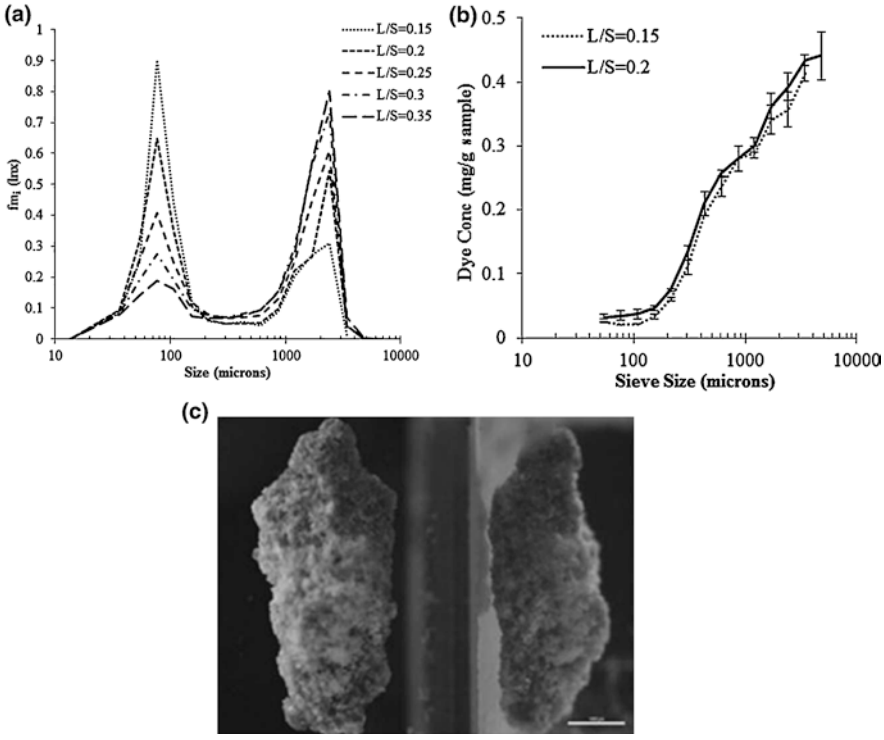


Fig. 8.3 (**a**) The change in size distribution of granules from the conveying section at different L/S ratios, (**b**) dye concentration in the different size fractions of granules from the conveying section; the error bar at each data point is the standard deviation of three measurements; (**c**) micrographs of top (left) and side (right) view of a representative granule from the conveying section. (Adapted from El Hagrasy and Litster (2013))

resulted in shorter liquid penetration time and relatively better granulating liquid distribution, leading to fewer small granules (<710 microns) and a larger fraction of medium and coarse granules (710–1000 microns and >1000 microns). Dhenge et al. linked the granulation process in the CE section of a TSG to flow behavior of dry and wet material with varying degrees of channel fill. Particle image velocimetry (PIV) and DEM studies showed that dry powder demonstrated the highest average particle velocity. The velocity decreased with the addition of low viscosity binder solutions, for example, water, because of increased cohesiveness of the formulation and subsequent dissipation of kinetic energy from the screws. Alternatively, the velocity behavior of the formulation approached that of dry powder as the viscosity of the granulating liquid increased due to the poor wettability ameliorated by the low shear environment in this zone.

The impact of material throughput in the CE section on granule size was less definitive and could not be generalized across different studies, where some reported an increase in the level of fines, while others reported the opposite effect (Djuric and

Kleinebudde 2008, 2010; Thompson and Sun 2010; Dhenge et al. 2013; Liu et al. 2015) The inconsistency among different studies could be attributed to the sensitivity of granule formation in this zone due to the low shear environment encountered, resulting in more pronounced impact of formulation, process, and equipment design parameters. Nevertheless, the design of CE upstream of the granulating elements, for example, KE, did not play a role in the granule attributes obtained downstream of KE (Liu et al. 2015). However, the CE used downstream of the kneading section influenced the granule properties with the studied 60F KE configuration.

Mixing Elements (KE and DME)

In a typical TSG operation, CE are used in combination with other types of mixing elements that provide the adequate shear and liquid distribution to produce granules suitable for downstream processing. Kneading elements (KE) provide a dispersive mixing zone for the incoming material from the conveying section, resulting in densification and breakage of large agglomerates and varying degrees of liquid distribution.

KE design variables: The KE configuration can be characterized in terms of (El Hagrasy and Litster 2013; Thompson and Sun 2010; Li et al. 2014):

1. Number of KE: defines the length of the kneading section.
2. Advance angle: offset angle between each successive pair of KE. Most common angles are 30, 45, 60, or 90. The AA defines the gap volume, which is the space in the middle channel between the tips of each pair of KE per the isometric drawings in Fig. 8.4. The divergence in the middle channel is expressed in terms of the diameter (D) of the barrel.
3. Angle orientation: forward (F) or reverse (R) depending on the rotational direction of the KE tips with respect to the flight of the preceding CE. The diagrams in Fig. 8.5a, b display the front view of the forward and reverse angle directions of KE at 60°, respectively.

El Hagrasy and Litster provided a detailed report on the impact of design variables of KE, including offset angle, angle orientation, and length of kneading section on granule physical properties and liquid distribution (El Hagrasy and Litster 2013). Figure 8.6 summarizes the dominant mechanisms occurring in the CE and KE sections of the TSG that govern the size, shape, and liquid distribution of the resulting granules. The two main granulation rate processes occurring in the kneading section can be summarized by:

1. Breakage, followed by layering.
2. Shear-elongation and breakage, followed by layering.

Breakage was the dominant mechanism in the 90° configuration, whereas shear-elongation was dominant in the 30 °R configuration. Other configurations demonstrated different degrees of combination between the two mechanisms. The breakage mechanism in 90° configuration can be attributed to the frequent intersection of the

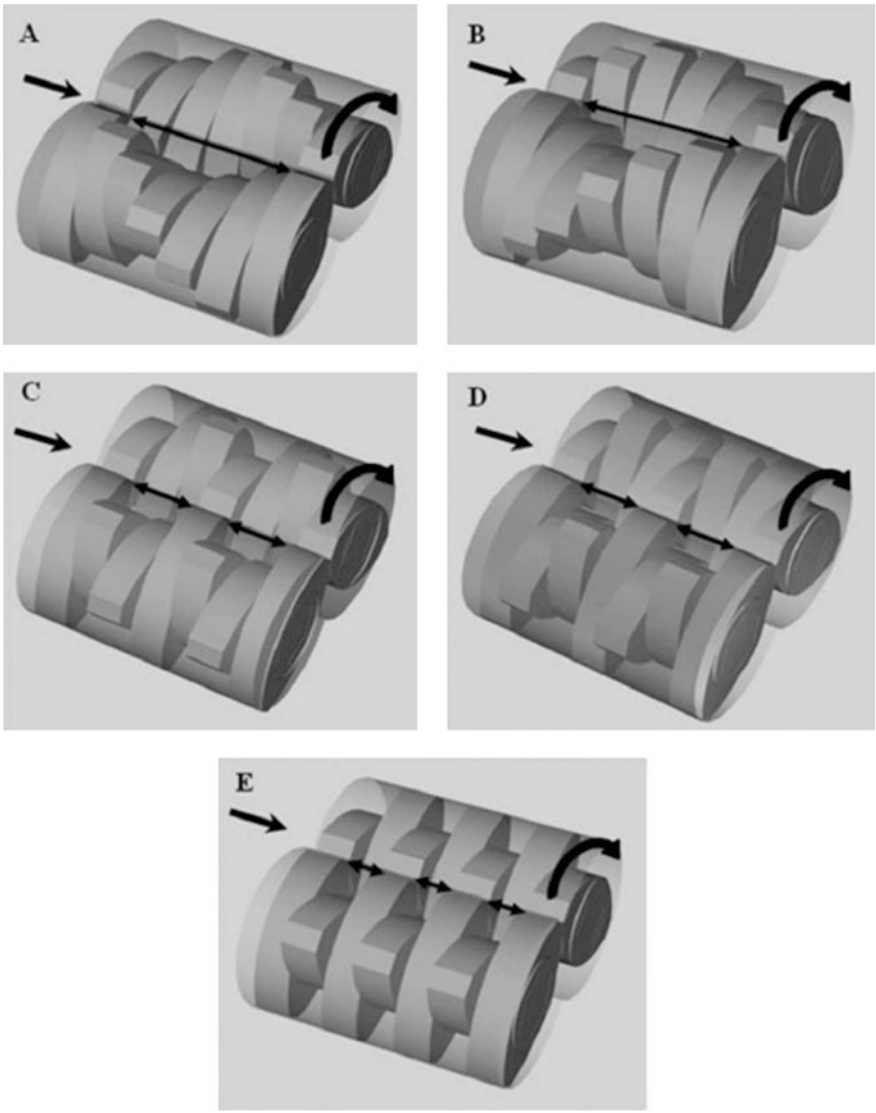


Fig. 8.4 Configurations of kneading section using 7 KE at (**a**) 30 °F, (**b**) 30 °R, (**c**) 60 °F, (**d**) 60 °R, and (**e**) 90°. The curved arrow indicates the direction of rotation of the shafts, and the straight arrow depicts the direction of material flow from the CE into the kneading section. The middle channel divergence length illustrated by the double-headed arrow is 1.25D for (**a**) and (**b**), 0.5D for (**c**) and (**d**), and 0.25D for (**e**). (Adapted from El Hagrasy and Litster (2013))

KE tips in the middle channel more than any other KE configuration (Fig. 8.6e). Breakage creates new wet surfaces that can pick up dry fines, thereby improving LD. The use of more KE in the 90 configuration enhances the breakage events, leading to simultaneous creation of more fines and improvement of LD.

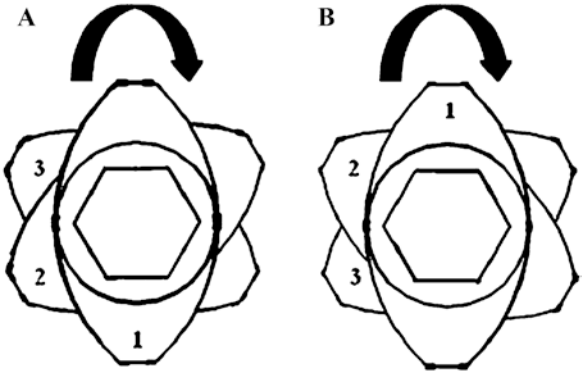


Fig. 8.5 Two-dimensional (2D) front view of the screw configuration for 3KE60 in the (**a**) forward angle direction and (**b**) reverse angle direction. The incremental order of KE shows the elements from front to back. The curved arrow indicates the directions of rotation of the shafts. (Adapted from El Hagrasy and Litster (2013))

Screw Type	Dominant Rate Processes	More KE	
		Liquid Distribution	Granule Growth
Conveying	Drop Nucleation (DN)	*NA*	*NA*
	Layering High L/S Ratio Droplet Coalescence Layering	*NA*	*NA*
Neutral	Breakage Layering	+	-
Reverse	30° Shear-Elongation and Breakage 30° Layering	+	+
	60° *Shear-Elongation and Breakage + Layering* *Breakage + Layering*	+	+
Forward	30° *Breakage + Layering + Conveying characteristics*	+	+
	60° *Conveying characteristics*	*Nil*	*Nil*

NA: Not applicable *+ Positive Change* *- Negative Change* *Nil: No Change*

Fig. 8.6 Proposed granulation rate processes in a twin screw granulator. (Adapted from El Hagrasy and Litster (2013))

In the reverse configurations, such as 30R and 60R, the wet agglomerates from the CE are smeared against the barrel wall. The shear-elongation mechanism, characteristic of 30 °R and 60 °R, improves liquid distribution inside TSG, especially with the addition of more KE. However, the resulting granules are dense with an undesirable flake-like shape that will likely cause powder handling problems during downstream processing. Other configurations demonstrated bimodal granule size distribution because of poor liquid distribution and/or breakage events, which requires the use of a downstream milling step of oversized granules to ensure drying homogeneity and avoid powder segregation issues.

Distributive mixing elements (DME), also known as comb mixing elements, are another type of mixing elements used in the TSG that have been studied by many researchers (Djuric and Kleinebudde 2008, 2010; Thiele 2003; Sayin et al. 2015a; Thompson and Sun 2010). Mixing in DME is attained through a series of cutting and recombination of inflowing granular material. Sayin et al. characterized the granulation rate processes in this type of mixing elements (Sayin et al. 2015a). The researchers investigated the impact of both screw orientation (forward versus reverse) and placement (adjacent versus spaced) on the resulting granule attributes. Figure 8.7 illustrates the design of the DME described in this work.

Liquid distribution and granule size significantly improved with the DME in the reverse orientation, regardless of placement. However, granule shape and porosity were not affected by DME orientation or placement. Flow visualization and RTD studies using a high-speed imaging camera showed that the reverse orientation improved the breakage of incoming granular material, leading to better liquid distribution through exchange between broken lumps and ungranulated fines. The dominant granulation rate processes occurring in the DME section are illustrated in Fig. 8.8.

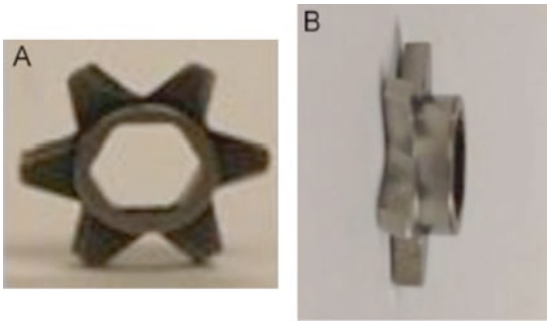


Fig. 8.7 Front view (**a**) and side view (**b**) of a distributive mixing element. (Adapted from Sayin et al. (2015a))

Screw Type	Dominant Rate Processes
DME - Forward	Breakage → + → Layering
DME - Reverse	Breakage → + → Layering

Fig. 8.8 Proposed granulation rate processes in distributive mixing elements in a twin screw granulator. (Adapted from Sayin et al. (2015a))

In contrast to KE, DME offer the potential of producing granules with more desirable attributes for downstream processing, including:

1. Monomodal size distribution without lumps that obviates the need for a downstream milling step.
2. Spherical-shaped porous granules with potentially improved flow properties.

The collective knowledge from the different studies used in characterization of screw elements in a TSG and the modeling work for simulation of the impact of screw configuration on granule attributes are yet to be utilized in optimizing the design of screw elements and configuration in a TSG process. To take it one step further would be to understand how to manipulate the functional role of different screw elements to custom-design desired granule attributes for different pharmaceutical applications a priori.

8.3 Real-Time Process Measurement and Process Control

In addition to the use of adequate science and engineering principles in process design, real-time process monitoring, implementation of process models, and process control tools are indispensable for continuous manufacturing. Real-time process measurement tools enhance process understanding that allows improved process design and optimization. The continuous monitoring of the process also ensures acceptable product quality throughout the manufacturing process. The real-time measurement of critical quality attribute (CQA) allows necessary process adjustments with minimal disruption of the continuous manufacturing line or diversion of nonconforming product without running the risk of rejecting the entire run.

NIR and Raman spectroscopy were used to examine the possibility of real-time measurement of solid-state changes of the active ingredient during twin screw granulation (Fonteyne et al. 2013). The spectroscopic measurements were collected on a static sample of granules at the outlet of the granulator. In addition, an in-line probe was placed in the stream of wet granules for particle size analysis. Both spectroscopic techniques provided an understanding of the impact of process parameters on the solid state of the active ingredient, with Raman technique showing more sensitivity than NIR in this study. However, probe fouling posed a challenge in using the particle size analyzer for in-line measurement of particle size distribution.

In-line monitoring of changes in granule size during a twin screw granulation process as a function of variation in process and formulation parameters was examined in another study (El Hagrasy et al. 2013a, b). An imaging camera (Eyecon™) was used to collect 3D surface maps of the sampled particles. An iterative process was then used to determine the equivalent diameter of the projected particles, from which the volume distribution was derived. The authors implemented statistical process control tools to evaluate the utility of different camera outputs,

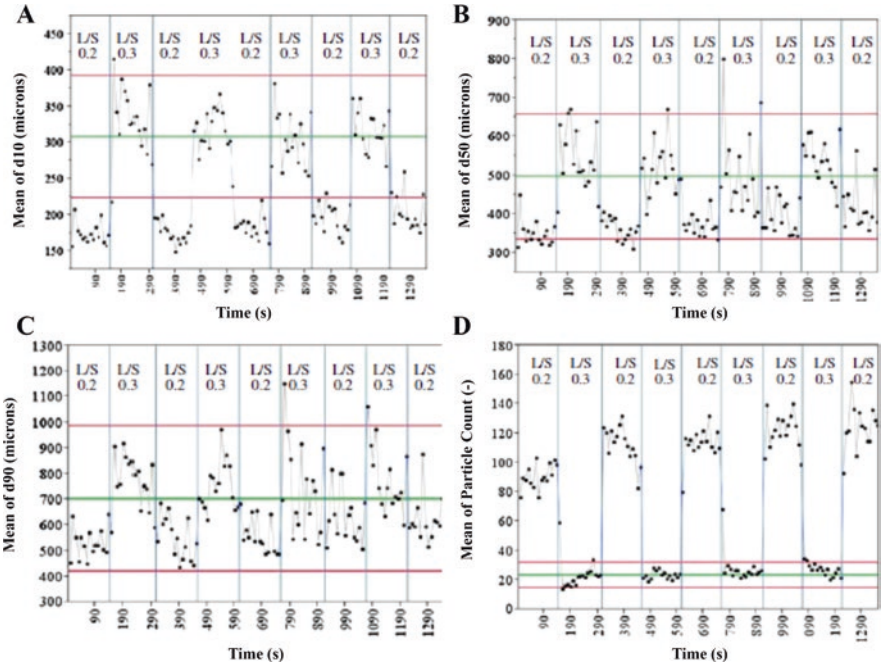


Fig. 8.9 Shewhart control limits applied to ($\overline{X}$) charts of d_{10} (**a**), d_{50} (**b**), d_{90} (**c**), and particle count (**d**). The upper and lower horizontal lines depict the upper control limit and lower control limit, respectively. The centerline depicts the process mean. The vertical lines mark the onset of change in L/S ratio. (Adapted from El Hagrasy et al. (2013b))

including statistics of particle diameter (d_{10}, d_{50}, and d_{90}) and particle count for granulation process monitoring and control. Figure 8.9 illustrates the ($\overline{X}$) charts for d_{10}, d_{50}, d_{90}, and particle count, in which the respective parameter in each case is plotted as a time series.

Particle count and d_{10} demonstrated the most robust performance among all studied particle quality characteristics. Unlike other calculated parameters such as d_{50} and d_{90}, d_{10} and particle count had inherently low measurement variability. Thus, they were more sensitive to changes in granule size attributes during routine process monitoring. An additional study from the same group confirmed the utility of this online technique with TSG (Sayin et al. 2015b).

Future optimization of the GSD characteristics from a TSG process is expected to further enhance the data quality obtained from direct measurement of granules using the camera or other process analyzers. Alternatively, such process monitoring tools can be interfaced with a downstream milling process to monitor milled granule size.

Moisture content is an important quality attribute of granules produced by a wet granulation process, especially that some formulations require tight control of moisture content to ensure optimal tablet characteristics. In-line NIR spectroscopy

was used for moisture content measurement on an integrated continuous granulation-fluid bed drying-milling process (Chablani et al. 2011). The NIR spectra were collected from a granule transfer hopper module situated between the fluid bed dryer and the mill. The work demonstrated the feasibility of using NIRS for measuring moisture content real time during processing.

However, the dynamic range of the calibration developed was limited because of the study design, in which the spectra were collected after the completion of the drying process. In another study, Fonteyne et al. placed the probe in the second cell of a six-cell segmented fluid bed dryer (Fonteyne et al. 2014). The probe placement earlier in the drying process allowed the variation in process parameters to produce sufficient difference in moisture levels. The calibration model developed was more robust due to wider range of moisture content values, and model validation showed that the NIR method was equivalent to the reference analytical method. Thus, it is important to carefully consider sensor-equipment interface with the continuous process to ensure that representative samples are collected for calibration development, validation, and method implementation.

Besides measuring granule attributes and physicochemical properties, process analytical technology (PAT) tools were also utilized to understand TSG process dynamics and material transport through measurement of RTD (Kumar et al. 2014, 2016; Vercruysse et al. 2014). Digital video recording was also used for measuring RTD in TSG (Sayin et al. 2015a; Li et al. 2014). Real-time process measurement tools provide an advantage over manual sampling techniques for measuring RTD, considering the relatively short residence time inside a TSG (Thompson 2015). Characterization of RTD is important for optimization of equipment design, formulation, and process parameters to achieve the desired degree of axial mixing (Gao et al. 2012). Experimental RTD measurement starts by adding a tracer to the inlet of the granulator when operating at steady state. The real-time sensor is then used to measure the signal from the tracer as it emerges from the granulator. The measured RTD represents the degree of axial dispersion along the length of the granulator covered in the study. Direct measurement of RTD is also possible through particle tracking using PEPT (Lee et al. 2012). In addition to measuring RTD of the entire screw length, it also provides RTD measurements within individual screw elements if desired.

8.4 Scaling Up and Scaling Out

Scaling up of continuous twin screw granulation is less of a concern than for batch HSWG because development work and manufacturing can often be done on machines of identical size and design. Smaller or larger amounts of product are simply made by running the equipment for shorter or longer periods of time. As the residence time in the TSG is of order 5–20 s, only a few minutes is needed to reach steady state. When undertaking a DOE for development work, the TSG need only be held a one set of conditions for 10–15 minutes to collect data, and quite short run

times are needed to provide product for clinical trials. For manufacturing, the campaign time can be many hours or even several days if required. This time-based "scaling out" is common.

Nevertheless, scale-up is sometimes necessary. The smallest available TSG systems have a barrel diameter of 11.2 mm, and equipment with barrel diameters of 16 mm, 24 mm, 25 mm, and 50 mm are available from different manufacturers. In addition, during scale-out, operators may wish to increase productivity by increasing the powder feed rate when compared to development operations, sometimes with corresponding increase in screw speed. So the important question remains: How do granule properties (critical quality attributes) vary with:

1. Changes to powder feed rate and screw speed for a given screw geometry and powder.
2. Changes to powder properties for given screw geometry and operating conditions.
3. Changes in screw barrel diameter for geometrically similar twin screws?

When considering scale-up and scale-out, dimensional analysis is a powerful tool. If we consider process parameters only, such an analysis leads to the following relationship between key dimensionless groups (Osorio et al. 2016):

$$\frac{d_{50}}{D} = g_1\left(LSR, PFN, Fr, \frac{L}{D}, F_1, F_2, \ldots\right) \tag{8.1}$$

$$\varepsilon = g_2\left(LSR, PFN, Fr, \frac{L}{D}, F_1, F_2, \ldots\right) \tag{8.2}$$

where ε is granule porosity and LSR is the liquid to solid ratio:

$$LSR = \frac{\dot{m}_l}{\dot{m}_p} \tag{8.3}$$

PFN is the powder feed number:

$$PFN = \frac{\dot{m}_p}{\rho_b \omega D^3} \tag{8.4}$$

and Fr is the Froude number:

$$Fr = \frac{D\omega^2}{2g} \tag{8.5}$$

Here D is the barrel diameter, ω is the angular velocity of the shaft, L is the barrel length after wetting addition of liquid, $\dot{m}_p$ *and* $\dot{m}_l$ are the mass flow rates of the powder and liquid, respectively, ρ_b is the bulk density of the powder, and F_1, F_2, …

are a series of geometric ratios that describe the geometry of the individual screw elements and the screw configuration.

If scale-up will be undertaken by keeping important geometric ratios similar (e.g., L/D for the whole TSG and for each screw element used), Eqs. (8.1) and (8.2) can be simplified to:

$$d_{50} = D.g_1\left(LSR, PFN, Fr\right) \tag{8.6}$$

$$\varepsilon = g_2\left(LSR, PFN, Fr\right) \tag{8.7}$$

PFN is the ratio of the volumetric powder feed rate to the turn over of volume in the screw barrel as the screw turns. It is closely related to the fill level in the barrel. Given the confined nature of the TSG and the screw operation, as screw speed or powder feed rate changes, the powder residence time changes very little with the fractional fill level adjusting up and down to match the change in PFN. *Fr* is an important parameter in mixer granulators, where it dictates the powder circulating pattern and mixing. However, it is less important for the TSG where operation is typically well below the critical value for centripetal flow.

Osorio et al. (2016) showed that the key granule attributes (d_{10}, d_{50}, d_{90}, ε) were largely insensitive to *PFN* and *Fr.* This is good news for scaling out in the same equipment. Granule attributes are highly sensitive to *LSR,* and care should be taken to keep *LSR* constant when scaling out. As predicted by Eqs. (8.1) and (8.6), the granule size distribution parameters, especially d_{90}, increase with increasing barrel diameter (see Fig. 8.10). This is because the largest granule size is that which can survive breakage events. For the TSG, breakage events are defined by the screw element geometry, and the largest gap through which a granule can pass unbroken will scale with barrel diameter for geometrically similar screw elements. Thus when scaling up, it will not be possible to keep the same granule size distribution, with the size of larger lumps increasing directly with scale.

In contrast, granule porosity is scale invariant and generally higher than achieved in HSWG. For a given formulation, porosity is sensitive only to *LSR*.

In summary, performance of the TSG during scaling out is robust with respect to both process parameters and formulation properties, except for LSR. We expect this will continue until a critical value of PFN when the free volume in the most constraining screw element section becomes completely full with powder and the TSG blocks. Anecdotally, we have observed that the critical PFN can be a function of both the screw element configuration and the powder flow properties. However, there are no rigorous published studies to confirm this. For free flowing powders, stable performance is possible over a wide range of screw speeds and powder flow rates.

Scaling up to larger screw barrels presents challenges in controlling the largest granule size. However, granule porosity does not change with barrel diameter and is robust on scale-up.

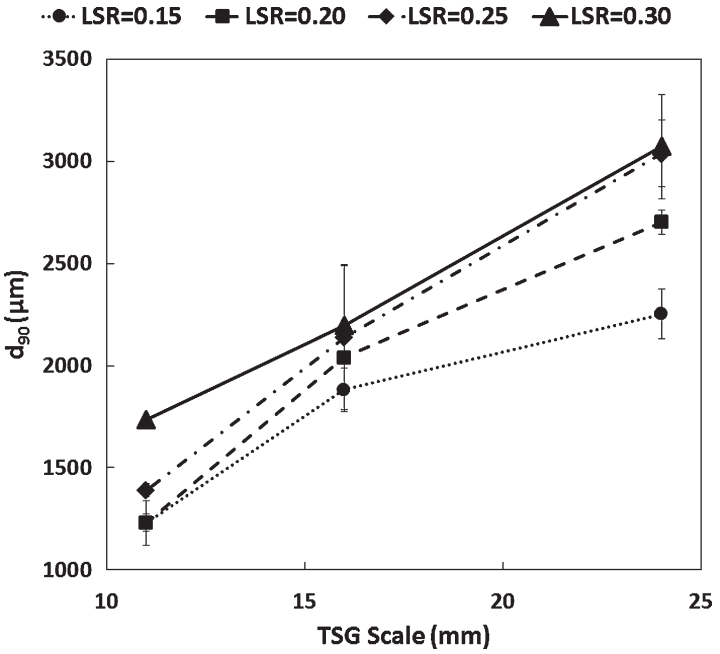


Fig. 8.10 Effect of TSG scale (barrel diameter) and *LSR* on d_{90} [$PFN = 1.30 \times 10^{-2}$] (Osorio et al. 2016)

8.5 Population Balance Model (PBM) of Twin Screw Granulation

8.5.1 A Generic Expression of the Population Balance Model

The population balance model (PBM) is a series of integrodifferential equations which describe the dynamic evolution of particles with respect to a certain attribute. The population balance model has been extensively used to model the granulating process as the particles are subjected to rate processes such as consolidation, aggregation, and breakage. For a well-mixed batch system with only one attribute (particle volume) considered, the population balance model is given by (Randolph and Larson 1971; Hounslow et al. 1988):

$$\frac{\partial n(v,t)}{\partial t} + \frac{\partial}{\partial v}\left[Gn(v,t)\right] = B(v,t) - D(v,t) \tag{8.8}$$

where $n(v, t)$ is the number density function, v is the vector representing particle property (herein volume), G is the growth rate, and $B(v, t)$ and $D(v, t)$ are the functions to denote the birth and death of particles due to aggregation, breakage, and nucleation.

Most PBMs in the literature are one-dimensional, which typically considers one particle property with regard to size (Barrasso et al. 2014). 1D PBM is limited in accuracy for predicting the real process of granulation due to the lack of variation of liquid binder and porosity. Consequently, multidimensional PBMs have been developed to incorporate the three properties: solid, liquid, and gas volume. A 3D population balance model evolves based on Eq. (8.8) (Poon et al. 2008):

$$\begin{aligned}&\frac{\partial}{\partial t}n\left(s,l,g,t\right)+\frac{\partial}{\partial s}\left[n\left(s,l,g,t\right)\frac{ds}{dt}\right]+\frac{\partial}{\partial l}\left[n\left(s,l,g,t\right)\frac{dl}{dt}\right]+\frac{\partial}{\partial g}\left[n\left(s,l,g,t\right)\frac{dg}{dt}\right]\\&=B_{\text{nuc}}\left(s,l,g,t\right)+B_{\text{agg}}\left(s,l,g,t\right)+B_{\text{break}}\left(s,l,g,t\right)-D_{\text{break}}\left(s,l,g,t\right)\end{aligned}\tag{8.9}$$

where (s, l, g) is the vector representing the solid, liquid, and gas volume of a granule and $n(s, l, g, t)$ is the population density of a granule with the three vectors over time. The three terms, with regard to partial differential functions, refer to the state change due to solid layering, liquid addition, and gas consolidation, respectively. $B_{\text{nuc}}(s, l, g, t)$ and $B_{\text{agg}}(s, l, g, t)$ denote the net rates of nucleation and aggregation. $B_{\text{break}}(s, l, g, t)$ and $D_{\text{break}}(s, l, g, t)$ denote the birth and death rate due to breakage, respectively.

Multicomponent formulations are standard in the pharmaceutical industry, and 3D PBM falls short of accounting for granule composition. Alternatively, one or more granule properties could be lumped into the remaining distributions in a reduced-order model (Hounslow et al. 2001). Another approach is proposed to add the fourth dimension (multicomponent) into the 3D population balance model. A 4D PBM was developed by Barrasso (2015) in the light of 3D PBM in Eq. (8.10). It gives:

$$\begin{aligned}&\frac{\partial}{\partial t}n\left(s_1,s_2,l,g,t\right)+\frac{\partial}{\partial l}\left[n\left(s_1,s_2,l,g,t\right)\frac{dl}{dt}\right]+\frac{\partial}{\partial g}\left[n\left(s_1,s_2,l,g,t\right)\frac{dg}{dt}\right]\\&=B_{\text{agg}}\left(s_1,s_2,l,g,t\right)+B_{\text{break}}\left(s_1,s_2,l,g,t\right)-D_{\text{break}}\left(s_1,s_2,l,g,t\right)\end{aligned}\tag{8.10}$$

where s_1 and s_2 are the volumes of two solid components in the granule. $B_{\text{agg}}(s_1, s_2, l, g, t)$, $B_{\text{break}}(s_1, s_2, l, g, t)$, and $D_{\text{break}}(s_1, s_2, l, g, t)$ are the net rates of aggregation, birth, and death rate due to breakage, respectively. Note that the nucleation and layering effects are excluded from this model.

8.5.2 *Rate Process Kernels in Twin Screw Granulation*

Design of the twin screw granulation differs from the traditional granulators where all the rate process kernels take place in a single granulation container. In the twin screw granulation however, due to its regime-separated nature, different rate process

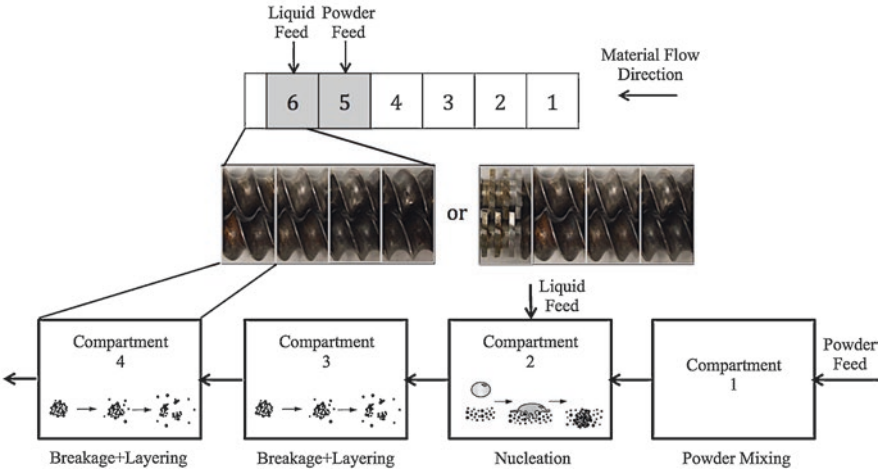


Fig. 8.11 Compartment-based rate mechanisms in a twin screw granulation (Sayin 2016)

kernels should be defined for each compartment. In Fig. 8.11, an example of the compartmental population balance modeling approach is depicted where rate mechanisms are specified based on the screw configuration.

The first compartment is the powder mixing zone in which no rate process kernel is defined. The second compartment is the conveying element where the liquid drips down to form nuclei with nucleation kernel considered. The third compartment is still the conveying element where the mixture of liquid and powder is transported. The fourth compartment is composed of either conveying element or a combination of other elements (KE or DME). In both the third and fourth compartment, the breakage and layering kernels are considered, whereas the agglomeration kernel is ignored due to the short residence time in the compartment. Considering the rate process kernels required in the population balance modeling of twin screw granulation, the key equations of nucleation, breakage, consolidation, and layering are presented below.

Nucleation Kernels

The nucleation mechanism takes place in the compartment where the droplet is fed into the powder bed. The rate of nuclei formation, B_{nuc}, is equal to the rate of droplet addition to the powder fines (Barrasso and Ramachandran 2016):

$$B_{\text{nuc}} = \frac{\dot{L}_{\text{in,powder}}\left(x,t\right)}{V_{\text{droplet}}} \tag{8.11}$$

where $L_{\text{in, powder}}(x, t)$ is the rate of liquid addition to the fine powder and V_{droplet} is the volume of a single liquid droplet, which is assumed to be constant and spherical.

It is assumed that each droplet reaching fine particles forms a single saturated granule with volume V_{nuc} and liquid volume $V_{L,\text{nuc}}$. During nucleation, solid material grows to form a granule due to the transfer from the fine powder phase. The death rate of fine powder particles, $D_{m_p,\text{nuc}}$, is equal to the product of the nucleation rate and the mass of solid associated with each nucleus (Barrasso and Ramachandran 2016):

$$D_{m_p,\text{nuc}} = \frac{\dot{L}_{\text{in,powder}}(x,t)}{V_{\text{droplet}}}\left(1-\epsilon_{\text{bed}}\right)V_{\text{nuc}}\rho_s \tag{8.12}$$

where ϵ_{bed} is the powder bed porosity and ρ_s is the density of the solid.

Breakage Kernels

The mass-based breakage equation is primarily used in the form (Hill and Ng 1995):

$$\frac{dM(v,t)}{dt} = \int_v^\infty S_M(w)\,b_M(v,w)\,dw - S_M(v)\,M(v,t) \tag{8.13}$$

where $M(v, t)$ is the mass of particles with volume v at the time t, $S_M(w)$ and $S_M(v)$ are the specific breakage rates of mass fraction of particles of volume w and v, and breakage function $b_M(v, w)$ denotes the fragment size distribution probability between the volume range v and w. $b_M(v, w) = B_{v-1,w} - B_{v-1,v}$.

The breakage kernel is composed of two important functions, namely, the selection function (S_m) and the breakage function (B_m). The selection function represents the breakage rate of particles, while the breakage function represents the breakage size distribution after collision. The selection function has been mainly studied in the application of milling. It should be noted that the majority of these selection functions are empirical and applicable to brittle material milling tests. However, their applicability to the breakage probability of granules in the wet granulation has not yet been fully tested. In terms of selection function used in the twin screw granulation, Barrasso et al. (2015a) proposed a size-dependent kernel including two fitting parameters P_1 and P_2:

$$S_M = \frac{1}{2}G_{\text{shear}}P_1^{P_2} \tag{8.14}$$

where G_{shear} is the shear rate, P_1 is a rate coefficient, and P_2 is a size-dependent exponent.

The breakage function used in the twin screw granulation is given by Barrasso and Ramachandran (2016):

$$b_M\left(v,w,v',w'\right)=2\delta\left(v-\frac{v'}{2},w-\frac{w'}{2}\right) \tag{8.15}$$

This fragment size distribution assumes that particles undergoing breakage would be broken into two equally sized particles. In another study by Sayin (2016), using the two equal volume fragments produces fairly good agreement in the conveying elements. However, a big deviation of the granule size prediction was found with the distributive mixing elements.

A recent study of breakage-isolated experiment in the twin screw granulation by Pradhan et al. (2017) indicates that the dominant breakage mechanisms in conveying and distributive mixing elements are chipping and fragmentation, respectively. This necessitates the development of different breakage kernels based on the screw elements. Furthermore, the critical granule size pertinent to the screw geometry beyond which the breakage probability increases dramatically was identified. The critical breakage size was determined as the largest available gap size by measuring the screw elements' open volume geometry.

Consolidation Kernels

Ennis et al. (1991) considered the effect of binder viscosity on granule consolidation. The amount of consolidation per collision gives:

$$\frac{\Delta x}{h}\cong 1-\exp\left(-St_v\right) \tag{8.16}$$

where Δx is the reduction in interparticle gap distance h and St_v is the Stokes number $(8\rho u_0 a/9\mu)$, where ρ and a are the density and radius of particle, respectively; u_0 is the collision velocity; and μ is the binder viscosity. Based on Eq. (8.16), increasing the binder viscosity and decreasing the particle size and the collision velocity would decrease the rate of consolidation.

Iveson et al. (1996) proposed a first-order consolidation model assuming that the consolidation rate at any time is inversely proportional to the porosity ε at that time:

$$\frac{d\varepsilon}{dt}=-k\left(\varepsilon-\varepsilon_{\min}\right) \tag{8.17}$$

where $\varepsilon_{\min}$ is the minimum porosity of the granules and k is the compaction rate constant.

A mechanistic consolidation rate expression used for twin screw granulation by Barrasso et al. (2015b) relating to the particle dynamics using DEM is given:

$$\frac{dV}{dt}=-kV_pC_{\text{impact}}\left(1-\exp\left(-St_v\right)\right) \tag{8.18}$$

where V is the particle volume, V_p and C_{impact} are per-particle pore volume and impact frequency, and St_v is the Stokes deformation number.

Layering Kernels

Layering refers to the deposition of fine powders on the surface of a wet granule. The mass balance model considering layering for a batch process follows:

$$\frac{dM_{\text{powder}}}{dt} = F_{\text{powder}}^{\text{in}} - 3G\int_0^{\infty}\frac{M(L)}{L}dL \tag{8.19}$$

and the mass balance model considering layering for continuous process gives:

$$\frac{dM_{\text{powder}}}{dt} = F_{\text{powder}}^{\text{in}} - \frac{M_{\text{powder}}}{t_R} - 3G\int_0^{\infty}\frac{M(L)}{L}dL \tag{8.20}$$

where G is the layering rate, $F_{\text{powder}}^{\text{in}}$ is the flow rate of additional powder stream in both batch and continuous process, t_R is the residence time for the powder, and $M(L)$ and L are the mass of particle and the size of particle, respectively.

Cameron et al. (2005) and Wang et al. (2006) proposed a model to calculate the layering rate according to:

$$G = G_m \frac{M_{\text{powder}}}{kM_{\text{granule}} + M_{\text{powder}}}\exp\left[-a\left(x_w - x_{wc}\right)^2\right] \tag{8.21}$$

where G_m is the maximum growth rate and M_{powder} and M_{granule} are the mass of fine powder below the lower bound of the particle classes and the mass of granule, respectively. x_w is the moisture content, while x_{wc} is the critical moisture content; k and a are fitting parameters. It is assumed that fine powder is small enough that powder particle size does not play a role in layering. The increase of granule size due to layering is divided among fine powder, liquid, and void components based on their volume fractions. The total increase in the mass of these components into the granule phase is equalized by the decrease in the mass of fine powder and liquid phases. Other forms of layering kernels, which have been applied in the twin screw granulation, could be found elsewhere (Barrasso and Ramachandran 2016).

8.6 Discrete Element Modeling (DEM) of Twin Screw Granulation

The discrete element method (DEM) was first proposed to describe the mechanical behavior of assemblies of spheres for the analysis of rock mechanics (Cundall and Strack 1979). However, DEM simulations are now widespread covering many areas, including but not limited to chemical, mineral, agricultural, and pharmaceutical engineering. DEM is based on an explicit numerical method in which the interaction of particles is monitored contact by contact. The particles are rigid but are subject to local deformation upon contact, and the resulting forces are calculated as a function of the contact law. A fundamental to DEM for the engineering application is the proper selection of the contact force models. Numerous contact models can be used to describe the particle collision behavior such as the hard-sphere model, the Hertz-Mindlin model, and the JKR model. For more details regarding the formulation and contact models, see Zhu et al. (2007). A characteristic feature of this method is that DEM can provide micromechanical insight into the particulate processing which is usually not available from experiments. Due to the rapid advancement of computational power, DEM can serve as an excellent tool to envisage the particle dynamics in the magnitude of millions of particles in the twin screw granulation. This section describes how the particle scale information by means of DEM would be utilized in the population balance model and how the coupling strategy between DEM and PBM could be implemented.

8.6.1 DEM Simulation of Residence Time Distribution

Residence time distribution characterization is a key step in the design and scale-up of twin screw granulation (Seem et al. 2015). Among the methods to measure the residence time distribution, DEM has been used to predict velocity spectrum of particles which is then used for the calculation of residence time distribution in the twin screw granulation. Barrasso and Ramachandran (2016) presented the application of DEM in estimating the residence time distribution in compartmental screw elements. It was found out that conveying elements show approximately equal residence times in each compartment. However, the kneading elements exhibited some variation in residence time along the granulator. It is worth noting that kneading elements have longer residence time than the conveying elements, indicating that the kneading elements are prone to inhibit the particle flow.

8.6.2 DEM Simulation of Particle Scale Dynamics

Apart from estimating the residence time distribution, DEM is also coupled with PBM by exploring the particle dynamics (i.e., impact frequency, impact, and collision rates) as input parameters in the rate process kernels. Barrasso and Ramachandran (2016) and Barrasso et al. (2015b) proposed a multiscale mechanistic model incorporating the mechanistic expressions for aggregation, breakage, and consolidation. In particular, these rates were evaluated based on the information from DEM such as particle velocity and collision data. In this study, nucleation and layering were omitted assuming that nuclei have formed with the focus on the later stage. The proposed DEM-PBM model was then used to predict the granulating performance in the twin screw granulator, which were consistent with experimental trends. The effects of screw configuration, material properties, and process parameters on the critical quality attribute were investigated. It was found that the screw configuration is the dominant factor in determining the rate process and product attributes. The mixing elements resulted in more aggregation, breakage, and consolidation as compared to the feed screw. However, qualitative and quantitative validation against the experimental results was lacking in this study. Considering the significant computational cost of running DEM as compared to PBM, this would negatively impact the overall time to solve the DEM-PBM coupling model. Hence, future work is recommended to develop the DEM-based rate kernels in light of particle dynamics with properly calibrated DEM contact models.

8.6.3 DEM-PBM Coupling for Twin Screw Granulation

The main advantage of PBM is the ability to provide a framework that can describe the evolution of a population of particles including subprocess kernels such as coalescence, nucleation, consolidation, and breakage. However, the detailed particle level information for these kernels, for example, spatial variation, is absent in PBM. DEM can provide this missing data of particle collision, particle residence time, and particle flux but is computationally intensive. Considering the complimentary advantages and disadvantages of the respective approaches, significant efforts have been made toward the coupling between DEM and PBM to predict various unit operations such as milling (Carvalho and Tavares 2009; Wang et al. 2012), mixing, and high shear granulation (Lee et al. 2016; Freireich et al. 2011). While there is an increasing trend of applying DEM to simulate the wet granulating process, DEM simulations and its coupling with PBM of twin screw granulator are quite limited in the literature. Barrasso and Ramachandran (2015) implemented bidirectional coupling between PBM and DEM to evaluate collision frequencies and liquid distribution as a proof of concept. Note that liquid droplets were modeled as solid particles with the properties identical to the primary powder

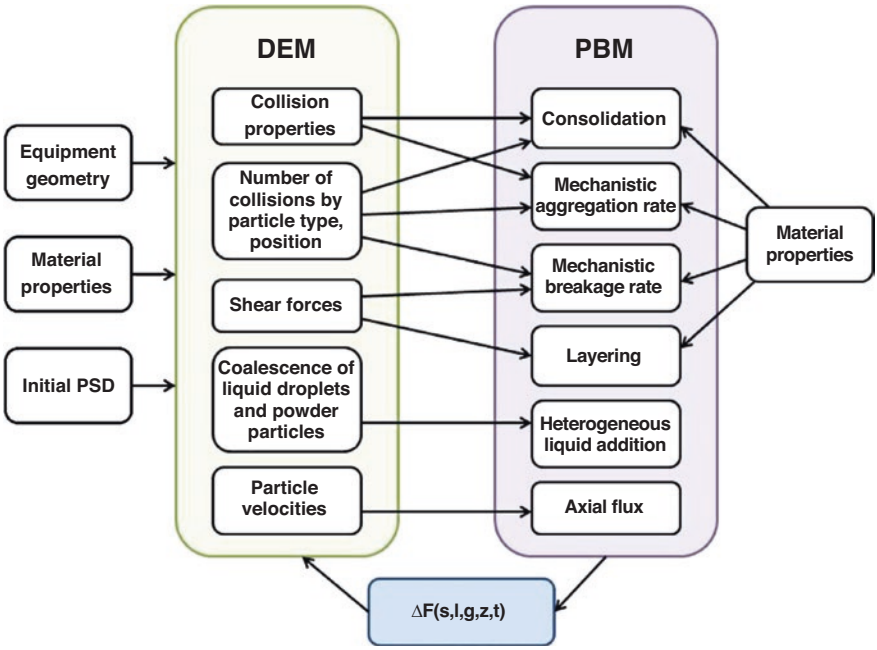


Fig. 8.12 Schematic of multiscale model and information exchange between DEM and PBM (Barrasso and Ramachandran 2015)

particles. A schematic of the coupling approach between DEM and PBM is shown in Fig. 8.12.

As shown in Fig. 8.12, the number of collisions between each pair of bins over the given time interval is recorded from DEM and then transferred to the PBM. The coupling is called two-way coupling on condition that the particle size distribution is updated back to DEM, that is, each time the PBM is solved, the corresponding change in the particle distribution is calculated in the context of DEM. The details of creating particles in DEM could be found in (Barrasso and Ramachandran 2015). The collision between different size ranges was investigated, and it was found that the collision frequency affects the aggregation rate kernel and the evolution of particle size distribution, indicating the need for bidirectional coupling between DEM and PBM. Note that the expression of kernels such as breakage and consolidation is modeled in an empirical way. The mechanistic understanding of these kernels from the perspective of particle level has not been developed in this coupling framework. Along with the advancement of computational capacity in DEM, more efforts are shifted into the development of kernels and PBM-DEM coupling methods. Even though there are many PBM-DEM coupling methods to predict the twin screw granulation (Barrasso and Ramachandran 2016; Barrasso et al. 2015b), the validation and calibration processes of these PBM-DEM models are seldom carried out.

8.7 Concluding Remarks

Twin screw granulation provides a unique approach to continuous wet granulation that gives granules with very different characteristics than those produced by batch high shear wet granulation. It has much shorter residence times, produces more porous granules, and is much less sensitive to formulation properties. The ability to go from development through to full-scale manufacture via scaling out on a time basis reduces risk and gives confidence that granule attributes can be maintained.

The confined nature of the TSG allows it to operate in regime-separated mode. Unlike batch granulation, the granulator geometry is easily changed by adjusting the choice of screw elements and their configuration. This gives almost infinite degrees of freedom, so it is very important to understand the controlling rate processes well in choosing a configuration. It is likely further improvements to screw element geometry and configuration, specifically targeted at granulation, as distinct from extrusion, are possible.

TSG is very amenable to process modeling using a population balance framework either on its own or coupled with particle-scale DEM simulations. As many of the rate processes, such as breakage, are dominated by screw element geometry, models are more robust with fewer arbitrary fitted parameters than those for batch granulation.

References

Barrasso, D.: Multi-scale modeling of wet granulation process. PhD Thesis, Rutgers University, New Jersey. (2015).

Barrasso D, Ramachandran R. Multi-scale modeling of granulation processes: bi-directional coupling of PBM with DEM via collision frequencies. Chem Eng Res Des. 2015;93:304–17.

Barrasso D, Ramachandran R. Qualitative assessment of a multi-scale, compartmental PBM-DEM model of a continuous twin-screw wet granulation process. J Pharm Innov. 2016;11:231–49.

Barrasso D, Tamrakar A, Ramachandran R. A reduced order PBM-ANN model of a multi-scale PBM-DEM description of a wet granulation process. Chem Eng Sci. 2014;119:319–29.

Barrasso D, El Hagrasy A, Litster JD, Ramachandran R. Multi-dimensional population balance model development and validation for a twin screw granulation process. Powder Technol. 2015a;270:612–21.

Barrasso D, Eppinger T, Pereira FE, Aglave R, Debus K, Bermingham SK, Ramachandran R. A multi-scale, mechanistic model of a wet granulation process using a novel bi-directional PBM-DEM coupling algorithm. Chem Eng Sci. 2015b;123:500–13.

Cameron IT, Wang FY, Immanuel CD, Stepanek F. Process systems modelling and applications in granulation: a review. Chem Eng Sci. 2005;60:3723–50.

Carvalho RM, Tavares LM. Dynamic modeling of comminution using a general microscale breakage model: Elsevier Inc; Amsterdam. 2009.

Chablani L, Taylor MK, Mehrotra A, Rameas P, Stagner WC. Inline real-time near-infrared granule moisture measurements of a continuous granulation–drying–milling process. AAPS PharmSciTech. 2011;12:1050–5.

Cundall PA, Strack ODL. A discrete numerical model for granular assemblies. Géotechnique. 1979;29:47–65.

Dhenge RM, Fyles RS, Cartwright JJ, Doughty DG, Hounslow MJ, Salman AD. Twin screw wet granulation: granule properties. Chem Eng J. 2010;164:322–9.
Dhenge RM, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw wet granulation: effects of properties of granulation liquid. Powder Technol. 2012a;229:126–36.
Dhenge RM, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw granulation: steps in granule growth. Int J Pharm. 2012b;438:20–32.
Dhenge RM, Washino K, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw granulation using conveying screws: effects of viscosity of granulation liquids and flow of powders. Powder Technol. 2013;238:77–90.
Djuric D, Kleinebudde P. Impact of screw elements on continuous granulation with a twin-screw extruder. J Pharm Sci. 2008;97:4934–42.
Djuric D, Kleinebudde P. Continuous granulation with a twin-screw extruder: impact of material throughput. Pharm Dev Technol. 2010;15:518–25.
Djuric D, Van Melkebeke B, Kleinebudde P, Remon JP, Vervaet C. Comparison of two twin-screw extruders for continuous granulation. Eur J Pharm Biopharm. 2009;71:155–60.
El Hagrasy AS, Litster JD. Granulation rate processes in the kneading elements of a twin screw granulator. AICHE J. 2013;59:4100–15.
El Hagrasy AS, Hennenkamp JR, Burke MD, Cartwright JJ, Litster JD. Twin screw wet granulation: influence of formulation parameters on granule properties and growth behavior. Powder Technol. 2013a;238:108–15.
El Hagrasy AS, Cruise P, Jones I, Litster JD. In-line size monitoring of a twin screw granulation process using high-speed imaging. J Pharm Innov. 2013b;8:90–8.
Ennis BJ, Tardos G, Pfeffer R. A microlevel-based characterization of granulation phenomena. Powder Technol. 1991;65:257–72.
Fonteyne M, Vercruysse J, Díaz DC, Gildemyn D, Vervaet C, Remon JP, De Beer T. Real-time assessment of critical quality attributes of a continuous granulation process. Pharm Dev Technol. 2013;18:85–97.
Fonteyne M, Arruabarrena J, de Beer J, Hellings M, Van Den Kerkhof T, Burggraeve A, Vervaet C, Remon JP, De Beer T. NIR spectroscopic method for the in-line moisture assessment during drying in a six-segmented fluid bed dryer of a continuous tablet production line: validation of quantifying abilities and uncertainty assessment. J Pharm Biomed Anal. 2014;100:21–7.
Freireich B, Li J, Litster J, Wassgren C. Incorporating particle flow information from discrete element simulations in population balance models of mixer-coaters. Chem Eng Sci. 2011;66:3592–604.
Gao Y, Muzzio FJ, Ierapetritou MG. A review of the Residence Time Distribution (RTD) applications in solid unit operations. Powder Technol. 2012;228:416–23.
Hapgood KP, Litster JD, Smith R. Nucleation regime map for liquid bound granules. AICHE J. 2003;49:350–61.
Hapgood KP, Lveson SM, Litster JD, Liu LX. Granulation. Handb Powder Technol. 2007;11:897–977.
Hill PJ, Ng KM. New discretization procedure for the breakage equation. AICHE J. 1995;41:1204–16.
Hounslow MJ, Ryall RL, Marshall VR. A discretized population balance for nucleation, growth, and aggregation. AICHE J. 1988;34:1821–32.
Hounslow MJ, Pearson JMK, Instone T. Tracer studies of high-shear granulation: II. Population balance modeling. AIChE J. 2001;47:1984–99.
Iveson SM, Litster JD, Ennis BJ. Fundamental studies of granule consolidation Part 1: effects of binder content and binder viscosity. Powder Technol. 1996;88:15–20.
Keleb EI, Vermeire A, Vervaet C, Remon JP. Twin screw granulation as a simple and efficient tool for continuous wet granulation. Int J Pharm. 2004;273:183–94.
Kumar A, Vercruysse J, Toiviainen M, Panouillot PE, Juuti M, Vanhoorne V, Vervaet C, Remon JP, Gernaey KV, De Beer T, Nopens I. Mixing and transport during pharmaceutical twin-screw wet granulation: experimental analysis via chemical imaging. Eur J Pharm Biopharm. 2014;87:279–89.

Kumar A, Alakarjula M, Vanhoorne V, Toiviainen M, De Leersnyder F, Vercruysse J, Juuti M, Ketolainen J, Vervaet C, Remon JP, Gernaey KV, De Beer T, Nopens I. Linking granulation performance with residence time and granulation liquid distributions in twin-screw granulation: an experimental investigation. Eur J Pharm Sci. 2016;90:25–37.

Lee KT, Ingram A, Rowson NA. Twin screw wet granulation: the study of a continuous twin screw granulator using Positron Emission Particle Tracking (PEPT) technique. Eur J Pharm Biopharm. 2012;81:666–73.

Lee KF, Dosta M, Mcguire AD, Wagner W, Heinrich S, Kraft M, Street P, Street P. Development of a multi-compartment population balance model for high-shear wet granulation with Discrete Element Method New Museums Site. Comput Chem Eng. 2016;99:171–84.

Li H, Thompson MR, O'Donnell KP. Understanding wet granulation in the kneading block of twin screw extruders. Chem Eng Sci. 2014;113:11–21.

Litster JD, Ennis BJ. The science and engineering of granulation processes: Kluwer Academic; 2004.

Liu Y, Thompson MR, O'Donnell KP. Function of upstream and downstream conveying elements in wet granulation processes within a twin screw extruder. Powder Technol. 2015;284:551–9.

Mu B, Thompson MR. Examining the mechanics of granulation with a hot melt binder in a twin-screw extruder. Chem Eng Sci. 2012;81:46–56.

Osorio J, et al. Scaling of continuous twin screw wet granulation. Am Inst Chem Eng. 2016;63:921–32.

Poon JM-H, Immanuel CD, Doyle FJ III, Litster JD. A three-dimensional population balance model of granulation with a mechanistic representation of the nucleation and aggregation phenomena. Chem Eng Sci. 2008;63:1315–29.

Pradhan SU, Sen M, Li J, Litster JD, Wassgren CR. Granule breakage in twin screw granulation: effect of material properties and screw element geometry. Powder Technol. 2017;315:290–9.

Randolph AD, Larson MA. Chapter 3 – the population balance. In: Theory of particulate processes; 1971, Elsevier, Amsterdam. p. 41–63.

Saleh MF, Dhenge RM, Cartwright JJ, Hounslow MJ, Salman AD. Twin screw wet granulation: effect of process and formulation variables on powder caking during production. Int J Pharm. 2015;496:571–82.

Sayin R. Mechanistic studies of twin screw granulation. PhD Thesis, Purdue University, Indiana. 2016.

Sayin R, El Hagrasy AS, Litster JD. Distributive mixing elements: towards improved granule attributes from a twin screw granulation process. Chem Eng Sci. 2015a;125:165–75.

Sayin R, Martinez-Marcos L, Osorio JG, Cruise P, Jones I, Halbert GW, Lamprou DA, Litster JD. Investigation of an 11 mm diameter twin screw granulator: screw element performance and in-line monitoring via image analysis. Int J Pharm. 2015b;496:24–32.

Seem TC, Rowson NA, Ingram A, Huang Z, Yu S, de Matas M, Gabbott I, Reynolds GK. Twin screw granulation – a literature review. Powder Technol. 2015;276:89–102.

Tan MXL, Hapgood KP. Foam granulation: effects of formulation and process conditions on granule size distributions. Powder Technol. 2012;218:149–56.

Thiele W. Twin-screw extrusion and screw design. In Pharmaceutical Extrusion Technology (Chapter 4), Edited By Isaac Ghebre-Sellassie, Charles E. Martin, Feng Zhang, James DiNunzio, Taylor and Francis, London. 2003.

Thompson MR. Twin screw granulation – review of current progress. Drug Dev Ind Pharm. 2015;41:1223–31.

Thompson MR, Sun J. Wet granulation in a twin-screw extruder: implications of screw design. J Pharm Sci. 2010;99:2090–103.

Van Melkebeke B, Vervaet C, Remon JP. Validation of a continuous granulation process using a twin-screw extruder. Int J Pharm. 2008;356:224–30.

Vercruysse J, Toiviainen M, Fonteyne M, Helkimo N, Ketolainen J, Juuti M, Delaet U, Van Assche I, Remon JP, Vervaet C, De Beer T. Visualization and understanding of the granulation liquid mixing and distribution during continuous twin screw granulation using NIR chemical imaging. Eur J Pharm Biopharm. 2014;86:383–92.

Wang FY, Ge XY, Balliu N, Cameron IT. Optimal control and operation of drum granulation processes. Chem Eng Sci. 2006;61:257–67.
Wang MH, Yang RY, Yu AB. DEM investigation of energy distribution and particle breakage in tumbling ball mills. Powder Technol. 2012;223:83–91.
Zhu HP, Zhou ZY, Yang RY, Yu AB. Discrete particle simulation of particulate systems: theoretical developments. Chem Eng Sci. 2007;62:3378–96.

Chapter 9
Continuous Fluidized Bed Drying: Advanced Modeling and Experimental Investigations

Ibrahim Alaathar, Stefan Heinrich, and Ernst-Ulrich Hartge

Abstract Fluidized bed drying is a technology which is widely applied in industry for the drying of particulate solids. One of the major advantages of fluidized bed system results from the fact that due to high heat, mass and momentum transfer, and intensive solids, mixing a good temperature control is possible, which allows an effective drying of heat-sensitive materials as, for example, food or pharmaceutical products, which are very often also porous and hygroscopic. A disadvantage of the nearly perfect solids mixing in the fluidized bed is the wide residence time distribution of the solids, that is, some particle will leave the dryer already a few seconds after they have been fed to it, while other particles stay for very long times inside the dryer, which may result in a wide moisture distribution of the solids at the outlet of the dryer.

Based on numerous works from the literature, a drying model for continuous fluidized bed drying has been developed and implemented as a module within the framework of the stationary flowsheet simulation program *SolidSim*, which allows the simulation of the drying process for many different solids, liquids, and gases.

In addition to the residence time distribution described by a population model, the influence of particle size distribution and the moisture distribution of the feed material was taken into account. The module computes the humidity and temperature of the gas versus the bed height and the moisture and temperature distribution of the solids by taking into account their distributed properties such as for the residence time distribution, particle size and initial moisture. Additionally, also the drying of mixtures of particles with different types of sorptive behavior and drying kinetics is possible while assuming that no particle segregation occurs. Furthermore,

I. Alaathar
Andritz Fliessbett Systeme GmbH, Ravensburg, Germany

S. Heinrich (✉) · E.-U. Hartge
Institute of Solids Process Engineering and Particle Technology,
Hamburg University of Technology, Hamburg, Germany
e-mail: stefan.heinrich@tuhh.de

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_9

the model has been extended by a simple approach to simulate the drying in an elongated fluidized bed dryer.

For the validation of the model, the main operating parameters were varied and compared with experiments. The model can calculate the moisture distribution at the dryer outlet, but due to technical restriction, only the average moisture of the particles was measured and compared with the model prediction.

Keywords Continuous process · Drying · Fluidized bed · Particle size distribution · Moisture distribution

List of Symbols

A	Surface area	[m^2]
c	Specific heat capacity	[J/kg K]
c_v	Solids volume concentration	[−]
d	Diameter	[m]
d_v	Bubble diameter	[m]
h	Specific enthalpy	[J/kg]
Δh_v	Specific enthalpy of evaporation	[J/kg]
$\dot{H}$	Enthalpy flow rate	[J/s]
$\dot{m}$	Drying rate	[g/m^2s]
M	Mass	[kg]
$\dot{M}$	Mass flow rate	[kg/s]
M	Molar mass	[kg/kmol]
N	Number of density	[−]
$\dot{Q}$	Heat flow rate	[W]
$\Delta Q_{3,k}$	Mass fraction of solids in particle class size k	[−]
$\Delta Q_{3,f}$	Mass fraction of solids in moisture class size f	[−]
T	Temperature	[K]
u	Velocity	[m/s]
$\dot{V}$	Visible bubble flow rate based on unit bed area	[m/s]
X	Particle moisture content (dry basis)	[kg_w/kg_s]
X_{eq}	Particle equilibrium moisture content (dry basis)	[kg_w/kg_s]
X_{cr}	Critical moisture content (dry basis)	[kg_w/kg_s]
Y	Gas moisture content (dry basis)	[kg_w/kg_g]
Y_{eq}	Gas equilibrium moisture content	[kg_w/kg_g]
z	Bed height coordinate	[m]

Greek Letters

α	Heat transfer coefficient	[W/m^2]
α	Crucial parameter	[−]
α_w	Water activity	[−]
β	Mass transfer coefficient	[m/s]

ε	Porosity	
η	Normalized particle moisture content	[−]
θ	Factor Eq. (9.40)	
ϑ	Temperature	[°C]
λ	The mean bubble life time	
$\dot{\nu}$	Normalized single particle drying rate (dimensionless)	[−]
$\dot{\nu}'$	Modified normalized single particle drying rate (dimensionless)	[−]
ξ	Normalized bed height	[−]
ρ	Density	[kg/m^3]
τ	Residence time	[s]
$\overline{\tau}$	Mean residence time	[s]
υ	Ratio of bubble to total gas flow rate	[−]
φ	Relative humidity of the gas (dimensionless)	[−]
ψ	Factor Eqs. (9.20) and (9.39)	

Subscripts

b	Bubble
cr	Critical
f	Index of discretized inlet moisture content coordinate
i	Index of discretized residence time coordinate
k	Index of discretized particle size coordinate
g	Gas
mf	Minimal fluidization
or	Orifice
p	Particle
s	Suspension
v	Vapor

Abbreviation

CSTR	Continuous stirred tank reactor
PFTR	Plug flow tubular reactor
RTD	Residence time distribution

9.1 Introduction

In numerous industrial applications, liquid has to be separated from the wet solid. This separation can be accomplished either mechanically without altering the liquid phase or thermally by adding heat to the wet material, which causes a phase

transformation from liquid to steam. While in the first case the driving force of the separation could be gravity, pressure, or centrifugal force depending on the machine design, the driving force in the other case is the pressure gradient of the steam at the interface between solids and the drying medium. In many applications mechanical drying or dewatering precedes thermal drying because by dewatering a large portion of the moisture can be more economically removed. Final drying to the required moisture must be carried out thermally.

This article emphasizes on thermal drying only.

Depending on the particle size and the kind of feeding, the preferred drying method is either spray drying (particle size 100–500 μm), flash drying (particle size 10–3000 μm), or fluidized bed drying (particle size 50–5000 μm) (Chandran et al. 1990). While liquid material is directly fed to spray and flash dryers, the fluidized bed dryer can only be fed with wet solids.

Fluidized bed drying is a technology, which is widely applied in industry for the drying of particulate solids. The fluidized bed drying can be carried out batchwise or continuously. The batch fluidized bed dryer is particularly suitable for smaller quantities and for heat-sensitive materials. The product has a uniform quality due to the homogeneity of the bed for batch operations. In continuous fluidized bed dryers, the fluidized bed exerts a strong mixing characteristic called "back-mixing" with the consequence that the product at the discharge of the dryer has the same characteristics as the product inside the entire dryer itself. While the continuous drying is widely used in chemical and food industry, the pharmaceutical industry still heavily relies on batch drying.

Our intention in this article is to provide an overview of continuous drying technologies describing recent advances in modeling, experimental investigations, and equipment design for these processes.

One of the major advantages of fluidized bed systems results from the fact that the high turbulence created in the bed provides high heat and mass transfer, as well as a good mixing of the solids inside the bed. Another important advantage is a good temperature control, which allows an effective drying of heat-sensitive materials as food-based or pharmaceutical products. A disadvantage of the nearly perfect solids mixing in the fluidized bed is the wide residence time distribution of the solids, that is, some particles will leave the dryer already within a few seconds after they have been fed, while other particles may stay for a very long time inside the dryer, which results in a wide moisture distribution of the solids at the outlet of the dryer as has been shown experimentally by Kettner et al. (2006) and by Peglow et al. (2007).

In the agricultural, food, and pharmaceutical industry, the main target of drying is to decrease the moisture content and then the water activity α_w in the products, which depends directly on the moisture content, in order to enable their storage at ambient temperature and to improve their shelf lives. By knowing the water activity as depicted in Fig. 9.1, it is possible to predict which microorganisms could be potential sources of spoilage. In addition to influencing microbial spoilage, water activity may play a significant role in determining the activity of enzymes and vitamins in foods and may have a major impact on their color, taste, and aroma. It may also significantly impact the potency and consistency of pharmaceuticals.

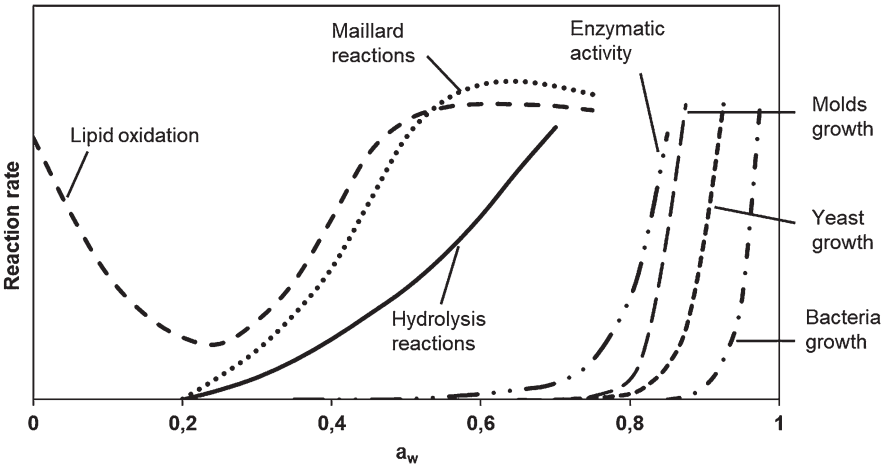


Fig. 9.1 Water activity and stability map for food material (Tsotsas and Mujumdar 2009)

The presence of a product moisture distribution during continuous fluidized bed drying cannot be avoided, but for certain products it is very important to narrow the moisture distribution to be well within the safe limit of the product.

Other factors that decisively influence the distribution of the product's moisture at dryer outlet are the particle size distribution and the moisture distribution of the feed material. Fine particles will be dryer on average at the dryer outlet than coarser particles (Kozanoglu et al. 2011; Palánczy 1983; Zahed et al. 1995). Also the particles with low inlet moisture content will be drier at the outlet than the ones with higher initial moisture (Garnavi et al. 2006).

Most modeling approaches describe the solids in the fluidized bed as ideally mixed, that is, assume that all particles have the same properties. These models are based on average properties of the solids, that is, average particle size, average moisture, and average residence time (e.g., Vanecek & Markvart 1966; Kannan et al. 1995).

In contrast to traditional modeling approaches, drying processes can nowadays be simulated with the help of computers and numerical methods. The recent modeling approaches describe the solid in the fluidized bed by population balances. In this case, the solid phase is divided into populations, and the properties of each particle population (temperature and moisture content) are a function of their residence time, particle size, and initial moisture content.

Burgschweiger and Tsotsas (2002) developed a model of the instationary drying of solids in fluidized beds which takes the distribution of residence times into account. Due to the complexity of the model of Burgschweiger and Tsotsas (2002), Cunäus et al. (2008) and Peglow et al. (2011) developed a simplified analytical solution of the population equation. Alaathar et al. (2013) extended the model developed by Burgschweiger and Tsotsas (2002) to consider the particle size distribution, the initial moisture distribution, as well as the drying of solid mixture in addition to the residence time distribution.

Because the drying process is normally not a stand-alone process but rather a final processing (finishing) step in a series of unit operations, the implementation of a dryer model as an equipment module for a flowsheet simulation program will not only enable the prediction of the drying process itself but also help optimizing the entire process chain.

In this work a generally applicable stationary model was implemented into the flowsheet simulation program SolidSim (now ASPEN). The model was generalized to allow its application to arbitrary combinations of solids and liquids and adjusted to consider the particle size and moisture distribution of the feed material and to be able to calculate a mixture of solids. Since SolidSim is a steady-state simulation tool, the formulation had to be adjusted accordingly.

Finally, the inclusion of the model as a unit operation into the flowsheet simulation program SolidSim allows the simulation of the interdependency of the operation of a dryer with preceding or subsequent solids processing units.

9.2 Kinetics of Drying

During drying, heat and mass transfer processes occur simultaneously as the heat that is introduced and transferred to the wet solid is used to evaporate the moisture, which in turn has to be transported outside of the system. In convective drying the same medium is used as heat carrier and vapor carrier from the system.

During drying the weight of the sample will decrease with time due to evaporation. The illustration of weight reduction versus time is called a drying curve (Fig. 9.2).

The other common means of illustrating the drying process is to plot the drying velocity (dx/dt) or the drying rate $\dot{m}_{pg}$ over time, which defines the evaporated moisture per unit of time and surface area of the particle, versus time (Fig. 9.3). It is thus possible to clearly distinguish between the different drying mechanisms and the respective drying regimes.

Generally, in a very short time period, the temperature of the solid will increase or decrease, based on the initial temperature of the solid. This period will end shortly

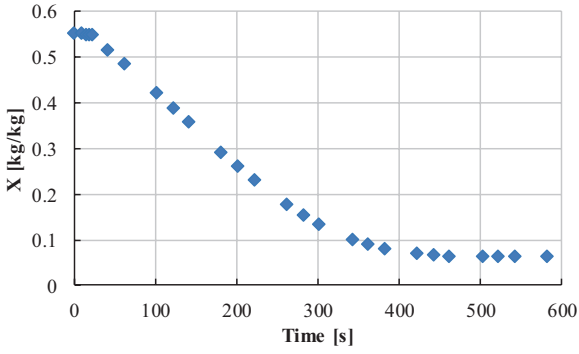


Fig. 9.2 Measured drying curve of a single particle of aluminum oxide (d_p = 1.8 mm, T_g = 41 °C, φ = 20% and u_g = 2.8 m/s)

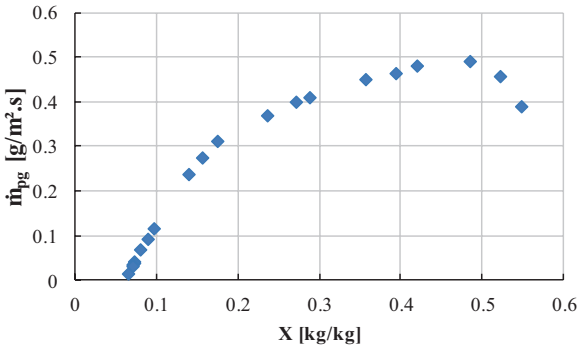


Fig. 9.3 Measured drying rate as function of the moisture content for a single particle of (d_p = 1.8 mm, T_g = 41 °C, φ = 20% and u_g = 2.8 cm/s) aluminum oxide

once reaching the wet bulb temperature on the particle surface (single particle) or the adiabatic saturation temperature in a deep fluidized bed. The adiabatic saturation point is reached when adding liquid water to humid air adiabatically until saturation is achieved, whereas the wet bulb temperature is a temperature of a small wet object (single particle) exposed to a sufficiently high air flow. In the water-air system, the two temperatures are almost the same at low temperatures. The moisture content of particles falls slightly during this phase.

In the following phase, or so-called first drying phase, the moisture content falls off linearly at a constant drying velocity and constant solid temperature (non-hygroscopic material). The drying rate in this phase can be written as:

$$\dot{m}_I = \rho_g \beta_{pg} \left(Y^* - Y \right) \tag{9.1}$$

In the equation above, ρ_g is the density of the drying medium, β_{pg} is the mass transfer coefficient, and Y^* is the absolute humidity at the wet bulb temperature or Y_{wb} in case of a single particle and the humidity at saturation in the other case.

For hygroscopic materials the drying velocity will also decrease during the first drying phase due to the reduction of the partial vapor pressure as the moisture content during drying decreases. This characteristic can be clearly seen in Fig. 9.3.

The drying rate in the first drying period can be written as follows while considering the sorption equilibrium $Y_{eq}(T_p, X_p)$ as:

$$\dot{m}_I = \rho_g \beta_{pg} \left(Y_{eq} \left(T_p, X_p \right) - Y \right) \tag{9.2}$$

The first drying phase ends when the critical moisture content X_{cr} is reached.

When reaching the critical moisture content, the outlet surface of the solid will be totally dry. When falling below the critical moisture content, the wet surface forming the border between dry and wet parts of the material where the liquid evaporates shifts toward the inside of the solid. The energy for evaporation must then be passed through an already dry layer of material with the thermal conductivity λ. On the other hand, the evaporated moisture must be transported from the wet surface to the outer surface across the dried layer. Compared to the evaporation at the solid surface the heat and diffusion resistances reduce progressively the drying rate as the

depth of the evaporation surface increases (Gnielinski 2009). Thus, the drying rate begins to decrease.

Van Meel (1958) used a normalized drying rate to consider the transport resistance in the second drying phase. The normalized drying rate is defined as the ratio of the actual drying rate $\dot{m}_{pg}$ to that in the first drying phase calculated by Eq. (9.1).

$$\dot{v}(\eta) = \frac{\dot{m}_{pg}}{\dot{m}_{I}} \tag{9.3}$$

η is the normalized moisture content:

$$\eta = \frac{X - X_{eq}}{X - X_{cr}} \tag{9.4}$$

with X_{eq} is the equilibrium moisture content and X_{cr} is the critical moisture content.

The normalized drying rate equals 1 when the normalized moisture content is higher than 1 and varies between 0 and 1 when η is less than 1.

The normalized drying rate according to van Meel (1958) was developed for non-hygroscopic material. Burgschweiger et al. (1999) modified the normalization concept according to van Meel (1958) to cover hygroscopic materials. Because the drying process rate is not constant in the first drying phase, and decreases with continuing drying, it is therefore important to distinguish between the reduction of the driving force due to the reduction of the partial steam pressure (hygroscopicity) and the transport resistance in the second drying phase. For this purpose the authors used the drying rate calculated according to Eq. (9.2) which already considers the reduction of steam pressure due to the hygroscopicity to obtain the normalized drying curve $\dot{v}'(\eta)$.

The drying rate in the second drying phase can be then written as:

$$\dot{m}_{II} = \rho_g \beta_{pg} \left(Y_{eq}(T_p, X_p) - Y \right) \dot{v}'(\eta) \tag{9.5}$$

Figure 9.4 illustrates schematically the difference between the two concepts. The calculated curve takes only into account the hygroscopic effect, that is, the

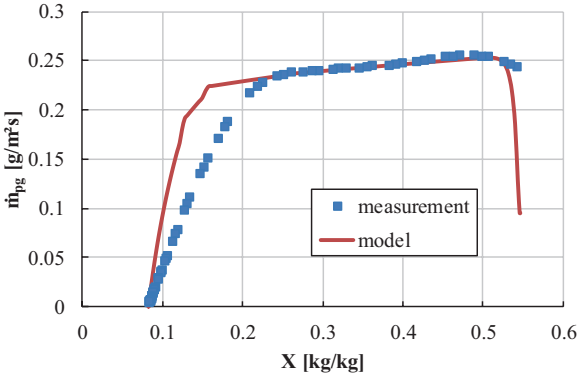


Fig. 9.4 Measured and calculated drying curve for a single particle of γ-alumina (only the sorptive effect was considered by calculation) (d_p = 3 mm, T_g = 41 °C, φ = 20% and u_g = 2.8 cm/s)

normalized drying rate being set to $\dot{v}'(\eta) = 1$. In contrast, the measured curve additionally includes the transport resistances in the second drying phase. So the distance between the two curves represents the moisture transport resistance in the second drying phase.

The second drying phase ends for non-hygroscopic materials with complete drying, with the hygroscopic equilibrium moisture for hygroscopic materials.

While the drying in the first drying phase depends essentially on the heat and mass transfer that is controlled by the external conditions such as temperature, humidity, and velocity of the drying medium, the drying rate in the second drying phase depends mainly on the heat and moisture transport resistances inside the dried solids layer.

9.3 Modeling

9.3.1 Classical Modeling

A schematic of a continuous fluidized bed dryer is depicted in Fig. 9.5.

The wet particles been discharged from a dewatering machine such as a filter or a centrifuge are transported into the dryer at a constant flow rate by means of, for example, scatterers, spreaders, and vibrating conveyors. Conveying helps to spread the wet solid over the feeding zone of the dryer and contributes to a slight delumping of wet agglomerates. The incoming wet product is instantly diluted to a noncritical moisture content through the action of the considerable back-mixing

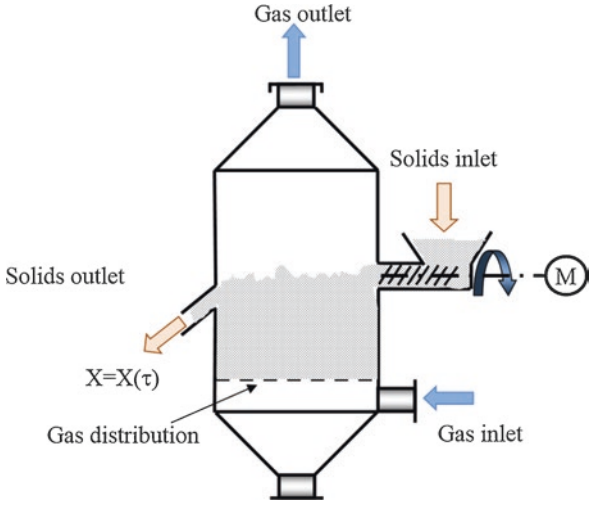


Fig. 9.5 Well-mixed continuous fluidized bed dryer

effect within the fluidized layer of material between already dried material and the wet product, which allows to instantly fluidize the wet product by means of hot air. Back-mixing is generally being observed in fluidized beds and instantly renders the product uncritical with respect to its fluidization properties and also ensures a uniform temperature profile within the bed. The product can be discharged through an overflow weir or a rotary valve if tightness is required.

Conventional modeling approaches describe the solids in the fluidized bed as ideally mixed with average properties, that is, average particle size, average moisture, and average residence time. Such a model will only predict the average moisture content of the solids in the dryer.

Mass Balance

Through continuous operation the solids enter and leave the dryer at a constant rate, that is, the total mass inside of the dryer will stay constant. The same is valid for the drying medium – the air or a gas – as long as the apparatus is gas-tight.

$$\dot{M}_{\mathrm{p,in}} - \dot{M}_{\mathrm{p,out}} + \dot{M}_{\mathrm{g,in}} - \dot{M}_{\mathrm{g,out}} = 0 \tag{9.6}$$

The moisture content of the solids and the humidity of the air can be written as follows:

$$X = \frac{M_{\mathrm{H_2O}}}{M_{\mathrm{drysolid}}} \tag{9.7}$$

$$Y = \frac{M_{\mathrm{H_2O}}}{M_{\mathrm{dryair}}} = \frac{M_{\mathrm{v}}}{M_{\mathrm{g}}} = \frac{\mathrm{Mass\,of\,vapor}}{\mathrm{Mass\,of\,dry\,air}} \tag{9.8}$$

Applying Eqs. (9.7) and (9.8) in Eq. (9.6) results in:

$$\dot{M}_{\mathrm{solid}} \left(X_1 - X_2 \right) = \dot{M}_{\mathrm{g}} \left(Y_2 - Y_1 \right) \tag{9.9}$$

By measuring the relative humidity, temperature, and the pressure of the drying medium at the outlet of the dryer, the mean moisture content of the dried material X_2 can easily be determined.

The absolute humidity of the gas can be written as:

$$Y = \frac{\tilde{M}_{\mathrm{v}}}{\tilde{M}_{\mathrm{g}}} \cdot \frac{p_{\mathrm{v}}}{P - p_{\mathrm{v}}} \tag{9.10}$$

In this equation p_{v} is the partial pressure of the vapor, and $\tilde{M}_{\mathrm{v}}$ and $\tilde{M}_{\mathrm{g}}$ are the molar masses of vapor and drying gas, respectively.

Defining the relative humidity as:

$$\varphi = \frac{p_v}{p_v^*} \tag{9.11}$$

and substituting into Eq. (9.10), the expression for Y can then be written as:

$$Y = \frac{\tilde{M}_v}{\tilde{M}_g} \cdot \frac{\varphi p_v^*}{P - \varphi p_v^*} \tag{9.12}$$

where P is the system pressure.

At saturation, the relative humidity is equal to 1 and the partial pressure of the moisture is equal to the saturation pressure:

$$Y^* = \frac{\tilde{M}_v}{\tilde{M}_g} \cdot \frac{p_v^*}{P - p_v^*} \tag{9.13}$$

Equation (9.13) is only valid for non-hygroscopic materials, and in this case the saturation pressure can be calculated with the Antoine equation. For hygroscopic materials the equilibrium vapor pressure is lower than when calculated with the Antoine equation, and Eq. (9.13) can be written as:

$$Y_{eq} = \frac{\tilde{M}_v}{\tilde{M}_g} \cdot \frac{p_{eq}\left(T_p, X\right)}{P - p_{eq}\left(T_p, X\right)} \tag{9.14}$$

with $p_{eq}(T_p, X)$ being the partial steam pressure at equilibrium, which can be calculated as follows:

$$p_{eq}\left(T_p, X\right) = \varphi_{eq}\left(T_p, X\right) p_{sat}\left(T_p\right) \tag{9.15}$$

The equilibrium relative humidity $\varphi_{eq}(T_p, X)$ can be obtained from the sorption isotherm.

To avoid condensation in the dryer hood or "freeboard" in technical terms as well as in any subsequent exhaust gas handling equipment (such as baghouse filters, cyclones, etc.), the humidity of the exhaust gas must be less than the value at saturation.

Energy Balance

$$\dot{H}_{p,in} - \dot{H}_{p,out} + \dot{H}_{g,in} - \dot{H}_{g,out} = 0 \tag{9.16}$$

The enthalpy of the wet solids results from the sum of the enthalpies of both phases:

$$\dot{H}_{p,in} = \dot{M}_p h_{p,in} = \dot{M}_p \left(h_p + h_l\right) = \dot{M}_p c_{p,p} \vartheta + \dot{M}_l c_{p,l} \vartheta = \dot{M}_p \vartheta \left(c_{p,p} + X c_{p,l}\right) \tag{9.17}$$

The enthalpy of the wet air results from:

$$h = h_g + Yh_v = c_{p,g}\vartheta + Y\left(c_{p,v}\vartheta + \Delta h_v\right) \tag{9.18}$$

Where the indices p, l, and v refer to solid particle, liquid, and vapor, respectively, h_v is the enthalpy of vapor and Δh_v is the enthalpy of evaporation.

By applying Eqs. (9.17) and (9.18) in Eq. (9.16), the mean temperature of the product can be calculated.

9.3.2 Advanced Modeling

The classical modeling as mentioned above considers the dryer as a closed box and focuses only on the phase boundaries or interfaces. Fluidized bed drying involves processes on different scales, that is, the heat and mass transfer between the gas and the single particle on the microscale on one hand and the more macroscopic heat and mass transfer between the bubble and the suspension phase in the fluidized bed on the other. Processes on both scales have to be modeled and to be combined in the fluidized bed dryer model.

The fluidized bed model used in this work is based on a two-phase model dividing the fluidized bed into a solid free bubble phase and a suspension phase consisting of the particles and a portion of the gas flowing through the suspension at the minimal fluidization velocity u_{mf}.

A schematic representation of the model is shown in Fig. 9.6.

The following assumptions were made in the model:

- The bubble phase is particle-free, and the gas in the bubble phase flows in plug flow mode.
- The particles in the suspension are perfectly mixed.
- The gas in the suspension phase flows in plug flow mode.
- Solid particles are added and removed at a constant rate.
- Heat and mass are transferred between the particles and the surrounding suspension gas.
- The bubble fraction is dependent on the bed height and calculated according to Hilligardt and Werther (1986) using the model developed by Werther and Wein (1994).
- The falling drying rate in the second drying period can be described as a modified normalized drying curve.
- There is no entrainment of solids.

Gas Phase Balances

As shown in Fig. 9.6, the gas phase is divided into a suspension phase and a bubble phase.

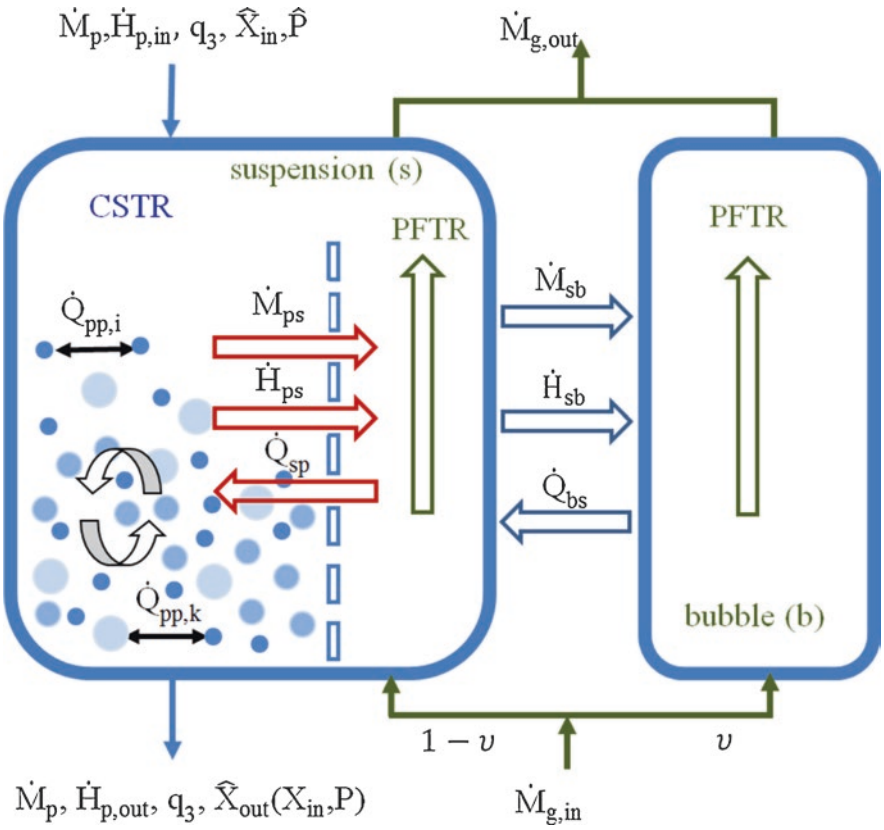


Fig. 9.6 The fluidized bed model

Suspension phase: the vapor loading Y_s of the gas in the suspension phase is changed by the mass flow $\dot{M}_{ps}$ of vapor transferred from the particles to the gas and by the mass flow $\dot{M}_{sb}$ of vapor exchanged between the bubble phase and the suspension gas:

$$(1-\upsilon)\dot{M}_g \frac{\partial Y_s}{\partial \xi} = \frac{\partial}{\partial \xi}\left(\dot{M}_{ps} - \dot{M}_{sb}\right) \tag{9.19}$$

In this equation υ is the bubble fraction, $\dot{M}_g$ denotes the total gas mass flow through the bed, and ξ is the height z above the gas distributor normalized with the bed height.

The volume flow ratio of the bubble υ can be calculated according to Hilligardt and Werther (1986)

$$\upsilon = \psi \frac{u - u_{\mathrm{mf}}}{u} \tag{9.20}$$

Here u denotes the superficial gas velocity, and the parameter ψ is a factor that describes the deviation of the visible bubble flow rate from the two-phase theory.

Similar to the vapor loading of the suspension gas, the specific enthalpy h_{s} of the suspension gas changes with height:

$$\left(1-\upsilon\right)\dot{M}_{\mathrm{g}}\frac{\partial h_{\mathrm{s}}}{\partial \xi} = \frac{\partial}{\partial \xi}\left(\dot{H}_{\mathrm{ps}} - \dot{Q}_{\mathrm{sp}} + \dot{Q}_{\mathrm{bs}} - \dot{H}_{\mathrm{sp}}\right) \tag{9.21}$$

with $\dot{H}_{\mathrm{ps}}$ denoting the enthalpy flow, which is due to the transfer of vapor from the particle to the suspension gas, $\dot{Q}_{\mathrm{sp}}$ is the heat flux from the suspension gas to the particle, $\dot{H}_{\mathrm{sb}}$ and $\dot{Q}_{\mathrm{bs}}$ are the respective enthalpy and heat flows exchanged between bubble phase and suspension gas.

Bubble phase: the gas in the bubble phase exchanges vapor, heat, and enthalpy with the suspension gas only. Thus the balances become:

$$\upsilon\dot{M}_{\mathrm{g}}\frac{\partial Y_{\mathrm{b}}}{\partial \xi} = \frac{\partial \dot{M}_{\mathrm{sb}}}{\partial \xi} \tag{9.22}$$

For the vapor loading:

$$\upsilon\dot{M}_{\mathrm{g}}\frac{\partial h_{\mathrm{b}}}{\partial \xi} = \frac{\partial}{\partial \xi}\left(\dot{H}_{\mathrm{sb}} - \dot{Q}_{\mathrm{bs}}\right) \tag{9.23}$$

Solid Phase Balance

Population balances have to be solved for the solid phase in order to describe the moisture at the outlet of the dryer dependent on the particle size and on the residence time. It is assumed, that the state variables of every particle, that is, its moisture content X and temperature T_{p}, are determined by its size, its inlet moisture content, and its residence time in the dryer, which means that the particles with a certain residence time having the same particle size and the same inlet moisture content will always have the same temperature and the same moisture content. Furthermore, it is assumed, that there is no classification occurring within the bed. This assumption is valid as long the particle size distribution is not too broad and the fluidization velocity is sufficiently high. In this case the mixing action of the bubbles is dominating the segregation tendencies. While the direct exchange of moisture between the particles is neglected, the transport of heat between the particles

belonging to different classes is taken into account. Finally, it is assumed, that particles undergo no shrinking, fragmentation, or abrasion, that is, there is no transfer of particles between the particle size classes.

For the computational treatment, the three-dimensional distribution of particle size, residence time, and inlet moisture content is discretized into different size, inlet moisture, and residence time classes. The size classes are in the following indexed by k, the inlet moisture by f, and the residence time classes by i.

With the assumptions as stated above, the equations describing the residence time distribution, moisture content, and enthalpy for every size, inlet moisture content, and residence time class under steady-state conditions for the component c are:

$$M_{\mathrm{p}} \cdot \frac{dN_{\mathrm{i}}}{d\tau} = -\dot{M}_{\mathrm{p,in}} \cdot N_{\mathrm{i}} \tag{9.24}$$

$$\Delta Q^{\mathrm{c}}_{3,k} \cdot \Delta Q^{\mathrm{c}}_{3,f} \cdot N_i \cdot M_{\mathrm{p}} \cdot \frac{dX^{\mathrm{c}}_{i,k,f}}{d\tau} = -\dot{M}^{\mathrm{c}}_{\mathrm{ps},i,k,f} \tag{9.25}$$

$$\Delta Q^{\mathrm{c}}_{3,k} \cdot \Delta Q^{\mathrm{c}}_{3,f} \cdot N_i \cdot M_{\mathrm{p}} \cdot \frac{dh^{\mathrm{c}}_{\mathrm{p},i,k,f}}{d\tau} = \left(\dot{Q}^{\mathrm{c}}_{\mathrm{sp},i,k,f} - \dot{H}^{\mathrm{c}}_{\mathrm{ps},i,k,f} + \dot{Q}_{\mathrm{pp}}\right) \tag{9.26}$$

Here $\Delta Q^{\mathrm{cC}}_{3,k}$ and $\Delta Q^{\mathrm{c}}_{3,f}$ are the mass fraction of particles component c in size class k and moisture class f, respectively, N_i the fraction of particles in the residence time class i, and $\dot{Q}_{\mathrm{pp}}$ the interparticle heat transfer, that is, the exchange of heat between particles in different residence time and/or size classes, which differ in temperature. The enthalpy of entering and leaving particles $\dot{H}_{\mathrm{p,in}}$ and $\dot{H}_{\mathrm{p,out}}$ is taken into account by the total energy balance of the fluidized bed dryer.

Kinetics

To solve the above-described system of balance equations, the drying kinetics, that is, the exchange flows (Fig. 9.6), are needed. According to the model assumption of an ideally mixed solid phase, plug flow in the gas phase and a distribution of the particle properties due to the residence time and particle size distributions the exchange flows between gas and solids have to be modeled four-dimensionally along the height-coordinate and the distributions of residence time, particle size, and inlet moisture content.

The mass transfer between particles and gas phase is given by:

$$\begin{aligned}\dot{M}^{\mathrm{c}}_{\mathrm{ps},i,k,f} &= \int_0^1 \int_{d_{\mathrm{p},k-1}}^{d_{\mathrm{p},k}} \int_{m_{f-1}}^{m_f} \int_{\tau_{i-1}}^{\tau_i} \frac{\partial^4 \dot{M}_{\mathrm{ps}}}{\partial\xi \partial d_{\mathrm{p}} \partial m_f \partial\tau} \cdot d\tau \cdot dm \cdot dd_{\mathrm{p}} \cdot d\xi \\ &= \int_0^1 N_i \cdot \dot{v}\left(\eta^{\mathrm{c}}_{i,k,f}\right) \cdot \rho_{\mathrm{g}} \cdot \beta^{\mathrm{c}}_{\mathrm{ps},k,f} \cdot A^{\mathrm{c}}_{\mathrm{ps},k,f} \cdot \left[Y_{\mathrm{eq}}\left(X^{\mathrm{c}}_{\mathrm{ps},i,k,f},\, T^{\mathrm{c}}_{\mathrm{ps},i,k,f}\right) - Y_{\mathrm{s}}\right] d\xi\end{aligned} \tag{9.27}$$

with β_{ps} being the mass transfer coefficient and $A^{\mathrm{c}}_{\mathrm{ps},k,f}$ the solids surface of the specified class. $\dot{\nu}\left(\eta^{\mathrm{c}}_{i,k,f}\right)$ is the modified normalized drying rate after Burgschweiger and Tsotsas (2002) as a function of normalized particle moisture content after van Meel (1958).

This normalized drying function has to be determined experimentally for every component.

The heat transfer between gas and particles is similarly defined by:

$$\begin{aligned}\dot{Q}^{\mathrm{c}}_{\mathrm{sp},i,k,f} &= \int_0^1 \int_{d_{\mathrm{p},k-1}}^{d_{\mathrm{p},k}} \int_{m_{f-1}}^{m_f} \int_{\tau_{i-1}}^{\tau_i} \frac{\partial^4 \dot{Q}_{\mathrm{sp}}}{\partial\xi \partial d_{\mathrm{p}} \partial m \partial\tau} \cdot d\tau \cdot dm \cdot dd_{\mathrm{p}} \cdot d\xi \\ &= \int_0^1 N_i \cdot \alpha^{\mathrm{c}}_{\mathrm{ps},k,f} \cdot A^{\mathrm{c}}_{\mathrm{ps},k,f} \cdot \left[T_{\mathrm{s}}(\xi) - T^{\mathrm{c}}_{\mathrm{p},i,k,f}\right] d\xi\end{aligned} \tag{9.28}$$

with α_{ps} being the mass transfer coefficient.

The enthalpy flow rate from the particle to the suspension $\dot{H}^{\mathrm{c}}_{\mathrm{ps},i,k,f}$ resulting from the transferred vapor mass flux is calculated according to:

$$\dot{H}^{\mathrm{c}}_{\mathrm{ps},i,k,f} = \dot{M}^{\mathrm{c}}_{\mathrm{ps},i,k,f} \cdot \left(c_{\mathrm{w,g}} \cdot T^{\mathrm{c}}_{\mathrm{p},i,k,f} + \Delta h_{\mathrm{v}}\right) \tag{9.29}$$

The mass and heat transfer between bubble phase and suspension gas are given by:

$$\frac{\partial \dot{M}_{\mathrm{sb}}}{\partial\xi} = \rho_{\mathrm{g}} \cdot \beta_{\mathrm{sb}} \cdot A_{\mathrm{sb}} \cdot \left[Y_{\mathrm{s}}(\xi) - Y_{\mathrm{b}}(\xi)\right] \tag{9.30}$$

$$\frac{\partial \dot{Q}_{\mathrm{bs}}}{\partial\xi} = \alpha_{\mathrm{sb}} \cdot A_{\mathrm{sb}} \cdot \left[T_{\mathrm{b}}(\xi) - T_{\mathrm{s}}(\xi)\right] \tag{9.31}$$

with the products of the mass transfer coefficient and the total bubble surface $\beta_{\mathrm{sb}} \bullet A_{\mathrm{sb}}$ and of the heat transfer coefficient and the total bubble surface $\alpha_{\mathrm{sb}} \bullet A_{\mathrm{sb}}$ calculated according to a correlation by Groenewold and Tsotsas (1997).

The enthalpy transfer between suspension and bubble phase is calculated analogously to Eq. (9.29) by:

$$\frac{\partial \dot{H}_{\mathrm{sb}}}{\partial\xi} = \frac{\partial \dot{M}_{\mathrm{sb}}}{\partial\xi}\left[c_{\mathrm{w,g}}(T_{\mathrm{sb}}) \cdot T_{\mathrm{sb}} + \Delta h_{\mathrm{v}}(T_{\mathrm{sb}})\right] \tag{9.32}$$

Finally, the heat exchange between particles belonging to different residence time, particle size, or inlet particle moisture content classes and having different temperatures is calculated according to Burgschweiger and Tsotsas (2002) by:

$$\dot{Q}_{\mathrm{pp},i,k,m} = \int_{d_{\mathrm{p},k-1}}^{d_{\mathrm{p},k}} \int_{m_{f-1}}^{m_f} \int_{\tau_{i-1}}^{\tau_i} \frac{\partial \dot{Q}_{\mathrm{pp}}}{\partial d_{\mathrm{p}} \partial m \partial\tau} \cdot d\tau \cdot dm \cdot dd_{\mathrm{p}} = \dot{N}_i \cdot \alpha_{\mathrm{pp}} \cdot A_{\mathrm{ps},k,m} \cdot \left(\overline{T}_{\mathrm{p}} - T_{\mathrm{p},i,k,m}\right) \tag{9.33}$$

$A_{ps,\,k,\,m}$ is the particle surface of all particles belonging to size class k and moisture class m and α_{pp} the heat transfer coefficient between the particles. As a first approach, the heat transfer coefficient α_{pp} was calculated with the correlation of Martin (1980) for the heat transfer between wall and suspension. Because the contact time and the contact area of two different particles is very small, therefore, the exchange of moisture between different classes is neglected. All equations required for calculation of heat and mass transfer coefficients are summarized in Appendix A.

It is worth mentioning that the heat and mass transfer coefficient and the gas properties such as density, viscosity, and thermal conductivity are calculated as a function of the actual temperature and pressure of the drying gas at each height. The new model does furthermore consider changes of the bed cross section with height (e.g., a widening of the bed) and the change of gas velocity resulting from the evaporated vapor mass flux from solids.

Fluid Dynamics of the Fluidized Bed

The fluid dynamics of fluidized beds are strongly influenced by the bubble formation. The bubbles flow through the bed at a rate which increases with the bubble diameter. In general, the bubble rise velocity in the fine-grained fluidized bed is substantially higher than the gas velocity in the suspension phase. This means that the bubbles act as a bypass. In contrast, the gas velocity in the suspension phase is greater than the bubble rise velocity in coarse grain fluidized beds, which means that the bypass effect is reduced. Bubble size and behavior depend on the type of bulk material in terms of fluidization behavior (Geldart classification (Geldart 1973)), on the height of the bed, on the fluidization gas flow, and on the type of gas distributor.

In order to calculate the bubble development in gas-solid fluidized beds, various empirical correlations are given in the literature. In this work, for the modeling of the fluidized bed, a model of bubbling bed according to Werther and Wein (1994) has been used. This model considers the combined action of coalescence and splitting of bubbles. It is assumed that the bed consists of two phases, a dense phase and solid-free bubble phase. All bed properties are assumed to be independent of the radius, that is, only cross-sectional average characteristics are considered.

The mechanism of bubble formation at an opening has been studied by Davidson and Harrison (1963); they found a simple relationship between the gas flow and the initial diameter of the resulting bubbles:

$$d_{v,o} = 1.3 \cdot \left(\frac{\dot{V}_{or}^2}{g} \right)^{0.2} \tag{9.34}$$

where $d_{v,o}$ is the diameter of the volume equivalent sphere.

With the superficial fluidizing gas velocity u, the volumetric flow $\dot{V}_{or}$ through a single opening of a nozzle of the air distribution plate is given as:

$$\dot{V}_{or} = \frac{A_t \cdot u}{n_{or}} \tag{9.35}$$

Gas-solid systems containing particles of Geldart group A and B (Geldart 1973) are characterized by "fast" bubbles and the appearance of splitting processes.

The bubble growth is described by Hilligardt and Werther (1987):

$$\frac{d\left(d_v\right)}{dh} = \left(\frac{2 \cdot \varepsilon_b}{9 \cdot \pi}\right)^{\frac{1}{3}} - \frac{d_v}{3 \cdot \lambda \cdot u_b} \tag{9.36}$$

where d_v is the mean local diameter of the bubbles, given as diameter of the volume equivalent sphere, ε_b the local bubble volume fraction, u_b the mean local bubble velocity, and λ the mean bubble lifetime.

The first term in Eq. (9.36) describes growth by coalescence and the second term accounts for bubble splitting. The mean lifetime λ of bubbles, representing the mean duration between two events of bubble splitting, can be expressed according to Hilligardt and Werther (1987) as follows:

$$\lambda = 208 \cdot \frac{u_{mf}}{g} \tag{9.37}$$

Here u_{mf} denotes the minimum fluidization and g is the acceleration of gravity.

The local bubble volume fraction ε_b is given by:

$$\varepsilon_b = \frac{\dot{V}_b}{u_b} \tag{9.38}$$

where $\dot{V}_b$is the visible bubble flow rate based on unit bed area and u_b is the average rise velocity of the bubbles at the height h above the distributor. $\dot{V}_b$ may be related to the excess gas velocity $(u - u_{mf})$ and can be calculated according to Hilligardt and Werther (1986) as:

$$\dot{V}_B \approx \psi \cdot \left(u - u_{mf}\right) \tag{9.39}$$

Here u denotes the superficial gas velocity.

The local bubble rise velocity may be calculated according to Hilligardt and Werther (1986):

$$u_B = \dot{V}_B + 0.71 \cdot \theta \cdot \sqrt{g \cdot d_v} \tag{9.40}$$

Table 9.1 The hydrodynamic parameter ψ for the Geldart groups A, B, and D (Geldart 1973) according to Hilligardt and Werther (1986)

A	B	D
$\psi = 0.8$ for $\frac{h}{d_t} < 1$	$\psi = 0.67$ for $\frac{h}{d_t} < 1.7$ $\psi = 0.51 \cdot \left(\frac{h}{d_t}\right)^{1/2}$ for $1.7 \le \frac{h}{d_t} \le 4$ $\psi = 1$ for $\frac{h}{d_t} > 4$	$\psi = 0.26$ for $\frac{h}{d_t} < 0.55$ $\psi = 0.35 \cdot \left(\frac{h}{d_t}\right)^{1/2}$ for $0.55 \le \frac{h}{d_t} \le 8$ $\psi = 1$ for $\frac{h}{d_t} > 8$

Table 9.2 The hydrodynamic parameter ϑ for the Geldart groups A, B, and D (Geldart 1973) according to Hilligardt and Werther (1986)

A	B	D
$\theta = 1.18$ for $d_t < 0.05$ m $\theta = 3.2 \cdot d_t^{1/3}$ for $0.05 \le d_t \le 1$ m $\theta = 3.2$ for $d_t > 1$ m	$\theta = 0.63$ for $d_t < 0.1$ m $\theta = 2 \cdot d_t^{1/2}$ for $0.1 \le d_t \le 1$ m $\theta = 2$ for $d_t > 1$ m	$\theta = 0.87$

According to Hilligardt and Werther (1986), the parameter ψ describes the deviation of the visible bubble flow rate (Eq. 9.39) from the two-phase theory according to Toomey and Johnstone (1952), and θ accounts for the deviation from the behavior of a single bubble and depends only on the bed diameter d_t.

Numerical values for the hydrodynamic parameters υ and θ for the Geldart groups A, B, and D (Geldart 1973) are illustrated in Tables 9.1 and 9.2 (Hilligardt and Werther 1986).

Coarse-particle or Geldart group D (Geldart 1973) fluidized bed is characterized by the appearance of slow bubbles. The bubble diameter varies with the bed height according to Hilligardt and Werther (1987) as:

$$\frac{d(d_v)}{dh} = \left(\frac{2 \cdot \varepsilon_b}{9 \cdot \pi}\right)^{\frac{1}{3}} \cdot \left(\frac{1}{1 - \xi \cdot \left(\frac{6 \cdot \varepsilon_b}{\pi}\right)^{1/3}}\right) \tag{9.41}$$

with

$$\xi = \begin{cases} 1 - \alpha^3 & \text{for } \alpha \le 1 \\ 0 & \text{for } \alpha > 1 \end{cases} \tag{9.42}$$

where α is the crucial parameter, which defines the border between the fast and slow bubbles; it is defined for a freely bubbling bed as follows:

$$\alpha = \frac{u_b}{u_s / \varepsilon_s} \tag{9.43}$$

u_s / ε_s denotes the average gas velocity in the interstices of the suspension phase.

The gas velocity in the suspension phase can be calculated by:

$$\frac{u_s - u_{mf}}{u - u_{mf}} = \begin{cases} 1/3 & \text{for porous plates} \\ 1/4 & \text{for industrial gas distributors} \end{cases} \tag{9.44}$$

The bubble rise velocity may be calculated as:

$$u_b = \max\left\{u_{b,i}, u - 2.71 \cdot u_{mf} + 1.36 \cdot u_{b,i}\right\} \tag{9.45}$$

$u_{b,i}$ is the velocity of an isolated single rising bubble, which can be calculated according to Davies and Taylor (1950) as:

$$u_{b,i} = 0.71 \cdot \sqrt{g \cdot d_v} \tag{9.46}$$

Richardson and Zaki (1954) suggested a simplified equation to calculate the porosity of the suspension phase ε_s as follows:

$$\varepsilon_s = \varepsilon_{mf} \cdot \left(\frac{u_s}{u_{mf}}\right)^{1/4.65} \tag{9.47}$$

The solids volume concentration c_v at a height h above the distributor can be determined according to Werther and Wein (1994) by:

$$c_v = \left(1 - \varepsilon_b\right) \cdot c_{vd} \tag{9.48}$$

The measurements carried out by Werther and Wein (1994) have shown that c_{vd} is independent of height h and depends only on gas velocity and bed material properties:

$$\frac{c_{vd}}{c_{vmf}} = 1 - 0.14 \cdot \mathrm{Re}_p^{0,4} \cdot \mathrm{Ar}^{-0,13} \tag{9.49}$$

In this equation c_{vmf} is the solids volume concentration at minimum fluidization and can be calculated from the bed porosity at the minimum fluidization point ε_{mf}:

$$c_{vmf} = 1 - \varepsilon_{mf} \tag{9.50}$$

In Eq. (9.49) Re_p is the Reynolds number based on the excess gas velocity $(u - u_{mf})$ and Ar is the Archimedes number:

$$\mathrm{Re}_p = \frac{\left(u - u_{mf}\right) \cdot d_p}{\nu} \tag{9.51}$$

$$\mathrm{Ar} = \frac{g \cdot d_\mathrm{p}^3}{\nu^2} \cdot \frac{\rho_\mathrm{p,w} - \rho_\mathrm{g}}{\rho_\mathrm{g}} \tag{9.52}$$

The required density of the solid for wet material with the moisture load X results from:

$$\rho_\mathrm{p,w} = \left(1 + X\right) \cdot \rho_\mathrm{p} \tag{9.53}$$

The porosity ε at each height above the distributor can be calculated under the assumption that in the suspension phase throughout the fluidized bed a uniform porosity ε_d prevails and the bubble phase is considered to be solid-free:

$$\varepsilon = \int_0^{L_\mathrm{T}} \left[1 - \left(1 - \varepsilon_\mathrm{b}\left(h\right)\right) \cdot \left(1 - \varepsilon_\mathrm{d}\right)\right] \cdot dh \tag{9.54}$$

The total mass of solids in the fluidized bed M_p can then be determined from, for example, the total pressure drop over the bed (Werther and Wein 1994):

$$M_\mathrm{p} = \int_0^{L_\mathrm{T}} \rho_\mathrm{P} \cdot c_\mathrm{v}\left(h\right) \cdot A_\mathrm{t}\left(h\right) \cdot dh \tag{9.55}$$

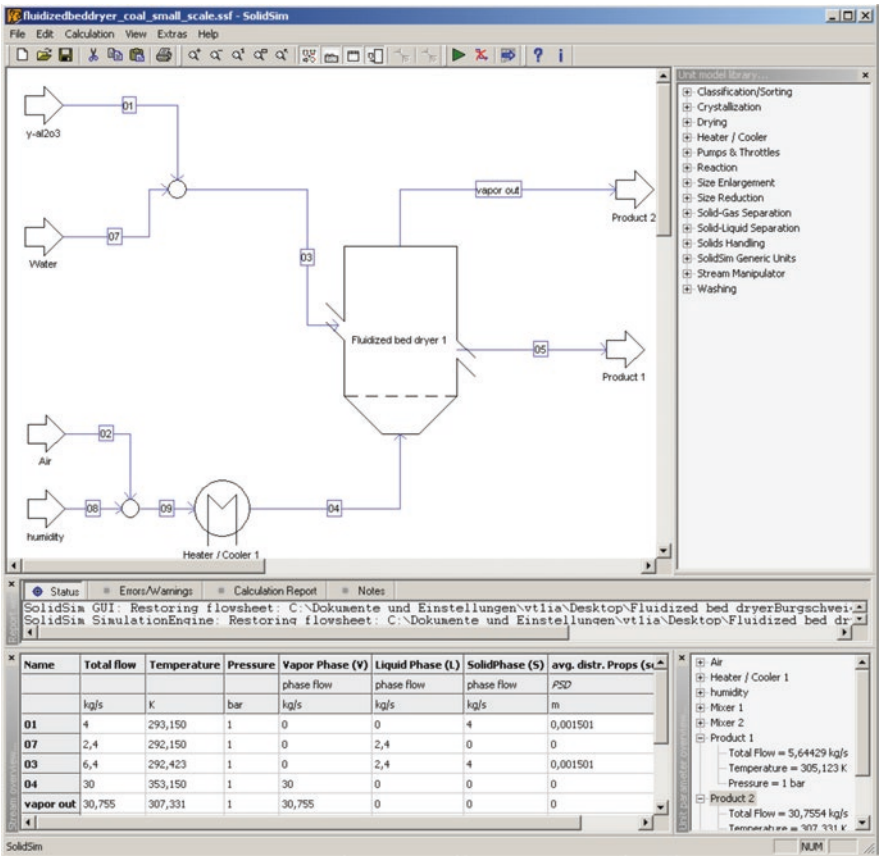
Model Application

The system of differential equations as described above has been implemented as a unit module into the flowsheet simulation program SolidSim (Hartge et al. 2006a, b), the latter providing the user interface for the definition of the feeds, that is, the gas and the solids feed, which especially provides the module with all solids and gas properties as function of temperature and pressure. This allows to easily apply the system to a wide range of solids, liquids, and gases. Furthermore, it allows to connect the fluidized bed dryer with other unit modules to investigate the interaction of different unit operations. A screenshot of SolidSim with the fluidized bed dryer module is shown in Fig. 9.7.

The following strategy has been used to solve the system of equations numerically:

- The residence time coordinate τ has been discretized geometrically into n classes and integrated by the Euler method.
- For the calculation of the gas moisture and temperature profiles, the dryer was divided into q finite height elements. A fourth-order Runge-Kutta algorithm was used to solve the height integration.
- l particle size classes have been used to describe the particle size distribution.
- The inlet moisture content coordinate has been discretized into r classes.

The user can change n, q, l, and r depending on the desired accuracy and calculation time.



Name	Total flow	Temperature	Pressure	Vapor Phase (V)	Liquid Phase (L)	SolidPhase (S)	avg. distr. Props (s
				phase flow	phase flow	phase flow	PSD
	kg/s	K	bar	kg/s	kg/s	kg/s	m
01	4	293,150	1	0	0	4	0,001501
07	2,4	292,150	1	0	2,4	0	0
03	6,4	292,423	1	0	2,4	4	0,001501
04	30	353,150	1	30	0	0	0
vapor out	30,755	307,331	1	30,755	0	0	0

Fig. 9.7 Screenshot of the SolidSim program with the fluidized bed dryer module

9.4 Experimental

9.4.1 Bed Material

Aluminum oxides in both alpha and gamma modifications with different particle sizes were used as bed material.

γ-Alumina (γ-Al_2O_3) is a white, water-insoluble, tasteless, odorless solid that is soluble in strong acids and bases. Due to their large internal surface area and crystal structure, γ-Al_2O_3 particles are used widely in industry as adsorbents, desiccants, catalyst carriers, as well as catalysts themselves. In this work, γ-Al_2O_3 from the producer "Sasol" was used as a model substance.

α-Al_2O_3 is a very hard material with a Mohs hardness of 9–9.5. α-Al_2O_3 has a trigonal crystal system which arises from calcining γ-Al_2O_3 at temperatures at above 1000 °C.

Table 9.3 Physical properties measured for both materials

Parameter	Unit	γ-Al_2O_3	α-Al_2O_3
Total pore volume	[ml/g]	0.59	0.373
Specific surface	[m^2/g]	213	80.3
Mean pore diameter	[µm]	0.0078	0.0142
Particle density	[g/ml]	1.08	1.507
Solid density	[g/ml]	3.06	3.44
Particle porosity	[−]	0.64	0.56

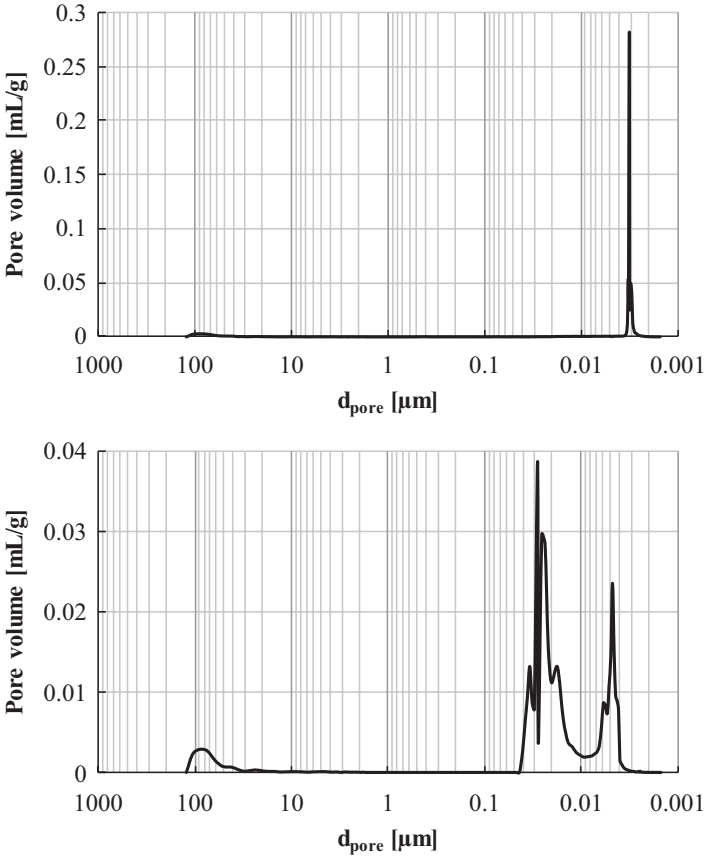


Fig. 9.8 Pore size distribution for γ-Al_2O_3 (top) and α-Al_2O_3 (bottom) measured by mercury porosimetry

α-Alumina is denser than γ-alumina and has significantly larger crystals and smaller porosity, however, larger pore diameter. Some physical properties are given in Table 9.3. Figure 9.8 presents the pore size distribution measured by mercury porosimetry.

9.4.2 *Fluidized Bed Installation*

The experimental work using continuous fluidized bed drying operation was carried out in the system as schematically shown in Fig. 9.9.

The system basically consists of a fluidized bed apparatus, with the conical fluidization chamber manufactured by the company "Glatt" having a diameter of 180 mm at the bottom and 250 mm at the top. The height of the fluidized bed apparatus is 320 mm. The fluidized bed apparatus is made of stainless steel with two sight glasses for process monitoring. A sieve tray was used as air distributor.

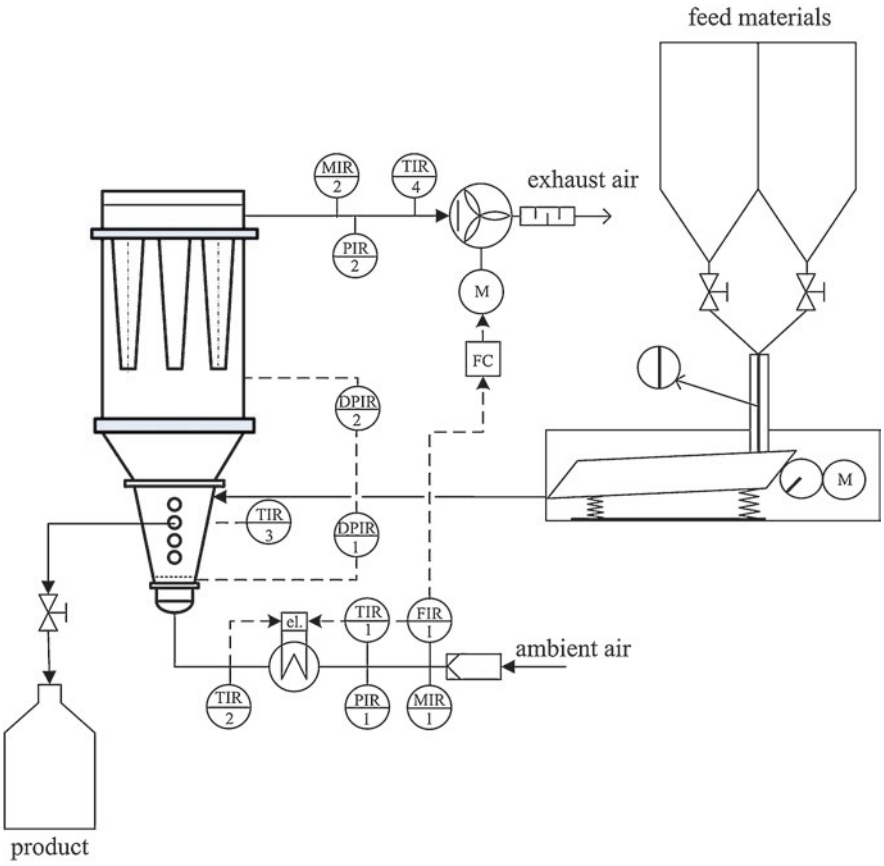


Fig. 9.9 Flowsheet of continuous fluidized bed installation

Ambient air is conveyed with a forced draft fan through a 4 kW electric heater. The desired air volume flow can be adjusted by controlling the speed of this fan using a variable frequency drive. After the air has passed the fluidized bed layer, it is decelerated in the subsequent settling zone featuring a conical enlargement of its cross section. The velocity of air is therefore reduced so that the largest portion of the fine particles will fall back into the fluidized bed. The discharged particles are separated from the exhaust air in a baghouse filter installed downstream of the settling zone.

The moist solid is dosed into the dryer by means of a vibrating conveyor. Two separate bins are used to prevent segregation in the case of occurrence of different particle sizes or possible moisture diffusion at different feed moisture contents as well as enabling feeding of various types of solids. The mass fraction of the conveyed materials is set by varying the respective gap height above the trough of the vibrating conveyor.

The discharge of dried particles happens continuously over a lateral opening (overflow weir) with a diameter of 16 mm, whereas the bed height can be adjusted by the changing the overflow height of this discharge opening. For this purpose, four holes at different heights were available. The dried particles are collected in a closed container.

In addition, samples from the bed material can be taken during the ongoing experiment through a lateral nozzle.

To monitor the fluidized bed drying process, the pilot plant was equipped with a measurement data recording system for displaying and online registration of temperature, humidity, mass flow, and pressure at several measuring locations.

Measuring System

The temperature measurements at the points TIR1, TIR2, TIR3, and TIR4 were done using PT100-RTD temperature probes. To calculate the absolute gas humidity from the measured relative gas humidities at inlet MIR1 and at outlet MIR2 as well as the mass flow rate from the volumetric flow rate measurement, the absolute pressures PIR1 and PIR2 had to be measured in addition to the temperatures indicated at TIR2 and TIR4.

For the intake air, a moisture sensor DKRF 400-10-200 from company Driesen + Kern GmbH (measuring range 0–100% RH, accuracy +/− 1.8% RH in the range of 20–80% RH) was used.

The relative humidity of the exhaust air is measured by a sensor type SHT75 from the company SENSIRION (measuring range 0–100% RH, accuracy +/− 1.8% RH). Both sensors were calibrated with a special salt solution kit. The relative humidities above the salt solutions were 11.3%, 33.1%, and 75.5%. The calibration results are depicted in Fig. 9.10.

The volumetric flow rate was measured with a vane anemometer at the measuring point (FIR).

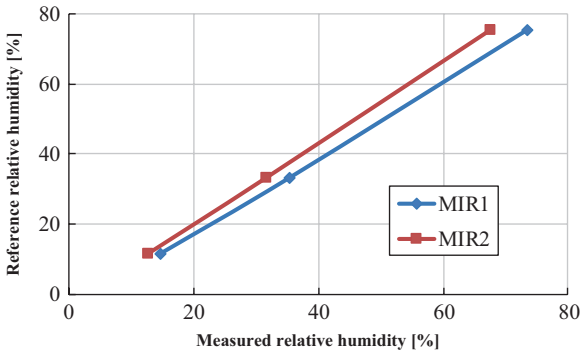


Fig. 9.10 Calibration of the humidity sensors

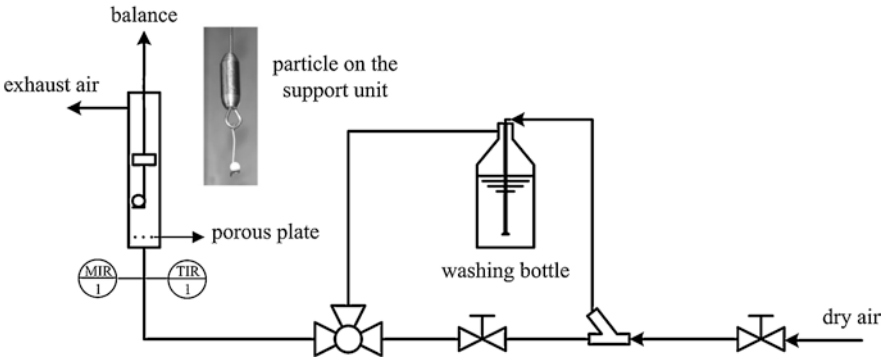


Fig. 9.11 Experimental setup for single particle moisture measurement

9.4.3 Single Particle Installation

To measure the equilibrium between the solid moisture and the moisture loading of the ambient air, the installation shown in Fig. 9.11 was used. The particle to be examined was placed on a support unit connected to the magnetic suspension balance and exposed to a stream of air having a defined temperature and humidity. The drying air was divided into two streams, one part flowing through a wash bottle and the other stream being bypassed. By changing the mixing ratio of both streams, the desired relative humidity could be adjusted. In addition, the temperature of the measuring chamber could be adjusted separately. To ensure uniform distribution of the drying air around the particle, the air had to pass through a porous plate.

Instead of a conventional gravimetric measuring instrument, a magnetic suspension balance with a magnetic coupling consisting of a permanent and an electromagnet is used to transfer the weight of the sample contact-free from the measuring chamber to the microbalance, which is located at the exterior of the chamber itself at ambient atmospheric conditions. Due to the fact that the measuring chamber and the microbalance are spatially separated, it is possible to determine the weight change of a single particle due to adsorption or desorption under industry-specific real-time conditions and at high resolutions of up to 1 µg accuracy.

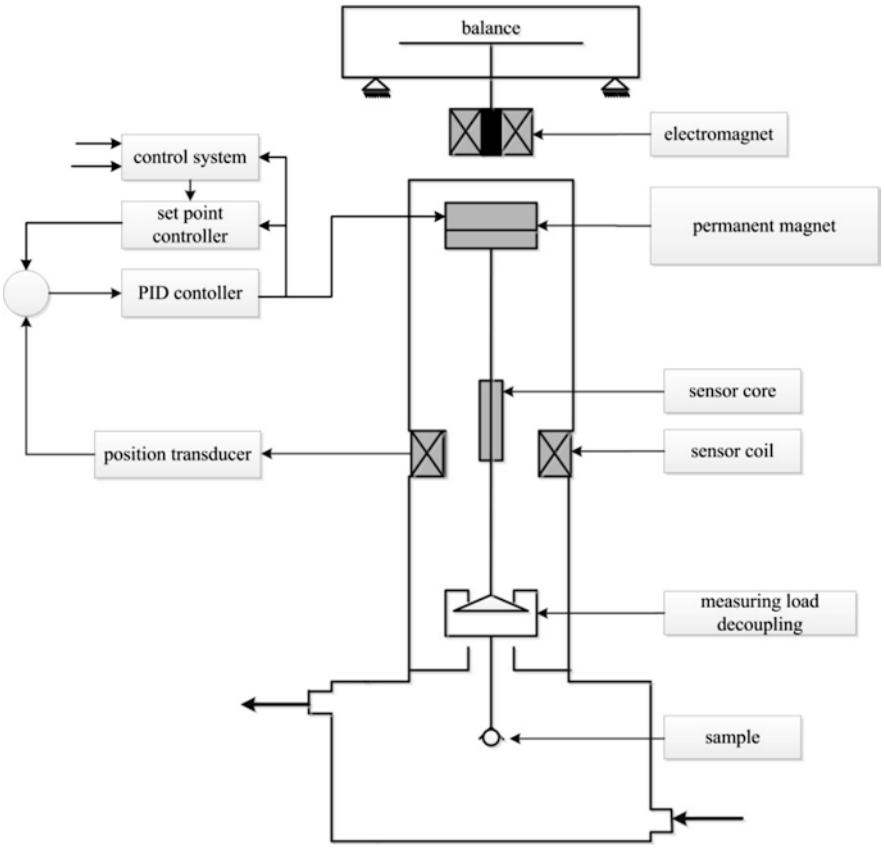


Fig. 9.12 Schematic representation of magnetic suspension balance from the company Rubotherm GmbH

In this work, a magnetic suspension balance from the company Rubotherm GmbH (Bochum, Germany) was used. Figure 9.12 shows schematic the magnetic suspension balance.

The same measuring devices for humidity, temperature, and flow rate as described above were used.

9.4.4 Single Particle Measurements

Determination of Sorption Isotherms

To measure the sorption isotherms, the experimental setup shown in Fig. 9.11 was used.

Starting with an initially dry particle and a stepwise humidification, the adsorption function is obtained by increasing the relative humidity surrounding the particle. Analogously, starting with an initially wet particle, the desorption isotherm is obtained through gradual drying by decreasing relative humidity.

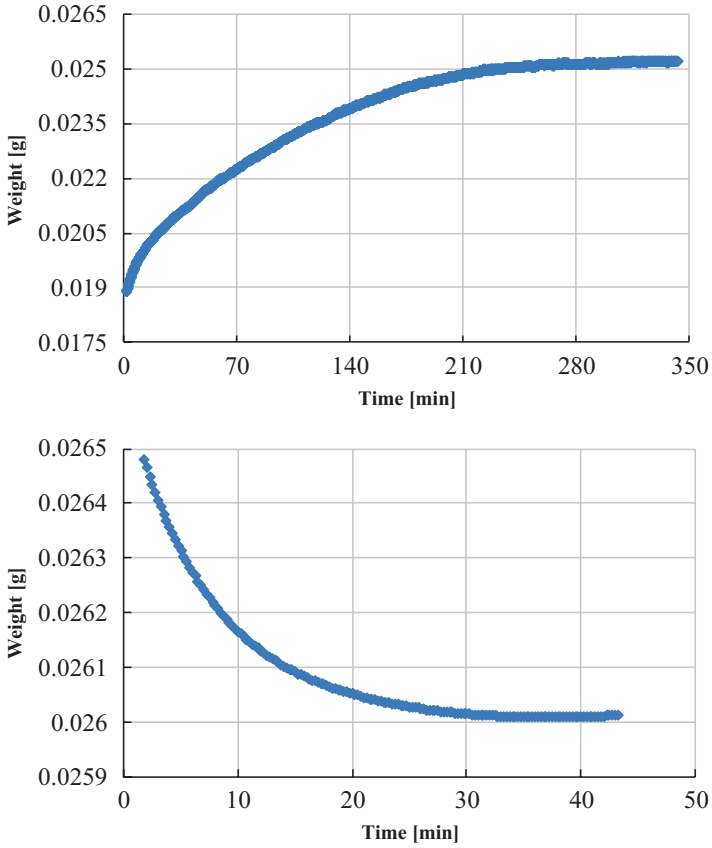


Fig. 9.13 Adsorption curve (top) and desorption curve (bottom) of 3 mm single particle of γ-Al_2O_3 (T_g = 41 °C, φ = 85%, u_g = 0.03 m/s)

For desorption, the magnetic suspension balance registers the weight resp. mass loss until a constant weight (mass) is reached (Peglow et al. 2009). The corresponding moisture content of the solid particle will represent the equilibrium moisture content. In Fig. 9.13, the weight decrease and increase as measured by the magnetic suspension balance due to desorption and adsorption of moisture from the sample are represented.

Figure 9.14 illustrates the measured desorption and adsorption isotherms of γ-Al_2O_3 particles for particle sizes of 1.8 and 3 mm, where the particle moisture content is a function of the relative humidity of the surrounding air at a constant temperature of 40 °C. In the range $\varphi < 0.6$, the two curves are hard to distinguish, but a typical hysteresis between adsorption and desorption in the second range $\varphi > 0.6$ is clearly recognizable.

The moisture content of γ-Al_2O_3 increases slowly up to a relative humidity of 60%. Further increase in relative humidity increases the moisture content sharply to

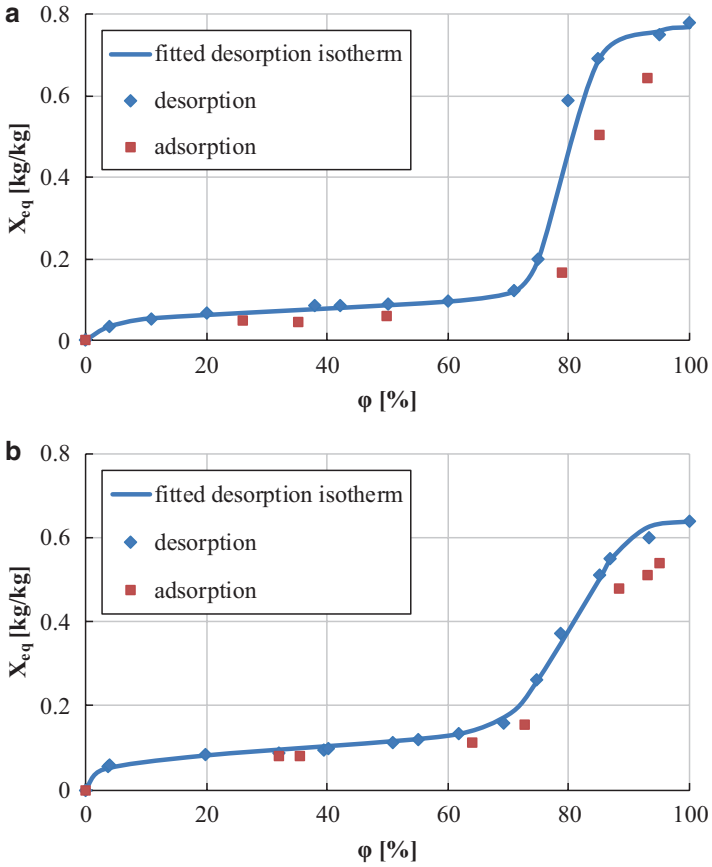


Fig. 9.14 Adsorption and desorption isotherms of γ-Al_2O_3 particles having a particle size of 1.8 (**a**) and 3 mm (**b**) measured by magnetic suspension balance

a maximum value of 0.78 kg/kg for the 1.8 mm particles and to 0.64 kg/kg for 3 mm particles. This maximum value depends directly on the particle porosity and increases with increasing porosity. The rapid increase in the moisture content at high relative air humidity can be attributed to the capillary condensation in the pores.

The problem with this method is the accurate adjustment and determination of the relative humidity. In addition, it was observed that at higher relative humidities, water vapor condensed on the particle surface as well as the supporting unit, which in turn affects the results. Therefore, the accurate determination of the equilibrium in this area is very difficult or even impossible using a magnetic suspension balance.

For the determination of sorption isotherms at higher relative humidities as well as for fine particles (restriction of a magnetic suspension balance), the desiccator method was used.

The samples (about 15 g) were stored in desiccators on top of different saturated aqueous salt solutions. The salt solutions were used to create discrete relative

humidities inside of the desiccators. Weight constancy of the samples would indicate the state of equilibrium with the desiccant. From the corresponding relative humidity and the water content, the solid at the equilibrium resulted in the sorption isotherm.

Figure 9.15 shows the adsorption and desorption isotherms of γ-Al_2O_3 charge having a particle size of 0.6 mm and α-Al_2O_3 charge measured with the desiccator method.

In the region $\varphi < 0.6$, all experiments show almost identical sorption isotherms; however, the curves shift slightly upward with increasing particle diameter (Fig. 9.14a, b). This can be attributed to the fact that with higher particle diameter, the crystal defects are greater, which in turn increases sorptive binding.

In the region $\varphi > 0.6$, the isotherms differ greatly, which can be attributed to the pore structure (pore diameter). As generally known, the smaller the pore diameter,

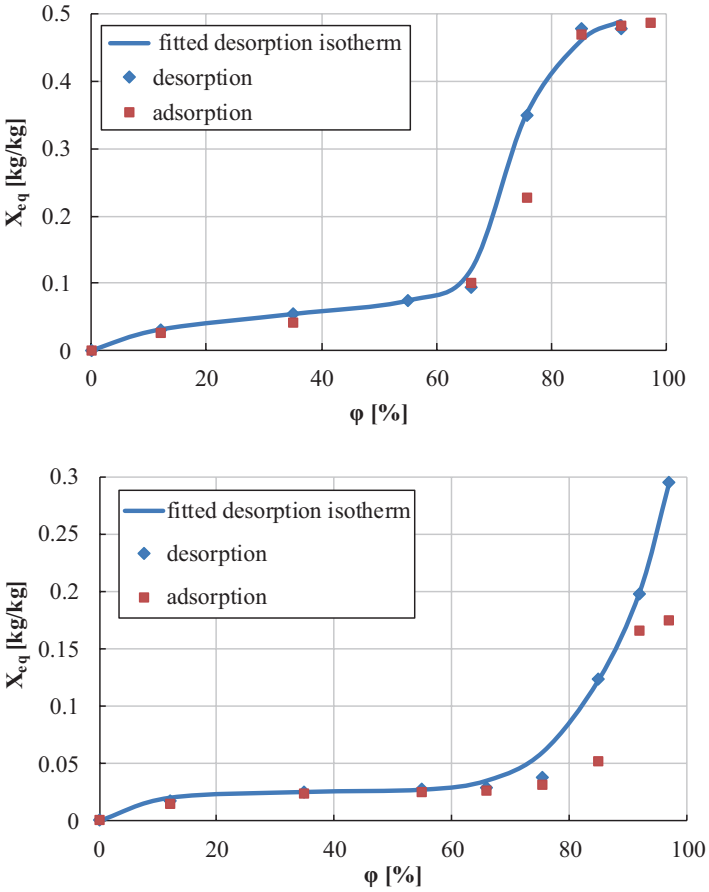


Fig. 9.15 Adsorption and desorption isotherms of γ-Al_2O_3 particles having a particle size of 0.6 mm (above) and α-Al_2O_3 (below) measured by the desiccator method

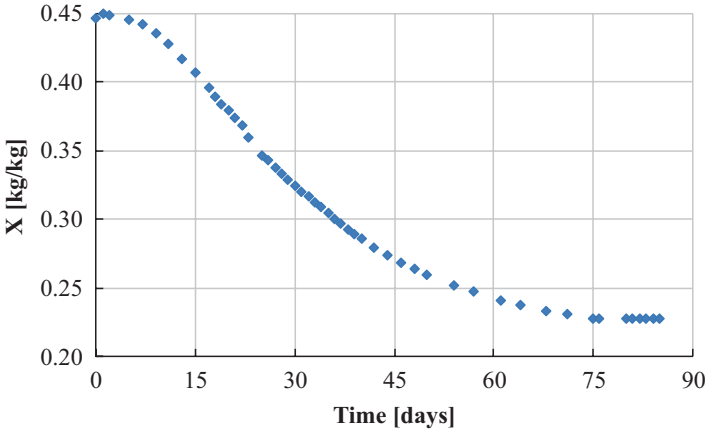


Fig. 9.16 Measured drying curve by means of desiccator method for γ-Al_2O_3 particles having a particle size of 0.6 mm by a relative humidity of 75%

the lower the vapor pressure inside the pore will be; hence, it would also be possible to condense at relative humidities of less than 100%. Additionally, in this area the porosity plays a dominant role, so particles with higher porosity (higher capacity to store moisture inside the particle) would absorb more moisture such as the 1.8 mm particles with a higher porosity of 0.7 would take a maximum humidity of 0.78 kg/kg, whereas the 3 mm particles having a porosity of 0.64 would only achieve a maximum moisture content of about 0.65 kg/kg (Fig. 9.14b).

The desiccator method yields very good and accurate results due to the accurate determination of the humidity and relatively large samples (about 15 g). However, the desiccator method is very time-consuming; for instance, the measuring of desorption at a relative humidity of 75% took 85 days (Fig. 9.16).

Drying Kinetics

Depending on the particle size, the kinetics of mass transfer inside the particles can be determined either with direct or indirect methods. The direct methods are based on a large particle or limited number of particles together. It produces good and reproducible results (Kwapinski and Tsotsas 2006).

Measurements should be carried out in a gradient-free environment so that the heat balance of the ambient air can be neglected (Groenewold 2005). The measurements with a microbalance (Burgschweiger et al. 1999), a drying channel (Hirschmann and Tsotsas 1998; Kemp et al. 2001; Hirschmann et al. 1998; Groenewold et al. 2002), an acoustic levitator (Groenewold et al. 2002), and the magnetic suspension balance (Suherman 2007; Cunäus 2010) are direct methods.

In indirect methods, the normalized drying curve of a single particle is derived by scale-down (Hirschmann et al. 1998; Fyhr and Kemp 1998a, b). This method is

especially suitable for particles that are smaller than 1 mm (Groenewold 2005; Tsotsas and Mujumdar 2009). In this case, the time incremental change of the absolute humidity of the drying air is recorded and converted using a fluid dynamic model to obtain the weight loss of the solid.

In this work the magnetic suspension balance is used to determine the normalized drying curve for the 1.8 and 3 mm-γ-particles. Figure 9.17 shows the measured drying curve for both particle sizes.

The 3 mm particle took about 1500 s, while the 1.8 mm particle only one-third the time to reach nearly the same equilibrium moisture content.

The fine particles showed a higher drying rate in both the first drying period and the second. This tendency may be attributed mainly to the increased heat and mass transfer area per unit volume. In addition, the mass transfer coefficient is inversely proportional to the particle diameter at a nearly equal Sherwood number, and the increased mass transfer coefficient with the reduction in the particle diameter thus supports this tendency. Even though the small particles have a higher drying rate

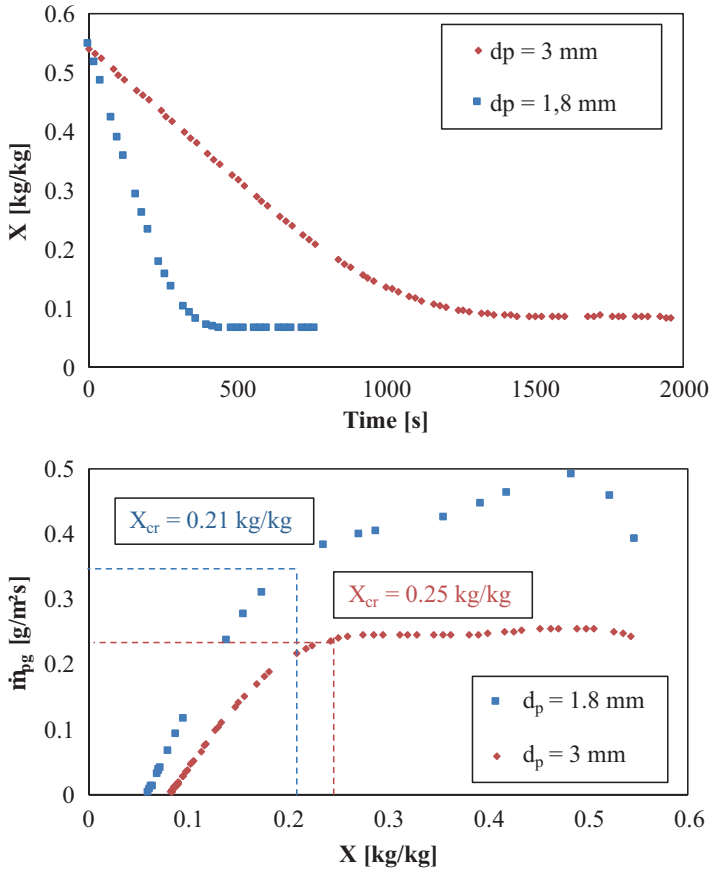
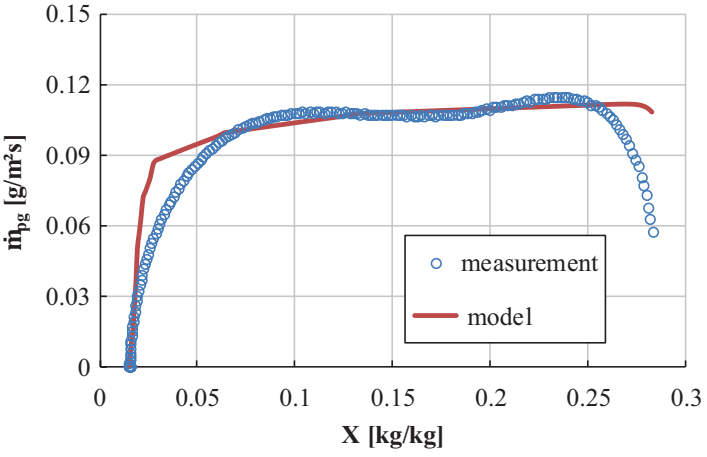


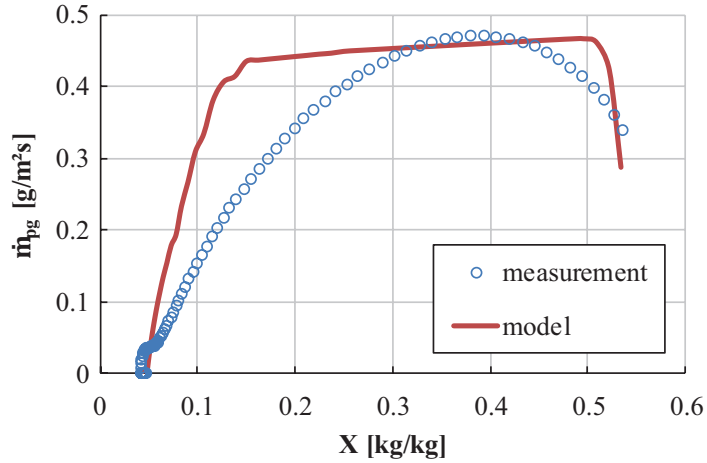
Fig. 9.17 Drying curve (top) and drying rate (bottom) of γ-Al_2O_3 particles having a particle size of 1.8 and 3 mm measured by magnetic suspension balance (T_g = 41 °C, φ = 20%, u_g = 2.8 cm/s)

(sharp drying), the second drying phase starts earlier by coarse particles because the capillary moisture diffusion due to higher transfer resistances is insufficient to offset the evaporated moisture at the particle surface. This results in a critical moisture content of 0.21 kg/kg, where it is 0.25 kg/kg for the coarse particles (Fig. 9.17).

The curves in Figs. 9.4 and 9.18 show the measured and calculated drying curves for γ- and α-single particle, respectively. The calculation was carried out by neglecting the particle internal resistances, that is, the normalized drying curve was set to be equal to 1.



(d_p = 2.6 mm, T_g = 22.4°C, φ = 27% and u_g = 0 cm/s)



(d_p = 3 mm, T_g = 41°C, φ = 3% and u_g = 10 cm/s)

Fig. 9.18 Measured and calculated drying curve for a single particle of α-Al_2O_3 (**a**) and γ-Al_2O_3 (**b**). (**a**) d_p = 2.6 mm, T_g = 22.4 °C, φ = 27%, u_g = 0 cm/s. (**b**) d_p = 3 mm, T_g = 41 °C, φ = 3%, u_g = 10 cm/s

Since the calculated curve (solid line) only represents the hygroscopic reduction of the driving force, the measured curve includes the kinetic hindrances during the second drying phase. The area between the two curves reflects the magnitude of the kinetic hindrances during the second drying period. This depends directly on the material properties as well as on the drying conditions. γ-Al_2O_3 is highly hygroscopic compared to α-Al_2O_3. Figure 9.18a, b clearly illustrates that the area between the calculated and measured curves is larger for γ-Al_2O_3 than that for α-Al_2O_3.

Furthermore it can be concluded that the critical moisture content or the beginning of the second drying phase is strongly influenced by the material properties and the drying conditions as well; while the critical moisture content of α-Al_2O_3 is about 0.1 kg/kg (Fig. 9.18), it changes between 0.25 and 0.32 kg/kg (Figs. 9.4 and 9.18b) depending on the drying conditions.

Please note that the vapor mass flow in the first drying phase of the α-particles (Fig. 9.18a) is about half of the experimental value depicted in Fig. 9.4 and about a quarter of the experimental value depicted in Fig. 9.18b. The reason for this is that α-Al_2O_3 has a small specific surface area per unit volume due to its higher density and lower porosity. In addition the measurement was carried out under ambient conditions and resting air (microbalance restriction), which results in a Sherwood number of about 2.3 (theoretical value is 2), while on the other hand, this value was 4 for both runs as depicted in Figs. 9.4 and 9.18b, respectively.

Determination of Normalized Drying Curves

The normalized drying curve of the single particles was determined according to the normalizing method suggested by Burgschweiger et al. (1999) and Burgschweiger and Tsotsas (2002):

Determination of Normalized Drying for α-Particles

The measurement of the drying curve of a particle took place on the weighting pan of a microbalance. The weight loss of the particle was recorded by a computer and smoothed. Figure 9.19 shows the measured and smoothed data of a single α-Al_2O_3 particle drying:

A total of four tests were conducted to determine the normalized drying curve of α-Al_2O_3 particles. The parameters of the experiments and the corresponding determined critical moisture content are listed in Table 9.4.

The normalized drying curve of α-Al_2O_3 is shown in Fig. 9.20. The corresponding critical moisture content is in all experiments about 0.095 kg/kg.

Determination of Normalized Drying for 1.8 and 3 mm γ-Particles

The drying curves were measured by the magnetic suspension balance described in Sect. 9.4.3. Unlike the microbalance the drying curve can be determined in the magnetic suspension balance under real operating conditions.

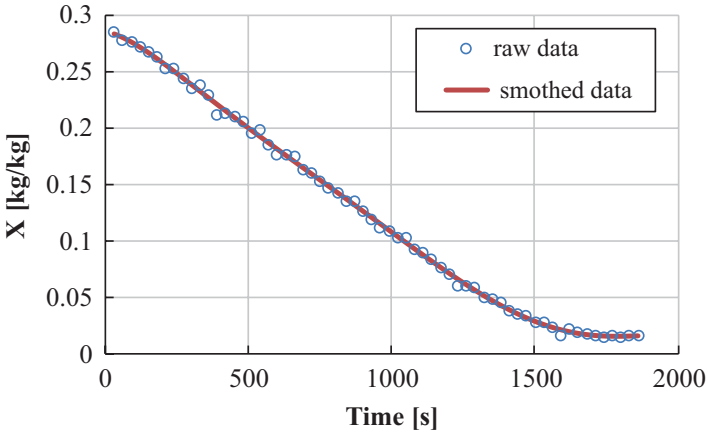


Fig. 9.19 Raw and smoothed data for an α-Al_2O_3 particle (d_p = 2.6 mm, T_g = 22.4 °C, φ = 27%, u_g = 0 cm/s)

Table 9.4 The parameters of all experiments carried out with single α-Al_2O_3 particles

Run no.	d_p (mm)	T_g (°C)	φ (%)	u_g (cm/s)	X_{cr} (kg/kg)
MB1	2.7	22.5	33	0	0.095
MB2	2.6	22.4	27	0	0.1
MB3	2.6	22.5	32	0	0.09
MB4	2.5	21.4	32	0	0.092

In Fig. 9.21, the normalized drying curve can be seen for two different 1.8 mm γ-particles under nearly identical conditions. It can be quite clearly seen that the curves are reproducible with good accuracy.

Figure 9.22 shows the measured normalized drying curve of a single particle under different conditions. The operating parameters of the experiments and the corresponding determined critical moisture content are listed in Table 9.5.

Figure 9.22 shows the measured normalized drying curves for 13 selected single particle experiments for γ-particles having a diameter of 3 mm at different drying conditions, which are summarized with the corresponding critical moisture content in Table 9.6.

It can be seen that for both 1.8 mm particles and for the 3 mm particles, the normalized drying curve shifts downward, when the drying takes place under severe conditions, that is, high temperatures and/or high gas velocities and/or low relative humidity The influence of the operating conditions is more dominant in case of fine particles than coarse as can be seen from the comparison between Figs. 9.22 and 9.23.

An important parameter for normalization is the critical moisture content, which represents the boundary between the first and second drying phase. As already mentioned the critical moisture content is a function of both material characteristic and drying conditions.

A comparison between the experiments MSW49 and MSW69 shows that the critical moisture content is higher, for sharper drying conditions. While the critical

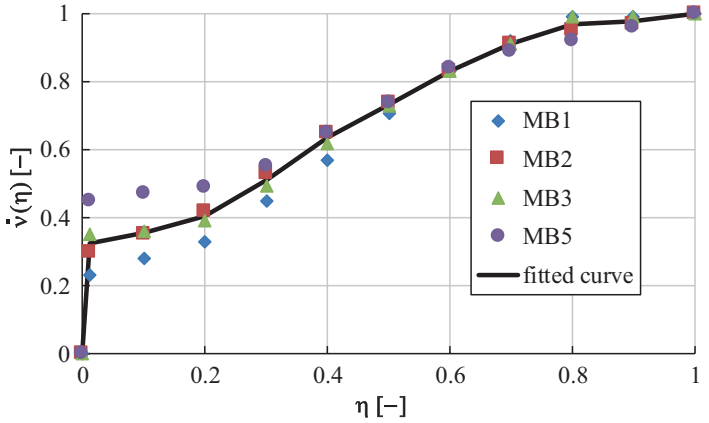


Fig. 9.20 Normalized drying curve of α-Al_2O_3 particles

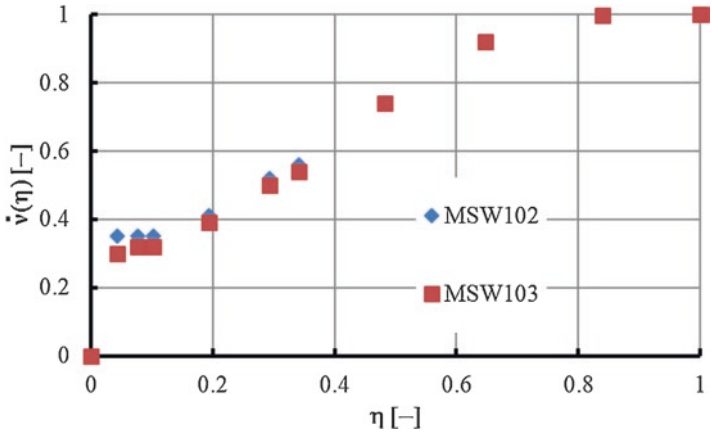


Fig. 9.21 Reproducibility of normalized drying curve for 1.8 mm γ-Al_2O_3 particles. (T_g = 25 °C, φ = 37%, u_g = 0 cm/s) and (T_g = 23 °C, φ = 45%, u_g = 0 cm/s)

moisture content in the first case (φ = 4%) is 0.33 kg/kg, this value is reduced to X_{cr} = 0.2 kg/kg at a relative humidity of φ = 60%.

The relationship between the critical moisture content and the particle diameter can be observed in the experiments MSW111 and MSW113. The two experiments were carried out under the same conditions. The critical moisture content for the 1.8 mm particles (MSW113) is X_{cr} = 0.24 kg/kg compared to X_{cr} = 0.27 kg/kg kg/kg for the 3 mm particles (MSW111).

The increase of the critical moisture content extends the second drying section and, consequently, increases the internal transport resistances. This means that with very coarse particles, the mass transport inside the particle is dominant; thus the external mass transfer is negligible compared to the internal mass transport.

Figure 9.24 shows the dependence of the critical moisture content on the drying conditions for 1.8 and 3 mm particles. On the *x*-axis the drying rate $\dot{m}_{pg}$ is applied

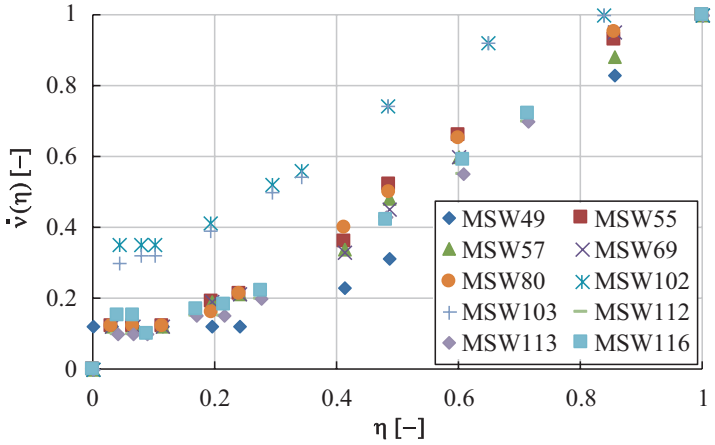


Fig. 9.22 Normalized drying curves for 1.8 mm γ-Al_2O_3 particles (operating parameters of the experiments to be found in Table 9.5)

Table 9.5 Operating parameters for the experiments carried out with 1.8 mm γ-Al_2O_3 particles

Run no.	d_p (mm)	T_g (°C)	φ (%)	u_g (cm/s)	X_{cr} (kg/kg)
MSW49	1.8	41	3.75	0.03	0.34
MSW55	1.8	41	11	0.03	0.33
MSW57	1.8	41	27.3	0.03	0.22
MSW69	1.8	40.5	50	0.03	0.2
MSW80	1.8	40	38	0.03	0.21
MSW102	1.8	25	37	0	0.2
MSW103	1.8	23	45	0	0.2
MSW107	1.8	41	4	0.03	0.3
MSW112	1.8	41	20	0.03	0.27
MSW113	1.8	41	20	0.03	0.24
MSW116	1.8	41	40	0.03	0.22

Table 9.6 Operating parameters for the experiments carried out with γ-particles having a diameter of 3 mm

Run no.	d_p (mm)	T_g (°C)	φ_g (%)	u_g (cm/s)	X_{cr} (kg/kg)
MSW26	3	60	48	0.03	0.26
MSW38	3	40.5	2.5	0.1	0.4
MSW54	3	40.5	11	0.03	0.28
MSW66	3	41	50	0.03	0.25
MSW100	3	24.5	41	0	0.24
MSW105	3	41	4.3	0.03	0.29
MSW108	3	41	4	0.03	0.29
MSW111	3	41	20	0.03	0.27
MSW114	3	41	40	0.03	0.27
MSW115	3	41	40	0.03	0.27
MSW124	3	41	63	0.03	0.19
MSW147	3	41	32	0.03	0.27

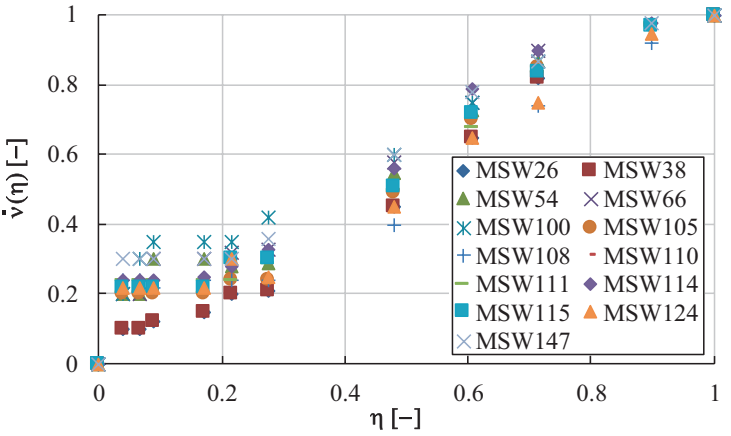


Fig. 9.23 Normalized drying curves for 3 mm γ-Al_2O_3 particles (operating parameters of the experiments to be found in Table 9.6)

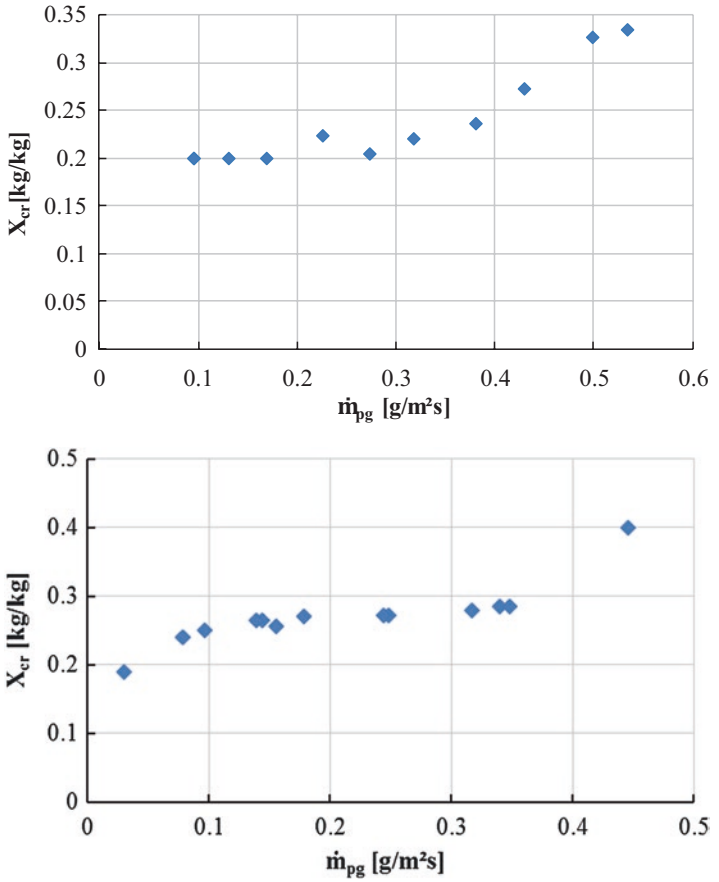


Fig. 9.24 Critical moisture content as a function of the drying rate for 1.8 mm (top) and 3 mm (bottom) γ-particles

(in *g* evaporated water per unit of evaporation area and time), and on the *y*-axis the critical moisture content *X* in kg of water per kg dry material is applied. It is clear that both curves show a range in which the critical moisture content remains nearly unchanged; however, this range is slightly wider for the coarse particles. This is consistent with the results discussed above that the influence of the operating conditions decreases with increasing particle diameter. The critical moisture content of coarse particles is about 25% higher for the same drying rate (same evaporation mass flow), that is, for the same operating conditions.

9.5 Simulation Results and Validation

The continuous drying experiments were performed using the fluidized bed bench scale plant described in Sect. 9.4 taking the following measures into consideration:

Preparation of test material: In order to ensure complete wetting of the interior of the particles, the material was stored in demineralized water for at least 24 hours. Independently of its initial moisture content, the material should be free of any surface moisture in order to reduce adhesion forces between the wet particles. Due to these adhesive forces, the fluidization in the next step would be too difficult, since the fluidizing air would channel through the product layer. In addition, agglomerates would be formed on the vibrating conveyor, which means a uniform feeding could not be ensured. For this purpose, a screen of suitable mesh size was used to drain the water.

Thereafter, the wet particles were mixed with sufficient dry particles having different particle sizes. The mixture was then screened. In this way a free-flowing and well-fluidizable material is obtained. For experiments that require a specific initial moisture content, the dewatered material was pre-dried in the fluidized bed to the desired moisture content in batch operation. The prepared material was then stored in a tightly sealed container. To determine the initial moisture content, several samples of approximately 50 g were taken.

Adjusting the test parameters: The air temperature, air mass flow rate, and particle mass flow were adjusted. The storage containers were filled with the wet material. Thereafter, the system was turned on and the measured values were recorded.

Sampling: Samples for determination of the average moisture content were carried out after reaching steady state. The steady state is characterized by a constant gas outlet humidity and temperature (Fig. 9.25). The sample was weighed and screened as quickly as possible to minimize further change of the moisture content. The dry substance was determined by drying the samples in a drying oven at a temperature of 180 °C until reaching constant weight. This high temperature was selected to ensure the relative humidity approaches zero, since the materials used in this work were highly hygroscopic especially γ-Al_2O_3. In addition, the residual moisture content is dependent on the condition of the surrounding air.

Termination of the trial: After taking the sample, the test was terminated by simultaneous shutdown of the air heater, the fan, and the continuous supply of particles. Then the static bed height and the mass of the solid in the dryer were determined.

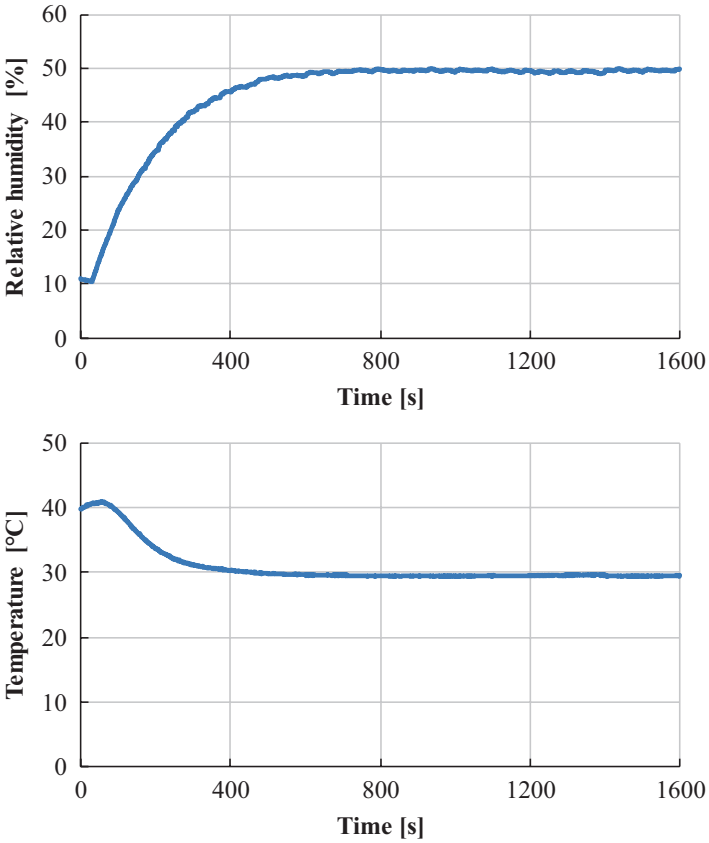


Fig. 9.25 Measured exhaust air relative humidity and temperature in a continuous fluidized bed. ($T_{g,in}$ = 60 °C, $\dot{M}_g$ = 67 g/s, φ = 37%, $\dot{M}_{p,in}$ = 1.85 g/s and M_p = 0.86 kg)

9.5.1 Model Validation

To validate the model, the influence of the operating parameters was studied. During continuous drying, the product has distributed characteristics because of its residence time distribution especially with respect to a moisture distribution. Cunäus et al. (2010), Kettner et al. (2006), and Peglow et al. (2011) have determined the moisture distribution of the product by single particle measurements. The measurement was performed by nuclear magnetic resonance spectroscopy.

However, in this work, only the average product moisture was determined. Since the calculated average moisture matches the experimentally determined average value, it is safe to assume that also the calculated moisture distribution represents the moisture distribution of the product. The SolidSim module is able to calculate the vapor loading, the temperature of the exhaust air (suspension and bubble gas), the product moisture distribution, as well as the temperature of the solid as a function of the residence time, particle size, and initial moisture content.

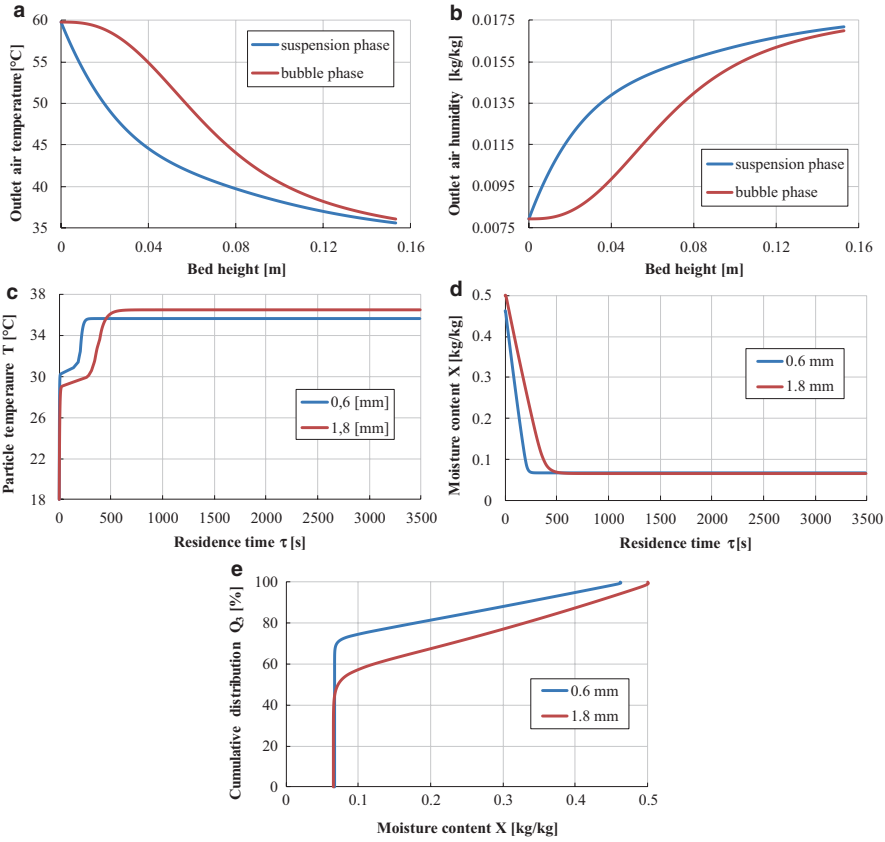


Fig. 9.26 Typical results obtained from SolidSim module during continuous, stationary drying. ($T_{g,in}$ = 60 °C, Y_g = 8 g/kg, two classes d_p = 0.6 mm (46%) and d_p = 1.8 mm (54%)). (**a**) Calculated gas temperature as a function of bed height, (**b**) Calculated gas humidity as a function of bed height, (**c**) Calculated temperature profile as a function of the residence time, (**d**) Calculated moisture profile as a function of the residence time, (**e**) Calculated moisture distribution of the product

Figure 9.26 shows typical SolidSim results. Figure 9.26a shows the calculated temperature profiles, and Fig. 9.26b shows the concentration profiles in the bubble and the suspension phase as a function of bed height.

Figure 9.26c and d shows the calculated temperature (C) and moisture content profiles (D) of the solid as a function of the residence time at steady and continuous drying. Remarkable is the temperature profile as a function of particle size. While at the beginning of drying and during the first drying phase fine particles show higher temperatures, they have lower temperatures than the coarse particles in the second drying period.

Finally, Fig. 9.26e describes the moisture distribution of the product for each of the shown particle sizes. Further discussions can be found in the following sections.

Influence of the Gas Mass Flow

In practical applications the air mass flow represents the operating parameter being most often varied. Therefore, experiments should investigate its influence on drying overall. Simultaneously, the influence of the particle size on drying will be examined. For this purpose, a mixture of γ-Al_2O_3 particles with a diameter of 1.8 and 0.6 mm was used. In order to avoid segregation as a result of vibration, both particle sizes were fed separately to the dryer, and the mixing was carried out directly on the vibrating conveyor.

The parameters of the tests and the mixing ratios are shown in Table 9.7.

Figure 9.27 shows the temperature and the vapor loading of the drying gas at the dryer outlet. The higher the gas mass flow, the higher the temperature and the lower the humidity of the gas. The measured gas temperatures tend to be somewhat lower than the calculated temperatures. This deviation increases with increasing the gas mass flow. The reasons for the deviation are the increasing heat loss and leak air through leaks in the testing plant. The amount of leak air is directly related to the prevailing pressure in the system. With increasing gas mass flow the pressure decreases and consequently the amount of sucked in air increases. This in turn increases the deviation between measured and calculated values. By increasing the gas mass flow at a constant gas temperature, the gas velocity increases, which results in increased heat and mass transfer coefficients leading to an improved drying behavior with drier product particles (Fig. 9.28).

The influence of the particle size is significant at low gas flow rates, while with increasing gas flow rates the agreement between experiments and model improves.

Figure 9.29 shows the change of moisture distribution for the two particle sizes depending on the gas mass flow. With increasing the gas mass flow, the curves shift upward, that is, the product leaves the dryer with lower average moisture and narrower moisture distribution.

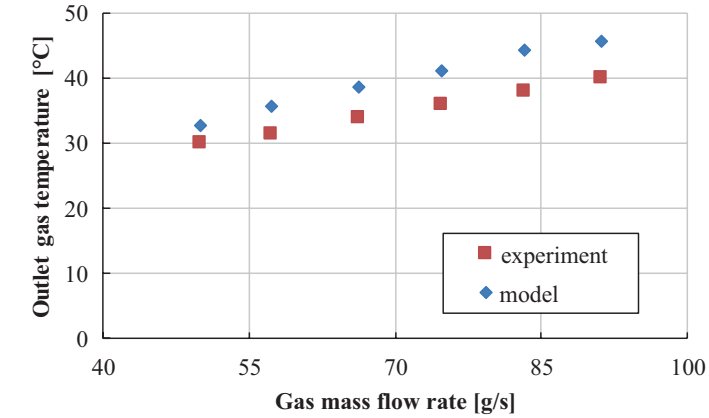
Since the drying gas exits warmer and in a drier state at higher mass flow, a lower equilibrium moisture content can therefore be expected (the curves shift to the left).

Influence of Gas Temperature

The temperature is the most important drying parameter. Both the product quality and the cost of operation are directly dependent on the temperature. Especially when heat-sensitive materials such as enzymes, vitamins in food, and active

Table 9.7 Operating parameters for the measurements of investigations into the influence of the gas mass flow

Run no.	$T_{g,\,in}$ (°C)	$Y_{g,\,in}$ (g/kg)	$\dot{M}_g$ (g/s)	$\dot{M}_p\left(g/s\right)$	Fine fraction (%)
72	60	7.85	50	1.58	46.7
73	60	7.94	57.4	1.58	46.4
74	60	7.77	66.2	1.58	46.6
76	60	7.92	74.7	1.6	47.2
71	60	5.67	83.3	1.55	44.8
78	60	7.9	91.2	1.55	45.4



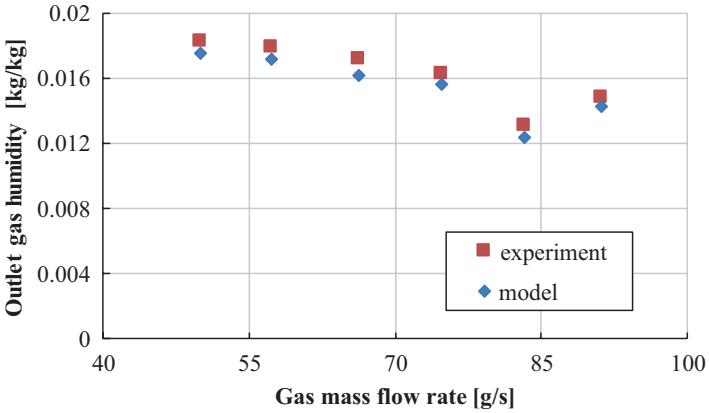


Fig. 9.27 Calculated and measured temperature and moisture concentration of the exiting drying gas. ($T_{g,in} = 60$ °C, $Y_g = 8$ g/kg, $\dot{M}_p = 1.6$ g/s two classes $d_p = 0.6$ mm (46%) and $d_p = 1.8$ mm (54%))

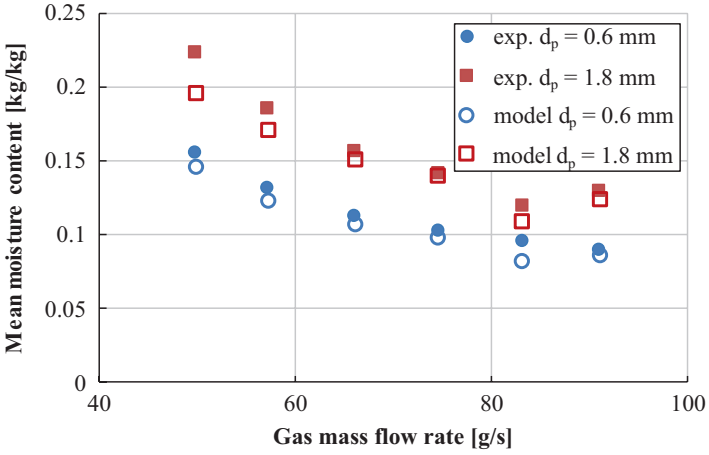


Fig. 9.28 Measured average particle moisture content and the corresponding model calculation of two particle sizes as a function of the gas mass flow rate. ($T_{g,in} = 60$ °C, $Y_g = 8$ g/kg, $\dot{M}_p = 1.6$ g/s two classes $d_p = 0.6$ mm (46%) and $d_p = 1.8$ mm (54%))

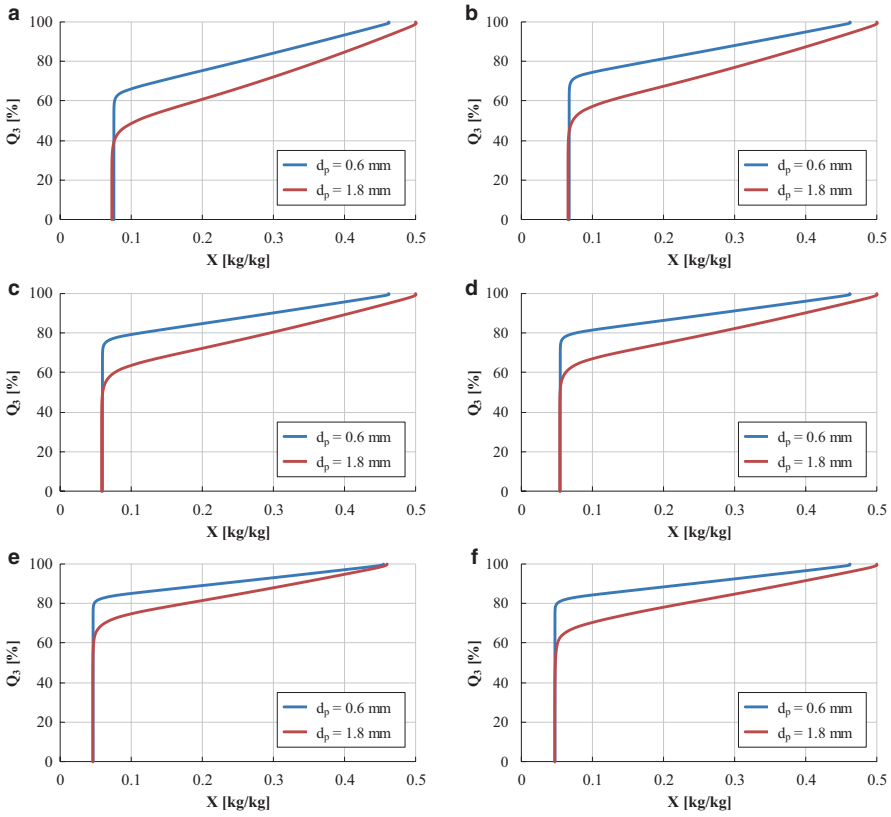


Fig. 9.29 Calculated particle moisture distribution of two particle sizes using SolidSim module (operating parameters of the experiments to be found in Table 9.7) (**a**)$\dot{M}_g = 50\,\mathrm{g/s}$, (**b**) $\dot{M}_g = 57.4\,\mathrm{g/s}$, (**c**) $\dot{M}_g = 66.2\,\mathrm{g/s}$, (**d**) $\dot{M}_g = 74.7\,\mathrm{g/s}$, (**e**)$\dot{M}_g = 83.3\,\mathrm{g/s}$, (**f**) $\dot{M}_g = 91.2\,\mathrm{g/s}$

pharmaceutical substances are involved, the temperature plays a crucial role for product quality.

At a higher temperature, the air has a higher saturation vapor pressure and thus a higher capacity for uptake of the water vapor.

The average moisture content as well as the exiting air properties with varying gas inlet temperatures is shown in Figs. 9.30 and 9.31.

The higher the temperature of the drying gas, the higher the temperature of the gas in the bubble and in the suspension phase, so that a higher evaporation rate can be expected. This in turn leads to an increase of the average temperature and a reduction of the average moisture content of the product at the dryer outlet.

In Fig. 9.31 it can be seen that the average moisture content of each particle size varies greatly with low gas inlet temperatures and only slightly at high temperature so that at a temperature of 100 °C the mean moisture contents for both particle sizes will approach each other.

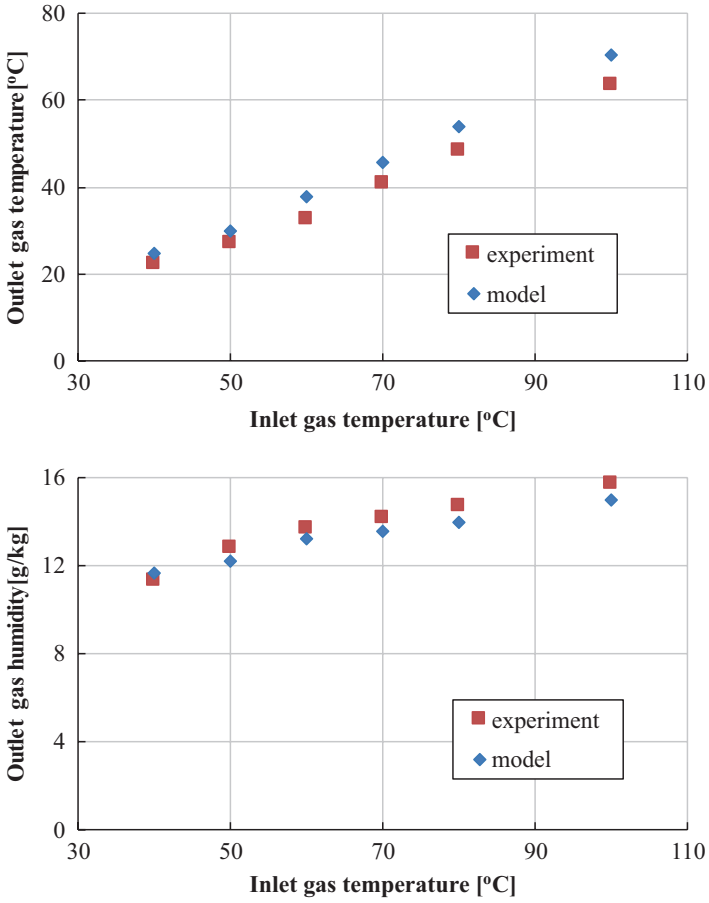


Fig. 9.30 Calculated and measured temperature (top) and moisture content (bottom) of the exiting drying gas. (Y_g = 4.5 g/kg, $\dot{M}_g$ = 67 g/s, $\dot{M}_p$ = 1.6 g/s and two classes d_p = 0.6 mm (50%) and d_p = 1.8 mm (50%))

The moisture distributions of each size classes at different gas temperatures are shown in Fig. 9.32. The higher the inlet gas temperature, the narrower the moisture distribution of the product, that is, the moisture distribution is shifted upward. On the other hand, the equilibrium moisture content decreases due to the prevailing high temperatures in the dryer, so the curves shift to the left with increasing temperatures. Comparing Fig 9.32a and f, it can be clearly seen that there is a significant difference between the two moisture distributions for each particle size at low temperatures.

The parameters of these experiments and the mixing ratios are given in Table 9.8.

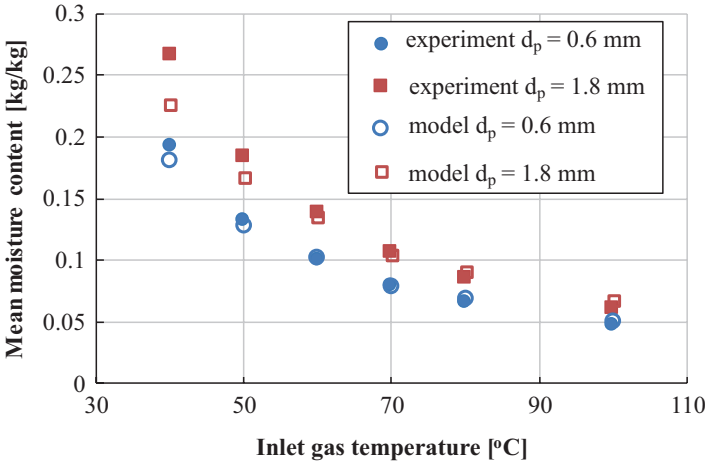


Fig. 9.31 Measured average particle moisture and the corresponding model calculation of two particle sizes as a function of the inlet gas temperature. ($T_{g,in}$ = 60 °C, Y_g = 8 g/kg, $\dot{M}_p$ = 1.6 g/s two classes d_p = 0.6 mm (46%) and d_p = 1.8 mm (54%))

Influence of Particle Mass Flow Rate

In these experiments, only the particle throughput through the continuous dryer was varied, while all other parameters were kept constant. The operating parameters are shown in Table 9.9.

The results of these measurements and the simulation results are given in Figs. 9.33 and 9.34. As can be expected for higher solids throughput (lower mean residence time), the moisture content of the solids at the outlet increases, that is, the moisture distribution is shifted downward and to the right. The minimal moisture content is lower for the low throughputs, since the outlet gas temperature and the solids temperature in the bed are higher, and therefore the equivalence moisture content is lower.

Influence of Inlet Moisture Distribution

The aim of these experiments is to investigate the influence of moisture distribution of the feed material on the moisture content of the product. An initial moisture distribution of the feed material can occur, for example, when two continuous dryers operate in series. For this purpose the 1.8 mm γ-particles are used. The material was pre-dried to two different moisture contents. To prevent moisture diffusion between the two wet materials, the materials were stored separately and also fed separately to the dryer. The mass fraction of the conveyed materials is set by varying the respective gap height above the trough of the vibrating conveyor.

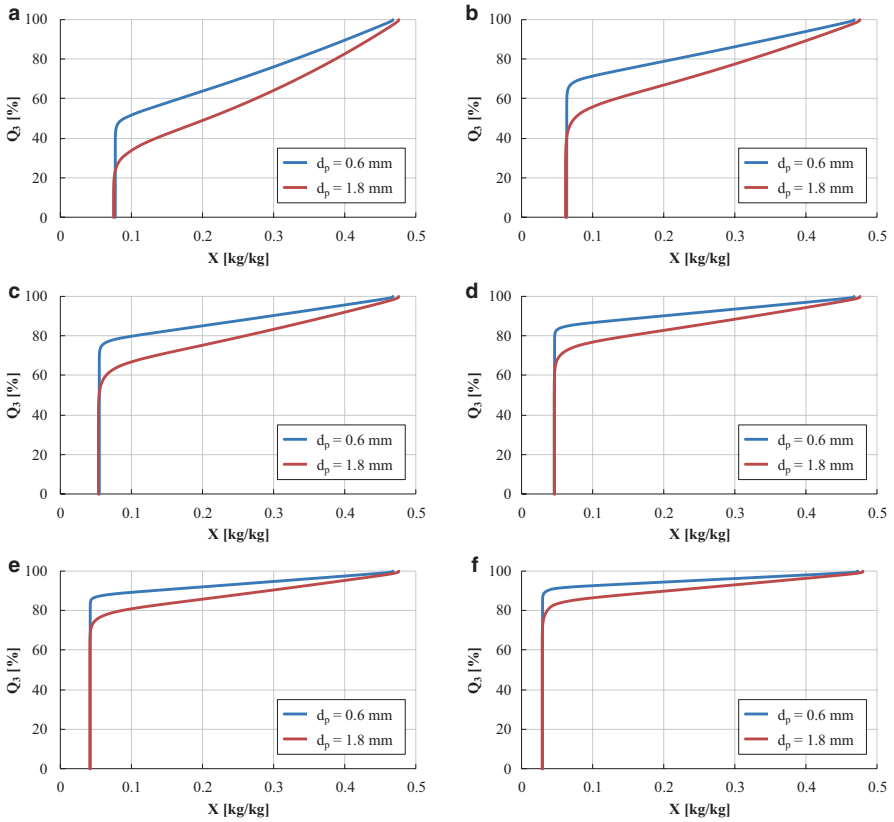


Fig. 9.32 Calculated particle moisture distribution of two particle sizes using the developed SolidSim module (operating parameters of the experiments to be found in Table 9.8). (**a**) $T_{g,\,in}$ = 40°C, (**b**) $T_{g,\,in}$ = 50°C, (**c**) $T_{g,\,in}$ = 60°C, (**d**) $T_{g,\,in}$ = 70°C, (**e**) $T_{g,\,in}$ = 80°C, (**f**) $T_{g,\,in}$ = 100°C

Table 9.8 Operating parameters for the experimental investigations of the influence of the gas temperature

Run no.	$T_{g,\,in}$ (°C)	$Y_{g,\,in}$ (g/kg)	$\dot{M}_g$ (g/s)	$\dot{M}_p\,(g/s)$	Fine fraction (%)
59	40	5.1	66.9	1.65	50.5
63	50	4.5	66.5	1.57	49.2
60	60	4.5	66.8	1.64	51
62	70	4.5	66.5	1.59	49.8
61	80	4.4	66.6	1.62	50.3
64	100	4.4	67	1.69	50.9

For comparison, the experiments were repeated with a uniform moisture content, which corresponds to the average moisture content of the first scenario. The operating parameters for the first and second scenarios are listed in Tables 9.10 and 9.11.

Table 9.9 Operating parameters for the measurements of the influence of particle throughput

Run no.	$T_{g,\,in}$ (°C)	$Y_{g,\,in}$ (g/kg)	$\dot{M}_g$ (g/s)	$\dot{M}_p\,(g/s)$	Fine fraction (%)
67	60	5.3	66.1	0.96	44.7
90	60	3.8	67.3	1.05	47.9
66	60	4.4	66.5	1.19	46.5
65	60	4.8	66.7	1.44	48.9
60	60	4.5	66.8	1.64	51
69	60	5.3	66.8	1.75	44
89	60	4.2	67.1	1.87	48.2
88	60	4.3	66.9	2.05	48.3

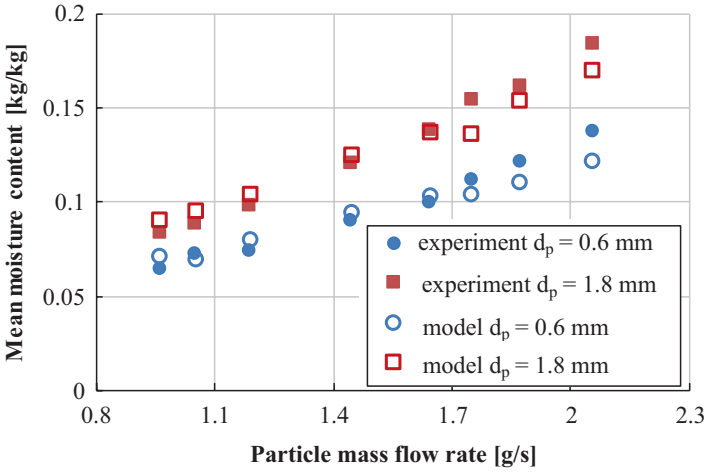


Fig. 9.33 Influence of the solids throughput on the mean moisture content

Figure 9.35 shows the measured mean particle moisture and corresponding model calculation as a function of the particle throughput. The difference between these two scenarios can hardly be seen. The reason could be that the measurement technology used with the defined accuracy cannot detect the difference, which means the variations between experimental and modeling results lie within the accuracy of the measuring devices.

Similar to the size distribution, the influence of the inlet moisture distribution plays an important role in the drying process as shown in Fig. 9.36. In Fig. 9.36 the actual values of the moisture distribution are plotted. It can be noted that with higher inlet moisture content, the moisture distribution shifts to the right, that is, the product has a higher moisture content at the outlet. The minimum moisture content is the same for both classes of the inlet moisture content, because the gas temperature and humidity are the same in the bed, and the particles with longer residence time will

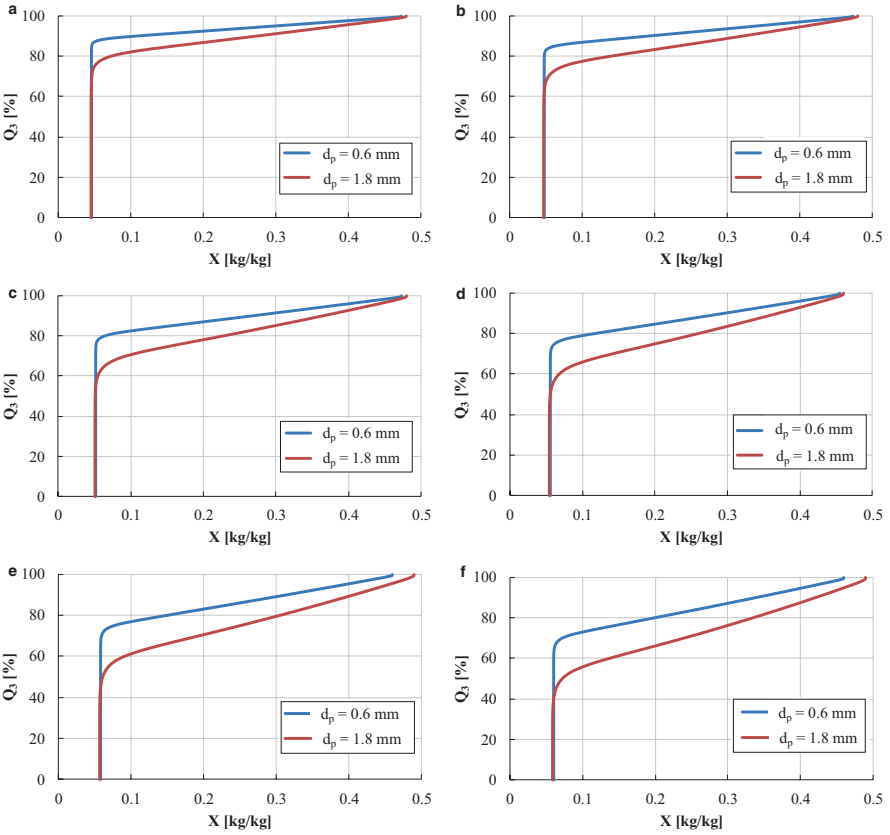


Fig. 9.34 Influence of the solids throughput on the moisture distribution (operating parameters of the experiments to be found in Table 9.9). (**a**) $\dot{M}_p = 0.96\,g/s$, (**b**) $\dot{M}_p = 1.19\,g/s$, (**c**) $\dot{M}_p = 1.44\,g/s$, (**d**) $\dot{M}_p = 1.75\,g/s$, (**e**) $\dot{M}_p = 1.87\,g/s$, (**f**) $\dot{M}_p = 2.05\,g/s$

Table 9.10 Operating parameters for investigation of the influence of moisture distribution of feed material (distributed moisture at the inlet)

Run no.	$T_{g,in}(°C)$	$Y_{g,in}(g/kg)$	$\dot{M}_g(g/s)$	$\dot{M}_p(g/s)$	$X_1(kg/kg)$	$X_2(kg/kg)$
94	60	4.7	50.5	1.38	0.257	0.695
95	60	4.6	50.7	2.23	0.257	0.695
96	60	4.5	50.7	1.56	0.257	0.695
97	60	5	50.7	1.76	0.257	0.695
98	60	5	50.8	1.14	0.253	0.706
99	60	4.8	50.7	0.91	0.253	0.706

in any case dry down to equilibrium moisture content. The maximum moisture content of each class is the inlet moisture content. These are particles which have a residence time close to zero and which will not be dried at all.

Table 9.11 Operating parameters for investigation of the influence of moisture distribution of feed material (uniform moisture at the inlet)

Run no.	$T_{g,in}$ (°C)	$Y_{g,in}$ (g / kg)	$\dot{M}_g$ (g / s)	$\dot{M}_p$ (g / s)	X_{in} (kg / kg)
100	60	4.2	50.9	1.07	0.48
101	60	4.5	50.8	1.01	0.48
102	60	5	50.7	1.53	0.49
103	60	5	50.7	1.71	0.49
104	60	6.5	5.6	1.94	0.49

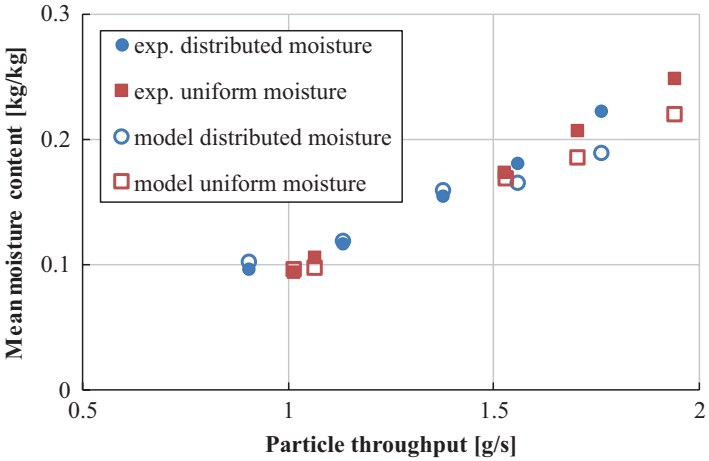


Fig. 9.35 Measured average particle moisture content at the dryer outlet and corresponding model calculation depending on the particle throughput for distributed and uniform moisture of the feed material

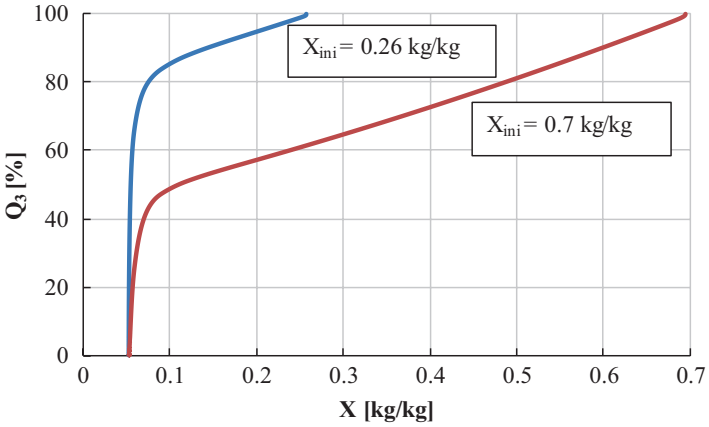


Fig. 9.36 Influence of the initial moisture distribution. ($T_{g,in}$ = 60 °C, T_g = 5 g/kg, $\dot{M}_g$ = 51 g/s, $\dot{M}_p$ = 1.55 g/s) with two moisture classes $X_{ini,1}$ = 0.26 kg/kg and classes $X_{ini,2}$ = 0.7 kg/kg each 50% mass fraction

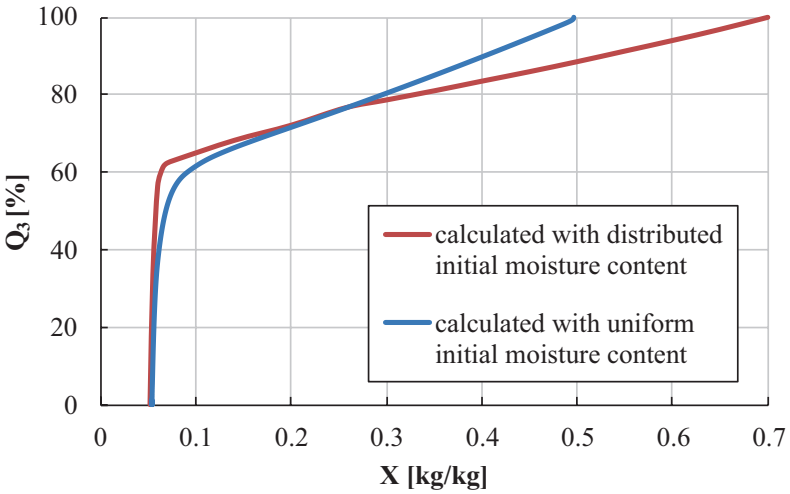


Fig. 9.37 Product moisture distribution calculated with distributed and average moisture content of the feed material. ($T_{g,in}$ = 60 °C, Y_g = 5 g/kg, $\dot{M}_g$ = 51 g/s, $\dot{M}_p$ = 1.55 g/s)

To study the influence of a moisture distribution at the inlet of the dryer on the drying process in Fig. 9.37, the results calculated for two different cases are shown. For both cases the average moisture of the feed is the same, but for case 1 the feed has two different moisture classes, while for case 2 all the particles have the same moisture content.

The results are shown in Fig. 9.37. There is a significant difference in the product moisture distribution of both cases. In case 2 (uniform initial moisture) the maximum moisture content of the product is naturally lower, as there are no particles with such high moisture content as the moist particles in the case of distributed initial moisture. In case 2 13% of the product shows a moisture content between 0.5 and 0.7 kg/kg, which could not be detected by the first case.

The equilibrium moisture content varies a little, because in the first case, the drying gas is more humid. It is noticeable that there is not only a difference between the product moisture distribution for the two cases, but also the average moisture with 0.164 kg/kg for the first case (distributed initial moisture) and 0.153 kg/kg for the second case differs (uniform initial moisture).

Drying of Mixture of Two Solids

The drying of a mixture of solids is encountered in the food industry. Quality and product shelf life are directly dependent on the moisture content of the individual particles. Particularly in the case of continuous drying, it is important to determine the moisture content and the moisture distribution of the individual components and the mixture to assure them to be well within the safe limit of the product.

In order to commence this simulation, a mixture of two different types of alumina particles was used. In addition to γ-alumina, α-alumina was used as a second type of solid. A detailed description of the two materials can be found in Sect. 9.4.1.

To prevent segregation caused by density and size differences, two separate bins are used. The mass fraction of the conveyed materials is set by varying the respective gap height above the trough of the vibrating conveyor. The operation parameters can be found in Table 9.12.

Figure 9.38 shows the measured and calculated average moisture content of the two materials as a function of total particle throughput.

In Fig. 9.39 the moisture content as function of residence time for α-Al_2O_3 and γ-Al_2O_3 is shown. γ-Al_2O_3 dries faster than α-Al_2O_3 at the beginning of the drying process. This is due to the lower density, which results in a larger surface area per mass unit of solids than for the α-Al_2O_3. But, due to the smaller porosity of the α-Al_2O_3, only a small amount of water has to be evaporated during the second drying period. Therefore, α-Al_2O_3 has lower moisture content than γ-Al_2O_3 after a residence time of more than 250 s.

Table 9.12 Operating parameter for drying of binary mixture

Run no.	$T_{g,in}$ (°C)	$Y_{g,in}$ (g/kg)	$\dot{M}_g$ (g/s)	$\dot{M}_p$ (g/s)	α portion (%)
110	60	4.2	50.88	1.07	53.3
111	60	4.5	50.83	1.01	53.3
113	60	5	50.67	1.53	52.4
114	60	5	50.71	1.71	51.3
115	60	6.5	5.56	1.94	54.7

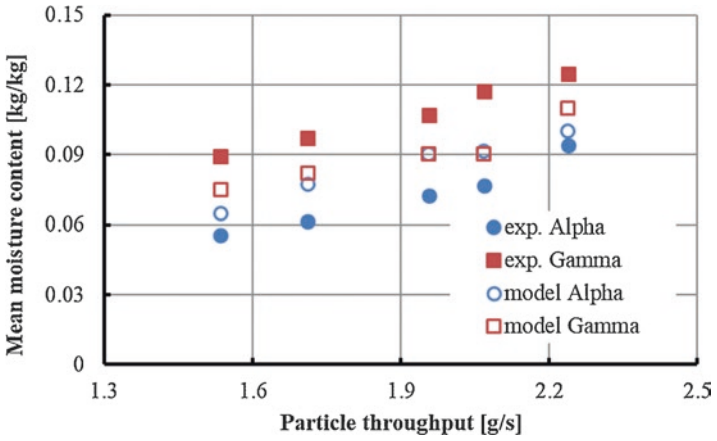


Fig. 9.38 Measured and calculated mean moisture content of a binary mixture of α- and γ-particles

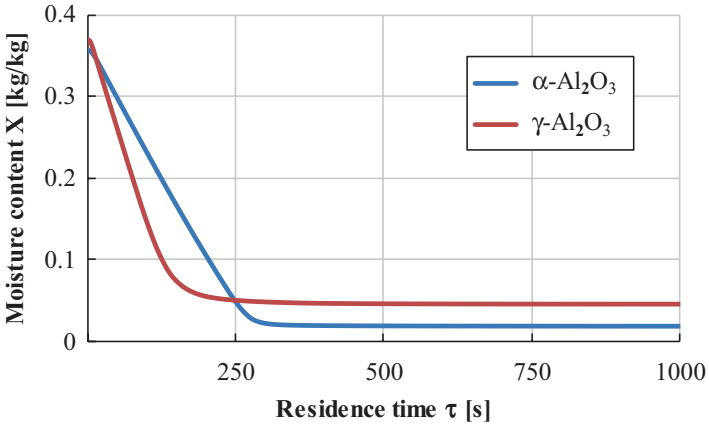


Fig. 9.39 Calculated moisture content as a function of residence time

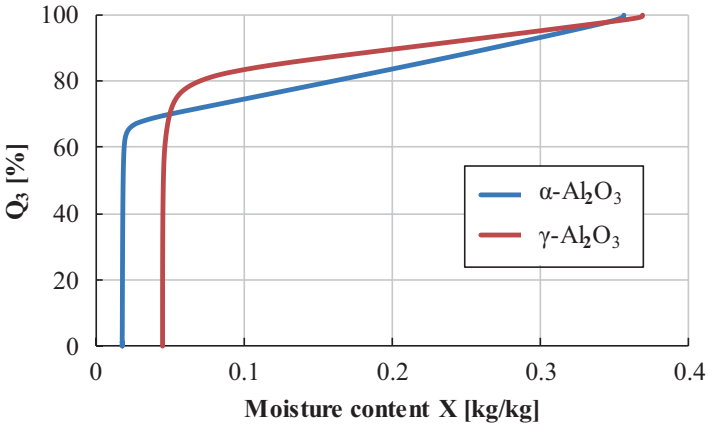


Fig. 9.40 Calculated moisture distribution of a binary mixture of α- and γ-alumina

In Fig. 9.40 the corresponding moisture distributions are shown. For the α-Al_2O_3 about 65% of the mass of the material is nearly completely dried down to about 0.02 kg/kg, but on the other hand more than 25% of the material has a remaining moisture content of more than 0.1 kg/kg.

For γ-Al_2O_3 there is no material with moisture less than 0.045 kg/kg; most of the material has moisture content between 0.045 and 0.06 kg/kg. Only 16% shows moisture contents greater than 0.1 kg/kg. In total this gives moisture content of 0.076 kg/kg for the α-Al_2O_3 and of 0.079 kg/kg for the γ-Al_2O_3.

9.6 Conclusion

Based on numerous works from the literature, especially the works by Burgschweiger and Tsotsas (2002), Cunäus et al. (2010), and Peglow et al. (2011), a drying model has been developed and inserted as a module in the flowsheet simulation program SolidSim. In addition to the residence time distribution described by a population model, the influence of particle size distribution and the moisture distribution of the feed material was taken into account. The module computes the vapor loading and temperature of the outflowing gas and the moisture and temperature distribution of the solids at the outlet taking into account their residence time distribution, their particle size, and their initial moisture distribution. The implementation within the flowsheet simulation program allows the simulation and optimization of the entire production process.

For the validation of the model, the operating parameters were varied, and the experimentally measured data were compared with those calculated by the model. The model can calculate the moisture distribution at the dryer outlet, but due to technical restriction only the average humidity was measured and compared with the model prediction.

For the experimental validation, aluminum oxide was used in both γ- and α-modifications.

The model generally shows a good agreement with the experimental results, only for the drying of several solids; it shows a slightly larger deviation due to a higher density difference and a broad particle size distribution. In this case the segregation had to be taken into account.

In this work, the influence of particle size and moisture distribution of the feed material on the continuous drying process of porous hygroscopic material was experimentally investigated and by means of the developed model with a good accuracy depicted. Furthermore, the drying behavior of different solid materials is examined.

The model enables the user to predict not only the average properties of the product but also the distributed properties such as moisture content and temperature of the entire product and each class.

Appendix A

A.1 Mass and Heat Transfer Between Particles and Suspension Gas

Mass and heat transfer between particles and suspension gas after Gnielinski (1980) for fixed bed	
$\mathrm{Re} = \frac{\mathrm{Re_{mf}}}{\varepsilon_{\mathrm{mf}}}$	$\mathrm{Re} = \frac{\mathrm{Re_{mf}}}{\varepsilon_{\mathrm{mf}}}$
$\mathrm{Sc} = \frac{\nu_{\mathrm{g}}}{\delta_{\mathrm{w,g}}}$	$\mathrm{Pr} = \frac{\nu_{\mathrm{g}} c_{\mathrm{g}} \rho_{\mathrm{g}}}{\lambda_{\mathrm{g}}}$
$\mathrm{Sh_{lam}} = 0.664 \mathrm{Re}^{\frac{1}{2}} \mathrm{Sc}^{\frac{1}{3}}$	$\mathrm{Nu_{lam}} = 0.664 \mathrm{Re}^{\frac{1}{2}} \mathrm{Pr}^{\frac{1}{3}}$
$\mathrm{Sh_{tur}} = \frac{0.037 \mathrm{Re}^{0.8} \mathrm{Sc}}{1 + 2.443 \mathrm{Re}^{-0.1} \left(\mathrm{Sc}^{\frac{2}{3}} - 1 \right)}$	$\mathrm{Nu_{tur}} = \frac{0.037 \mathrm{Re}^{0.8} \mathrm{Pr}}{1 + 2.443 \mathrm{Re}^{-0.1} \left(\mathrm{Pr}^{\frac{2}{3}} - 1 \right)}$
$\mathrm{Sh_p} = 2 + \sqrt{\mathrm{Sh_{lam}^2} + \mathrm{Sh_{tur}^2}}$	$\mathrm{Nu_p} = 2 + \sqrt{\mathrm{Nu_{lam}^2} + \mathrm{Nu_{tur}^2}}$
$\mathrm{Sh_{ps}} = [1 + 1.5(1 - \varepsilon_{\mathrm{mf}})]\mathrm{Sh_p}$	$\mathrm{Nu_{ps}} = [1 + 1.5(1 - \varepsilon_{\mathrm{mf}})]\mathrm{Nu_p}$
Back-mixing effect after Groenewold and Tsotsas (1999)	
$\mathrm{Sh_{ps}^{'}} = \frac{\mathrm{Re_0 Sc}}{A/F} \ln\left(1 + \frac{\mathrm{Sh_{ps}} A/F}{\mathrm{Re_0 Sc}}\right)$	$\mathrm{Nu_{ps}^{'}} = \frac{\mathrm{Re_0 Pr}}{A/F} \ln\left(1 + \frac{\mathrm{Nu_{ps}} A/F}{\mathrm{Re_0 Pr}}\right)$
$\mathrm{Sh_{ps}^{'}} = \frac{\beta_{\mathrm{ps}} d_{\mathrm{p}}}{\delta}$	$\mathrm{Nu_{ps}^{'}} = \frac{\alpha_{\mathrm{ps}} d_{\mathrm{p}}}{\lambda_{\mathrm{g}}}$

where Re Reynolds number, Sc Schmidt number, $\varepsilon_{\mathrm{mf}}$ minimal fluidization velocity, and Pr Prandtl number.

A.2 Mass and Heat Transfer Between Suspension Gas and Bubble Phase

The Number of Transfer Unit (NTU) is given as

$$\mathrm{NTU_{sb}} = \frac{\rho_{\mathrm{g}} \cdot \beta_{\mathrm{sb}} A_{\mathrm{sb}}}{\dot{M}_{\mathrm{g}}}$$

It is assumed that NTU increases linearly with the height; for a bed height of 5 cm, a value of 1 was set (Groenewold and Tsotsas 1999); therefore

$$\mathrm{NTU}_{\mathrm{sb}} = \mathrm{NTU}_{\mathrm{sb}}^{0} \cdot \frac{L_{\mathrm{bed}}}{50\,mm}$$

$$\frac{\alpha_{\mathrm{sb}} A_{\mathrm{sb}}}{c_{\mathrm{g}} \dot{M}_{\mathrm{g}}} = \frac{\rho_{\mathrm{g}} \beta_{\mathrm{sb}} A_{\mathrm{sb}}}{\dot{M}_{\mathrm{g}}} \mathrm{Le}^{1-m}$$

$$\mathrm{Le} = \frac{\lambda_{\mathrm{g}}}{c_{\mathrm{g}} \rho_{\mathrm{g}} \delta_{\mathrm{g}}}$$

$$m = \frac{1}{3}$$

A.3 Heat Transfer Between the Particles and the Wall

The heat transfer coefficient between the particles and the wall is calculated after (Martin 1980).

$$\mathrm{Nu}_{\mathrm{pw}} = \frac{\alpha_{\mathrm{pw}} d_{\mathrm{p}}}{\lambda_{\mathrm{g}}}$$

$$\mathrm{Nu}_{\mathrm{pw}} = \left(1-\varepsilon\right) Z \left(1 - e^{-N}\right)$$

With the coefficients

$$Z = \frac{1}{6} \frac{\rho_{\mathrm{p}} c_{\mathrm{p,wet}}}{\lambda_{\mathrm{g}}} \sqrt{\frac{g d_{\mathrm{p}}^{3} \left(\varepsilon - \varepsilon_{\mathrm{mf}}\right)}{5\left(1-\varepsilon_{\mathrm{mf}}\right)\left(1-\varepsilon\right)}}$$

$$N = \frac{\mathrm{Nu}_{\mathrm{pw(max)}}}{C_{\mathrm{K}} Z}$$

$$C_{\mathrm{K}} = 2.6$$

The maximal Nusselt number can be calculated from

$$\mathrm{Nu}_{\mathrm{pw(max)}} = 4\left\{\left(1 + \frac{2l}{d_{\mathrm{p}}}\right) \ln\left(1 + \frac{d_{\mathrm{p}}}{2l}\right) - 1\right\}$$

with

$$l = 2\left(\frac{2}{\gamma} - 1\right)\sqrt{\frac{2\pi \tilde{R} T}{\tilde{M}_\mathrm{g}}}\,\frac{\lambda_\mathrm{g}}{p\left(2c_\mathrm{g} - \tilde{R}/\tilde{M}_\mathrm{g}\right)}$$

$$\lg\left(\frac{1}{\gamma} - 1\right) = 0.6 - \frac{\frac{1000\,K}{T_\mathrm{g}} + 1}{C_\mathrm{A}}$$

$$C_\mathrm{A} = 2.8$$

where l is the modified free path of the gas molecules.

References

Alaathar I, Hartge E-U, Heinrich S, Werther J. Modelling and flowsheet simulation of continuous fluidized bed dryers. Powder Technol. 2013;238:132–41.

Burgschweiger J, Tsotsas E. Experimental investigation and modelling of continuous fluidized bed drying under steady-state and dynamic conditions. Chem Eng Sci. 2002;57(24): 5021–38.

Burgschweiger J, Groenewold H, Hirshmann C, Tsotsas E. From hygroscopic single particle to batch fluidized bed drying kinetics. Can J Chem Eng. 1999;77(2):333.

Chandran AN, Rao SS, Varma YBG. Fluidized bed drying of solids. AICHE J. 1990;36(1):29–38.

Cunäus U. Populationsdynamische Beschreibung der kontinuierlichen Wirbelschichttrocknung. Barleben: Docupoint-Verl; 2010.

Cunäus U, Peglow M, Tsotsas E, Metzger T. Modelling of continuous fluidized bed drying using a population balance approach. In: The Ethan international drying symposium IDS 2008, Hyderabad, India. 2008.

Cunäus U, Peglow M, Tsotsas E. Application of population balance equations for continuous fluidized bed drying. In: The 17th international drying symposium (IDS2010), Magdeburg, Germany. 2010.

Davidson JF, Harrison D. Fluidized particles. New York: Cambridge University Press; 1963.

Davies RM, Taylor G. The mechanics of large bubbles rising through extended liquids and through liquids in tubes. Proc R Soc Lond A. 1950;200(1062):375–90.

Fyhr C, Kemp IC. Comparison of different drying kinetics models for single particles. Dry Technol. 1998a;16(7):1339–69.

Fyhr C, Kemp IC. Evaluation of the thin-layer method used for measuring single particle drying kinetics. Chem Eng Res Des. 1998b;76(7):815–22.

Garnavi L, Kasiri N, Hashemabadi SH. Mathematical modeling of a continuous fluidized bed dryer. Int Commun Heat Mass Transf. 2006;33(5):666–75.

Geldart D. Types of gas fluidization. Powder Technol. 1973;7(5):285–92.

Gnielinski V. Wärme- und Stoffübertragung in Festbetten. Chem Ing Tech. 1980;52(3):228–36.

Gnielinski L. Beitrag zum EInfluss der Gutsdicke auf den Trocknungsverlauf bei der Trocknung von porösen Stoffen. Doctoral thesis, University Magdeburg. 2009.
Groenewold H. Wirbelschichttrocknung mit indirekter Beheizung. Doctoral thesis, University Magdeburg. 2005.
Groenewold H, Tsotsas E. A new model of fluid bed drying. Dry Technol. 1997;15(6–8): 1687–98.
Groenewold H, Tsotsas E. Predicting apparent Sherwood numbers for fluidized beds. Dry Technol. 1999;17(7–8):1557–70.
Groenewold C, Moser C, Groenewold H, Tsotsas E. Determination of single-particle drying kinetics in an acoustic levitator. Chem Eng J. 2002;86(1):217.
Hartge EU, Pogodda M, Reimers C, Schwier D, Gruhn G, Werther JG. SolidSim – a tool for the flowsheet simulation of solids processes. Aufbereitungs-Technik: AT = Processing Preparation. 2006a;(1/2):42–51.
Hartge EU, Pogodda M, Reimers C, Schwier D, Gruhn G, Werther JG. Flowsheet simulation of solids processes. Kona Powder Part. 2006b;24:146–58.
Hilligardt K, Werther J. Local bubble gas hold-up and expansion of gas/solid fluidized beds. Ger Chem Eng. 1986;9:215–21.
Hilligardt K, Werther J. Influence of temperature and properties of solids on the size and growth of bubbles in gas fluidized beds. Chem Eng Technol. 1987;10(1):272–80.
Hirschmann C, Tsotsas E. Impact of pore structure on particle-side drying kinetics. In: Drying ’98: proceedings of the 11th international drying symposium, Thessaloniki. 1998.
Hirschmann C, Fyhr C, Tsotsas E, Kemp IC. Comparison of two basic methods for measuring drying curves: thin layer method and drying channel. In: Drying ’98: proceedings of the 11th international drying symposium, Thessaloniki. 1998.
Kannan C, Thomas P, Varma Y. Drying of solids in fluidized beds. Ind Eng Chem Res. 1995;34(9):3068–78.
Kemp I, Fyhr BC, Laurent S, Roques M, Groenewold C, Tsotsas E, Sereno A, Bonazzi C, Bimbenet J-J, Kind M. Methods for processing experimental drying kinetics data. Dry Technol. 2001;19(1):15–34.
Kettner C, Peglow M, Metzger T, Tsotsas E. Distributed product quality in fluidized bed drying. In: 15th international drying symposium 2006, Budapest, Hungary. 2006.
Kozanoglu B, Martinez J, Alvarez S, Guerrero-Beltrán JA, Welti-Chanes J. Influence of particle size on vacuum–fluidized bed drying. Dry Technol. 2011;30(2):138–45.
Kwapinski W, Tsotsas E. Characterization of particulate materials in respect to drying. Dry Technol. 2006;24(9):1083–92.
Martin H. Wärme- und Stoffübertragung in der Wirbelschicht. Chem Ing Tech. 1980;52(3):199–209.
Paláncz B. A mathematical model for continuous fluidized bed drying. Chem Eng Sci. 1983;38(7):1045–59.
Peglow M, Cunäus U, Kettner C, Metzger T, Tsotsas E. A population balance approach for continuous fluidized bed dryers. In: 6th European congress of chemical engineering (ECCE-6), Copenhagen. 2007.
Peglow M, Metzger T, Lee G, Schiffter H, Hampel R, Heinrich S, Tsotsas E. Modern drying technology. In: Tsotsas E, Mujumdar AS, editors. Experimental techniques, vol. 2. Weinheim: Wiley-VCH; 2009.
Peglow M, Cunaus U, Tsotsas E. An analytical solution of population balance equations for continuous fluidized bed drying. Chem Eng Sci. 2011;66(9):1916–22.
Richardson JF, Zaki WN. Sedimentation and fluidisation: part I. Chem Eng Res Des. 1954;75:S82–S100.
Suherman. Drying kinetics of granular and powdery polymers. Magdeburg: Docupoint-Verl; 2007.
Toomey RD, Johnstone HF. Gaseous fluidization of solid particles. Chem Eng Prog. 1952;48: 220–6.

Tsotsas E, Mujumdar AS. Modern drying technology. Volume 2. Experimental techniques. Wiley-VCH Verlag GmbH & Co. KGaA. 2009.

van Meel DA. Adiabatic convection batch drying with recirculation of air. Chem Eng Sci. 1958;9(1):36–44.

Vanecek V, Markvart M. Fluidized bed drying. London: Leonard Hill; 1966.

Werther J, Wein J. Expansion behavior of gas fluidized beds in the turbulent regime. AIChe Symp Ser. 1994;90(301):31–44.

Zahed AH, Zhu JX, Grace JR. Modelling and simulation of batch and continuous fluidized bed dryers. Dry Technol. 1995;13(1–2):1–28.

Chapter 10
Statistical Methods in Quality by Design and Process Analytical Technologies for Continuous Processes to Enable Real-Time Release

Vanessa Cárdenas, Juan G. Rosas, Carlos Pinzón, and Rodolfo J. Romañach

Abstract The large data sets obtained in real-time monitoring of batch and continuous manufacturing processes require statistical methods to extract information on the state of critical process parameters and quality attributes. This chapter presents useful statistical methods such as design of experiments (DoE), principal component analysis (PCA), and partial least squares (PLS) regression for monitoring both physical and chemical properties in pharmaceuticals. A discussion of variographic analysis is also included, presenting a promising method to evaluate the mixing of pharmaceutical formulations and address sampling errors. This chapter also presents statistical methods that are currently used to improve both batch and continuous processes. These methods are now used in commercial manufacturing of pharmaceutical products.

Keywords Quality by design · Process analytical technology · Design of experiments · Near-infrared spectroscopy · Raman spectroscopy · Principal component analysis · Partial least squares regression · Chemometrics

10.1 Introduction

Statistical methods are essential to fulfill the goals of quality by design (QbD), process analytical technologies (PAT), and continuous manufacturing. In the QbD era, continuous manufacturing should be visualized as (1) an integrated manufacturing

V. Cárdenas · C. Pinzón · R. J. Romañach (✉)
Center for Structured Organic Particulate Systems (C-SOPS), Department of Chemistry, Mayagüez, Puerto Rico
e-mail: rodolfoj.romanach@upr.edu

J. G. Rosas
Merck, Sharp & Dohme Limited, Cramlington, UK

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_10

process, (2) which includes process analytical technology for real-time process monitoring, and (3) implements engineering systems to control the quality of drug products (Lee et al. 2015). The ambitious goals of continuous manufacturing start with process design and development and require a solid understanding of material properties. The continuous process could involve the use of a Comil for dry coating of a poorly flowing active pharmaceutical ingredient (API) (Huang et al. 2015; Deng et al. 2015). The dry coated particles will then become part of the pharmaceutical process, and real-time analysis will be necessary to monitor manufacturing (Ropero et al. 2009; Beach et al. 2010). The control of continuous processing requires integrated specialized information and engineering systems with PAT to monitor and control the process in real time and achieve high-quality products throughout the entire manufacturing run (Singh et al. 2013, 2014; Joglekar et al. 2014; Fonteyne et al. 2015). In turn, this integrated system provides a significant amount of data that can be used for process optimization. Statistical methods are required to understand, compress, and extract the information contained in the data generated with PAT sensing technologies such as near-infrared (NIR) and Raman spectroscopy. These statistical methods and integrated specialized information and engineering systems are now used in commercial pharmaceutical manufacturing where a 28-hour continuous manufacturing run was recently reported (Vargas et al. 2018).

The term "statistical methods" is broadly defined within this chapter as all efforts to extract information from data. The term includes design of experiments and multivariate data analysis. The scientific approach, philosophy, and algorithms taught in the field of chemometrics are encompassed within the statistical methods discussed in this chapter.

The ideas discussed in this chapter are still new within the pharmaceutical industry, but are more widely adopted in the agricultural, food, chemical, and petrochemical industries. The "Making money with chemometrics" article published by Seasholtz in 1999 provides good insight into the adoption of statistical methods within the worlds' largest chemical company (Seasholtz 1999). These ideas are not new to the petrochemical industry which has for many years discovered new ways to obtain information to optimize processing parameters and obtain better control of processes. The petrochemical industry has also learned that this increase in process understanding brings greater profits (Petersen and Esbensen 2005). The much discussed PAT pharmaceutical initiative is based on the ideas first developed by Kowalski and collaborators at the Center for Process Analytical Chemistry who visualized process analytical chemistry (PAC) as a "worthy subdiscipline of analytical chemistry" requiring an interdisciplinary systems approach (Callis et al. 1987; Pell et al. 2014). The term PAC was not used by FDA to avoid giving the impression that these valuable concepts were limited to the chemistry of processes. Instead the term PAT emphasizes a systems approach for a broad industrial audience from different disciplines and organizational units (Hussain 2015). PAT has become a regulatory initiative, and a guidance has been developed to submit methods to the Food and Drug Administration (FDA) (U.S. Department of Health and Human Services FDA 2004).

Real-time monitoring of chemical and pharmaceutical processes is not easily achieved. Although a significant amount of data is generated, the information is usually not easily accessible. The information is often hidden within the massive data

acquired and must be extracted using multivariate "chemometric" methods. The importance of the chemometrics or statistical methods is widely recognized in the agricultural, food, and chemical industry to make the best use of the companies' assets and maximize productivity (Miller 2000). The pharmaceutical industry is quickly adopting these statistical methods for the analysis of data for both batch and continuous processes. These statistical methods are part of the training needs necessary to continue advancing pharmaceutical manufacturing science (Reklaitis et al. 2010; Rantanen and Khinast 2015; Aksu et al. 2012; de Matas et al. 2016; Moghtadernejad et al. 2018).

10.2 Designing PAT for a Continuous Manufacturing Process

Figure 10.1 shows a sketch of continuous tablet manufacturing processes developed within the Center for Structured Organic Particulate Systems (C-SOPS). The continuous manufacturing system involves mixing and direct compression in this scheme, but could also include roller compaction or continuous granulation (Osorio et al. 2015). The C-SOPS developed a system that consists of multiple loss in weight feeders on the top part of the platform. The feeders provide the multiple formulation components, and a Quadro Comil is used to break large agglomerates and improve flow properties. The milled material passes to a continuous mixer equipped with a horizontally rotating shaft to mix the blend as it travels through the tubular body. Once the material is mixed, it moves directly to the tablet press.

The continuous system includes PAT providing faster and nondestructive analysis that facilitates real-time quality control. Continuous processes are designed to operate at a steady state when the desired conditions are met. In continuous mixing, the near-infrared (NIR) instrument could be used to monitor and control an entire manufacturing run for 24 hours or more by taking spectra after powders exit the blender. A typical batch blending process could be monitored until an acceptable drug distribution is obtained. The monitoring of a batch process is designed to reach a state where materials are well mixed or transformed to the desired species. Once this goal is achieved, the process is stopped and the materials are moved to the next stage. Thus, a continuous process may be monitored for a longer time than a batch process. A continuous direct compression process of a commercial pharmaceutical product was recently monitored for 28 hours (Vargas et al. 2018).

A recent mathematical approach uses the residence time distribution (RTD) analysis in combination with PAT to meet regulatory requirements such as raw material traceability. For these determinations, a trace material is sent through the continuous system, and the dispersed pulses are monitored using NIR spectroscopy in specific locations of the integrated equipment. The data obtained in RTD experiments is further analyzed using mathematical modeling and statistical methods which will be explained throughout the chapter (Engisch and Muzzio 2016).

Advanced control strategies for continuous manufacturing were implemented by Singh et al. where NIR spectrometers are connected to a control platform that provides feedback to the engineering system (Singh et al. 2014). The spectroscopic

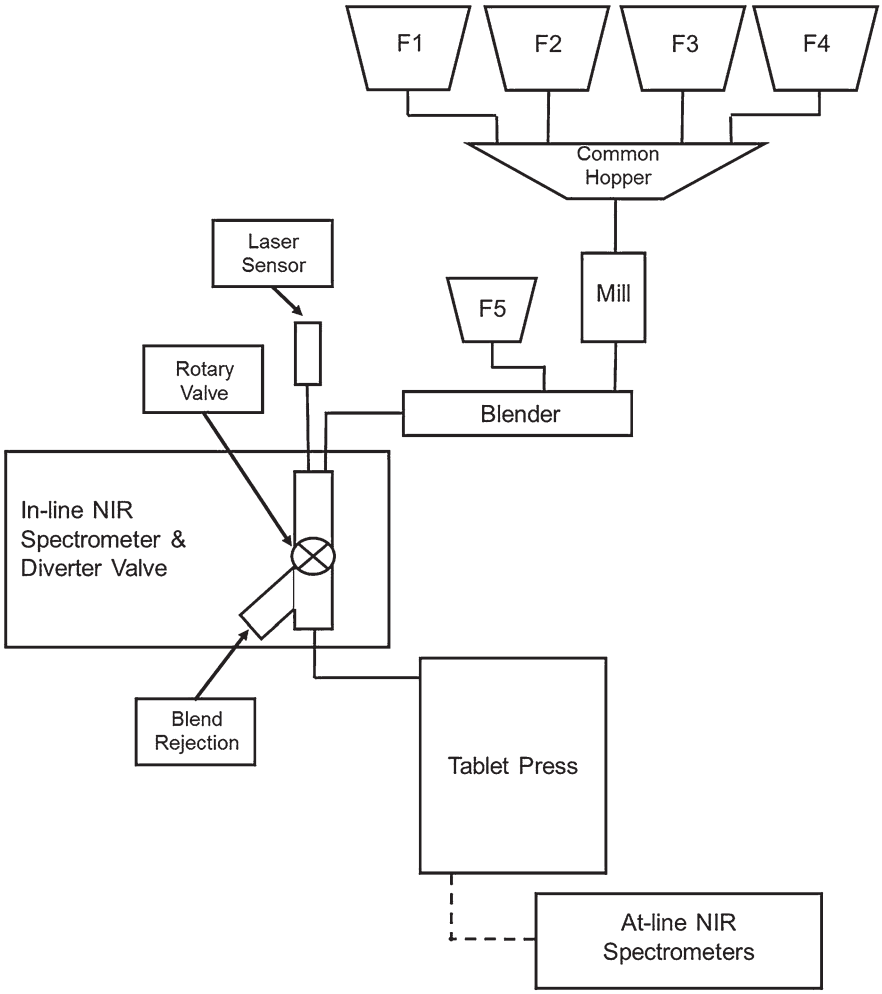


Fig. 10.1 Scheme of a continuous tablet manufacturing process. (Vargas et al. 2017)

signal and other type of alarms are interconnected to a specialized software, which in turn controls the manufacturing hardware, and the process can be stopped when a malfunction occurs. Figure 10.2 summarizes such a control system.

Continuous mixing processes have been monitored by real-time determination of the drug concentration during blending (Vanarase et al. 2010; Colón et al. 2014; Martinez et al. 2013). The statistical methods usually used are principal component analysis (PCA) and partial least squares (PLS) regression (Esbensen and Geladi 2009; Geladi and Kowalski 1986). These methods will be discussed below, but their advantages are described in this section. PCA allows the identification of the formulation components and may also indicate whether materials are adequately mixed or if an incorrect excipient has been used. PCA makes it possible to detect the absence

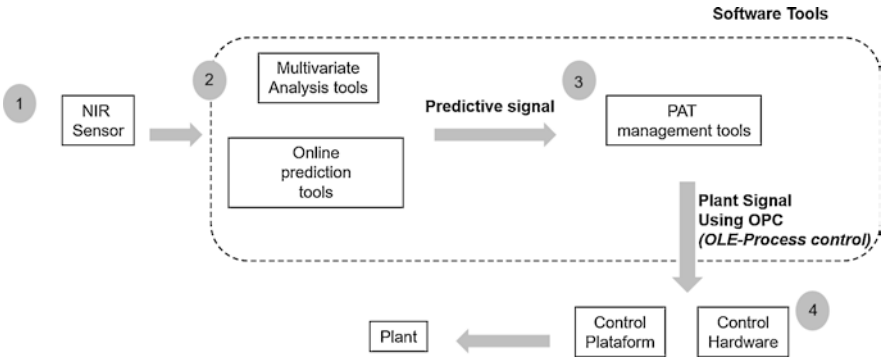


Fig. 10.2 Integration of NIR spectrometer, software, and hardware for continuous manufacturing control system

of material in front of a NIR probe. PLS permits the direct monitoring and quantification of a critical quality attribute such as drug concentration, particle size, and humidity and the detection of possible feeder malfunctions, reducing risk in continuous manufacturing.

NIR spectroscopy has been used to monitor continuous systems since spectra can be obtained while the process is occurring. NIR spectra can be obtained for pharmaceutical materials in their native state "as is" condition without the need to mill or dissolve particles as usually required by many quality control methods (Alcalà et al. 2012). NIR spectra may be obtained from powders within a blender or granulator, as the use of fiber optics facilitates the spectra acquisition in a noninvasive manner. The use of fiber optics permits the analysis of potent drugs through remote analysis, safeguarding the health of plant personnel, or the analysis of a hygroscopic material that has to be protected from moisture (Zhou et al. 1998). The same NIR spectrum can be used to determine physical and chemical properties of the samples such as drug concentration or density in a pharmaceutical formulation. Thus, NIR spectroscopy has been widely used in PAT applications. Table 10.1 describes a number of important elements for the design of the PAT system for this continuous manufacturing system.

NIR spectra could be obtained in several seconds or in the sub-second scale, depending on equipment design. The spectral acquisition time is an important parameter in the design of PAT for the continuous system as discussed in Table 10.1. A common design goal is for the spectral acquisition time to be at least 5–6 times faster than the fastest significant change expected in the system. This design goal is aimed at detecting a possible feeder malfunction or detecting a drug aggregate in time to control the system and avoid quality problems. The application of PAT is not simple and requires a thorough understanding of the system's dynamics.

The design of PAT for the continuous system does not stop with the consideration of the spectral acquisition time. The design should also consider the sample mass that interacts with the NIR radiation during the time that a spectrum is collected. For example, in continuous mixing, the goal is to develop the NIR method to

Table 10.1 Elements of design to be considered in PAT systems for continuous manufacturing

Design	Description
Spectral acquisition time	Varies according to instrument used, often selected to sample approximately a unit dose. Should be at least 5–6 times faster than the fastest change expected in the system
Sample mass analyzed by sensor	Varies according to a bulk density, irradiation spot diameter, sample movement, and number of spectra that are averaged (Rantanen and Khinast 2015)
In-line prediction of drug concentration	Development and validation of spectroscopic calibration model to determine drug concentration
Response time for control system	Drug concentration determined and relayed to control system to divert blend before compressing machine (Singh et al. 2013)
Data provenance	Storage of data in a way that can be associated with each of the process steps

analyze approximately the mass of a single dose unit. This mass can be estimated as shown in a number of publications, in which depth of penetration and bulk density must be determined (Berntsson et al. 1998). However, it must be emphasized that this is an *estimate*. The interaction of light with particles in diffuse reflectance methods is quite complex (Dahm and Dahm 2007). The exact sample mass analyzed is not known, only estimated. However, this estimate is necessary, as without it the analyst would not have an estimate of the mass of material being measured. The sample mass analyzed by the NIR spectrometer was estimated in a recent study by obtaining spectra of powder mixtures placed over an acrylic surface. The thickness of the powder bed was reduced, until the bands of the acrylic were observed at a thickness of 0.5 mm. The approximate sample mass – 312 mg – was calculated with the equation below that considers the sample movement while the spectra are being obtained. The estimate includes the sample bulk density (ρ), diameter (H), and depth of penetration (d) of the NIR beam, the sample displacement (Δx) (Colón et al. 2014):

$$M = \rho\left[\pi\left(\frac{d}{2}\right)^2 + d\left(\Delta x\right)\right]H \tag{10.1}$$

The design for an estimated sample mass usually requires the acquisition and averaging of a number of spectra (scans) to approximate the mass of a unit dose. The averaging of spectra is advantageous since this reduces the random spectral noise. In Fourier transform instruments, the noise is reduced by approximately the square root of the number of scans that are averaged as one spectrum (Griffiths and de Haseth 2007). The reduction in spectral noise leads to a greater signal to noise ratio that eventually helps in developing robust calibration models to determine the drug concentration. The spectra may also be obtained while powder flows through the system. Thus, each spectrum that is averaged corresponds to different increments of the powder blend. The spectral acquisition of the blend may be designed in this way to increase the mass analyzed to approximately the mass of a unit dose.

The average of the scans also constitutes a composite sample of the powder blend. This is important since the final product will also constitute a composite sample obtained from the powder blend. A recent study compared calibration models for drug concentration in powder blends obtained with (1) spectra obtained from different sections of a tray where a powder blend was deposited and (2) spectra obtained from a system where powder flowed as spectra were obtained (Mateo-Ortiz et al. 2014; Sierra-Vega et al. 2018). The accuracy of the calibration model improved by a factor of about 3 for the systems where powders flowed. This is likely due to the composite sampling that occurs when spectra are obtained while the powder flows.

The uncertainty in the sample mass may be overcome by the analysis of the final tablets (Vargas et al. 2018; Martinez et al. 2013). The drug concentration from tablets could be compared to the drug concentration results obtained with the in-line NIR system. If the results obtained from the in-line method and final product are similar, then the in-line method is being effective in spite of the uncertainty in the sample mass analyzed.

Quality by design and PAT are not limited to tablets, as shown in a continuous manufacturing system for strip films designed to deliver poorly soluble drugs (Zhang et al. 2014). This study showed the suitability of Raman spectroscopy for the in-line and off-line quantification of API concentration in strip films. The system allows the deposition of the nano-suspensions containing API in the film, followed by continuous drying, with monitoring in three different stages of the integrated system with Raman spectroscopy.

The implementation of PAT for the continuous mixing system for direct compression requires the development and validation of a calibration model to predict drug concentration in the blends and tablets obtained from the system. This challenge requires a careful development plan encompassing the expected variations in drug concentration and excipients in the manufacturing process (Romañach et al. 2016). Drug concentration methods are usually developed based on multivariate methods such as principal component analysis (PCA) and partial least squares (PLS) regression (10.6). Both algorithms may be implemented within the integrated equipment in Fig. 10.2 to provide data that is used as input to the control system. PCA is performed first to evaluate whether the spectra obtained are similar to those used in the PLS calibration model. This PCA evaluation is an example of the spectral quality test discussed in the EMA guidelines on submission of NIR tests, and several systems are commercially available for performing this calculation (European Medicine Agency E 2014). If the spectra are similar, the drug concentration is determined with the PLS calibration model. PCA could also indicate that the spectrum obtained is different from that expected. The root cause of this difference then needs to be determined. Thus, PCA-based supervisory control is now developed for a continuous manufacturing system (Singh et al. 2014).

The NIR system can also determine the drug concentration at the blender outlet, which then becomes the input for the controller to generate the set point for the ratio of the feeders shown in Fig. 10.1 (Singh et al. 2014). The control system then makes the necessary adjustments. At this point, the full benefits of the PAT system are being realized as the process is being monitored and controlled. The drug

concentration value obtained with the NIR spectrometer has become the input for controlling the blending process.

The drug concentration data must be captured and stored in context. Context is the key to relevant data as it documents the specific process steps that create the information and the conditions associated with each step. The process steps constitute the work flow, also called the provenance of the information (Joglekar et al. 2014). The real-time data may be compared to those of previous processes, and this knowledge could be used for making decisions. These decisions could include fault detection (Beach et al. 2010) and activation of alarms. Thus, an entire information structure has evolved to manage the data obtained from PAT systems (Singh et al. 2014; Vargas et al. 2018; Markl et al. 2013).

10.3 Information from NIR and Raman Spectroscopy

A basic background on NIR and Raman spectroscopy is needed to understand the continuous manufacturing data analyzed by statistical methods. The first method discussed is NIR spectroscopy which is based on the interaction of radiation with matter (Alcalà et al. 2012). The light which includes wavelengths that vary from 780 to 2500 nm interacts with particles, and the radiation is then transmitted, absorbed, or diffusely scattered by the particles. The real-time applications for continuous systems are based on diffuse scattering systems. In diffuse reflectance the light source and detector are on the same side of the sample.

NIR spectra are obtained from the ratio of the radiation that returns to a detector from the sample and background materials. The absorbance in NIR spectroscopy is usually termed as:

$$\text{Absorbance} = \log\left(1/R\right) = \log\left(I_0 / I_r\right) \tag{10.2}$$

$$R = I_r / I_0 \tag{10.3}$$

where I_r is the intensity of the radiation remitted to the diffuse reflectance detector and I_0 is the total radiation intensity remitted from the reference ceramic used in the NIR spectrometer for the background scan. The NIR baseline will be set at zero if I_r and I_0 have similar intensities. However, this is seldom the case for NIR spectra obtained from a pharmaceutical process. In pharmaceutical processes, spectra are obtained from flowing powders, formulations with different particle sizes, and uneven surfaces. In all of these cases, the fraction of light received by the detector is constantly varying. As powder flows and the radiation encounters an uneven surface, some of the radiation will be scattered in a direction where it is not collected by the optical system and not focused on the detector. Figure 10.3a shows the differences in baseline observed in NIR spectra. When only a fraction of the remitted radiation (I_r) reaches the detector (I_{det}), this relationship may be described by:

$$I_{\text{det}} = 1/c \times I_r \tag{10.4}$$

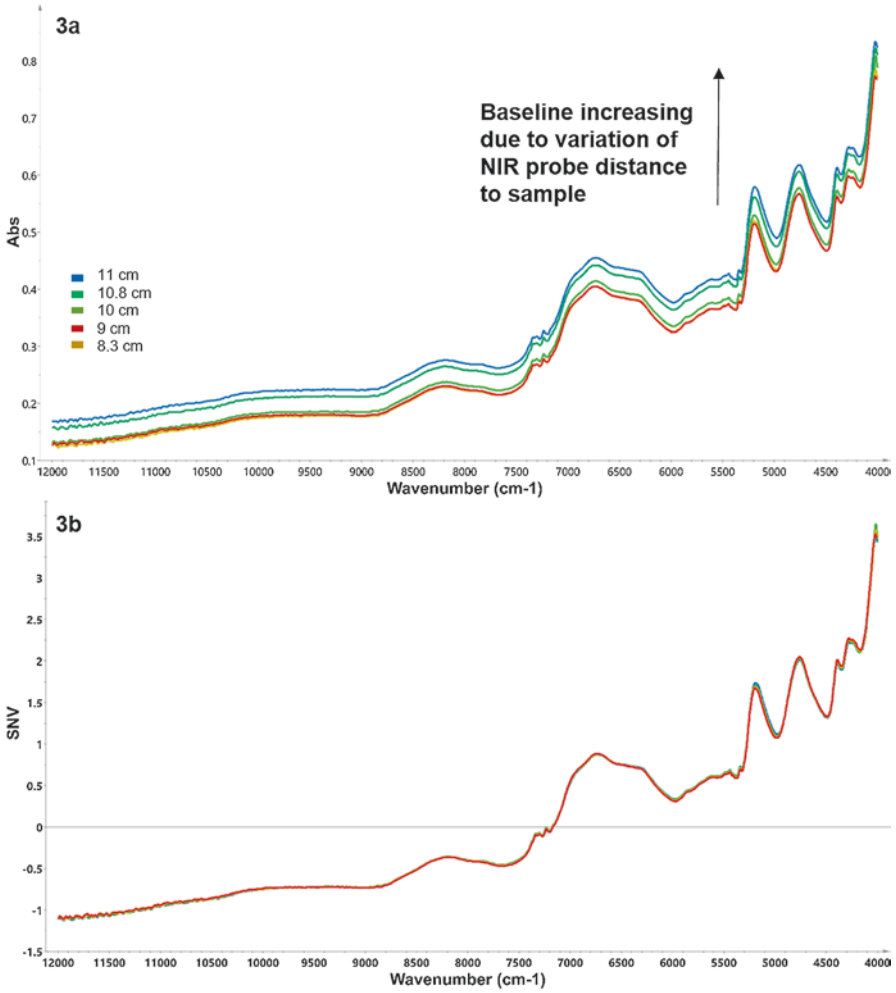


Fig. 10.3 (**a**) Differences in spectral baseline of pharmaceutical blends, (**b**) SNV scattering spectra correction

If I_r is substituted into Eq. (10.1) by I_{det} then:

$$\text{Absorbance} = \log\left(I_0 / I_{\text{det}}\right) = \log c + \log\left(I_0 / I_r\right) = c' + A \tag{10.5}$$

where A represents the spectrum and c' is on offset that explains the variation observed in the baseline (Ropero et al. 2009; Markl et al. 2013; Naes et al. 2002; Romañach et al. 2014). Baseline differences must be eliminated when the objective is to determine the concentration of a component in the formulation. Otherwise, the absorbance readings will be related to both the concentrations of the formulation components and the baseline differences. The statistical methods used require

working with spectral pretreatment or preprocessing methods to reduce spectral effects that hide the relevant information or to enhance bands related to the component of interest. The most common preprocessing methods are the standard normal variate (SNV) shown in Fig. 10.3b and derivatives that eliminate the baseline differences (Naes et al. 2002; Beebe et al. 1998).

The NIR spectral bands are overtones and combination bands of fundamental bands observed in the mid-infrared region. NIR bands are observed for hydrogen atoms bonded to larger atoms, namely, C-H, N-H, and O-H. These fundamental bands have high stretching frequencies that are very close to the NIR region, and thus their overtones are observed in the NIR region. NIR spectra are usually composed of wide overlapping bands, although some materials such as talc present sharp narrow bands. The bands observed in the NIR are 10–100 times weaker than the corresponding mid-IR fundamental bands, but this is usually not a concern for pharmaceutical applications where most of the components are present at the percent level. The differences for spectra obtained throughout a pharmaceutical process are often very subtle. However, NIR spectroscopists are well trained to work with these subtle differences, and statistical methods facilitate the extraction of information from NIR spectra (Esbensen and Geladi 2009).

A NIR spectrum can also provide information on the physical properties of pharmaceutical materials. NIR spectra can elucidate the particle size distribution of materials even if particles are flowing (Barajas et al. 2007). The NIR spectrum contains information of both the chemical and physical properties of materials, but the deconvolution of the two types of information is not easy. The spectral features related to the physical properties and chemical composition are so entwined that it is difficult to discern between the two, but easy to come to wrong conclusions. The use of statistics and multivariate analysis and an understanding of the physics of diffuse reflectance are necessary to discern between the impact of chemical changes and material property changes on the NIR spectrum (Dahm and Dahm 2007). Esbensen has indicated that "multivariate data analysis" could be compared to "driving a car" since both require mandatory theoretical and practical training (Esbensen 2012; Esbensen and Swarbrick 2018). This advice should be extended to NIR spectroscopy.

Many PAT applications include Raman spectroscopy which depends on scattering of the radiation (Smith and Dent 2005). Raman spectra may also be obtained for samples in their native form without milling or dissolving. The radiation from a laser interacts with an electron cloud, and the electrons in covalent bonds form a virtual state. The virtual state is not stable and only forms when the radiation interacts with the electrons in the covalent bond. Raman scattering will be stronger for bonds with higher electron density, such as double or triple bonds. If the electrons are pulled toward the nucleus and there is a strong dipole, Raman scattering will be low. The bonds with a strong dipole moment will show strong mid-infrared bands. Thus, it is often said that mid-infrared and Raman spectroscopy are complementary techniques.

In Raman spectroscopy the sample receives the radiation from the laser beam, and this radiation is quickly re-irradiated. Most of the radiation is elastically

scattered and does not provide useful information. Only one in every 10^6 or 10^8 photons is scattered at optical frequencies different from that of the incident laser radiation providing useful information. The weak Raman scattering is not a limiting factor due to the high power density of the lasers and the highly sensitive detectors. The Raman spectrum is obtained by subtracting the scattered signal from the laser signal and is expressed as a shift in energy from that of the exciting radiation. The difference in energy corresponds to frequencies in the infrared region providing an information-rich spectrum that may be used in elucidating the structure of pharmaceutical compounds.

The physical state of materials affects their Raman spectrum (Smith and Dent 2005). Crystalline materials give sharp strong spectra, while liquids and vapors give much weaker spectra. Raman spectra may also show subtle changes due to differences in pressure, crystal size, and orientation. The subtle changes observed may be used to understand intramolecular and crystal lattice vibrations (Żarów et al. 2011; Rozo et al. 2011). Thus, Raman spectroscopy is extremely valuable for the analysis of compounds that have one or more crystal structures (polymorphs) (Hu et al. 2005). Raman spectroscopy is also extremely useful in monitoring reactions or crystallizations in aqueous media due to the minor effect of water on the spectra obtained (Mercado et al. 2008; Wang et al. 2000).

Raman spectra often include a low-frequency noise that results from fluorescence of the compound analyzed or the presence of fluorescent impurities as shown in Fig. 10.4. The low-frequency noise is frequently observed in spectra obtained for formulations with microcrystalline cellulose. The effect of fluorescence can be overcome through baseline correction and pretreatment of the spectra with the standard normal variate approach. In spite of the problems with fluorescence, several Raman methods have been developed to determine coating thickness and uniformity (Romero-Torres et al. 2005, 2006; El Hagrasy et al. 2006).

10.4 Design of Experiments (DOE)

The development of a new continuous manufacturing process will involve a number of experiments to determine the optimum conditions for the system. The optimum conditions could be determined in just 2–3 days instead of the 30 days which could be required for a batch process (Osorio et al. 2015). The feeders and mixer could be programmed for the different conditions needed through a design of experiments (DOE) to optimize the manufacturing process. DOE involves a series of trials in which deliberate changes in the input variables (factors) of a process are effected to observe and identify the causes of changes in the output response. The methodology of design of experiments (DOE) comprises the systematic application of statistics to the experimentation. There are different types of experimental designs that have been applied in pharmaceutical research as shown in Table 10.2. In the pharmaceutical industry, the inputs might be the critical process parameters (CPP) and the outputs are critical quality attributes (CQA). DOE contributes to reducing the number

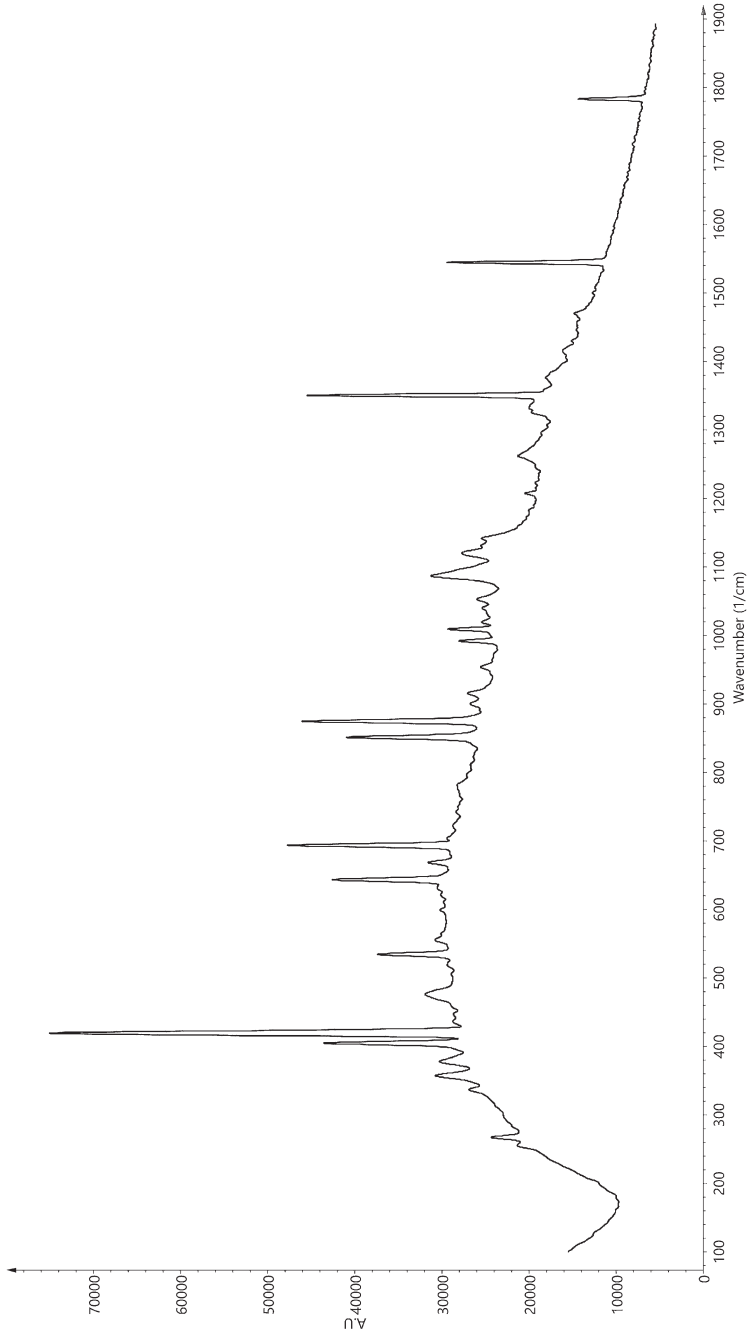


Fig. 10.4 Raman spectra of pharmaceutical blends. Note low frequency noise due to fluorescence

Table 10.2 Characteristics of the common DoE used in pharmaceutical research

Purpose of study	N° of factors to include	Model	Design
Screening	10–30	Qualitative variables	Fractional factorials (asymmetrical) (Box et al. 2005)
			D-optimal (Goupy and Creighton 2007)
		Linear (with some interactions)	Hadamard design (Agaian 1985)
			Fractional factorials (symmetrical) (Box et al. 2005)
			D-optimal
Study of effect	5–15	Linear with interactions	Fractional factorials (Box et al. 2005)
			Full factorials (Box et al. 2005)
			D-optimal
Optimization	2–7	Quadratic	Central composite designs (Box and Draper 2006)
			Doehlert designs (Doehlert 1970)
			D-optimal
			Fractional factorials (3 levels) (Box et al. 2005)
			Box-Behnken design (Box et al. 2005)
Mixtures	>3	Linear or quadratic	Type I (Schéffé designs) (Smith 2005)

of experiments without sacrificing the information obtained from the systems or obtaining a design space with the lowest number of experiments. The smaller volume of the mixers also makes it possible to perform the DOE with a much lower amount of the API and excipients than in a batch process. Thus, the continuous manufacturing setup facilitates the use of DOE to optimize a new process.

The choice of DoE depends on the number of factors that can be desired. The fundamental idea that should be considered in the experimental procedure is that DoE must be adapted to respond a specific problem. The following steps must be performed:

1. Define the aims of the study and criteria.
2. Identify the controlled and uncontrolled factors to include in the study and the experimental domain.
3. Build the DoE.
4. Conduct the experiments.
5. Analyze the results.
6. If linear model is inadequate, additional experiments might be added and a quadratic model can be evaluated.
7. Validate the results.
8. Obtain the conclusions on the study.

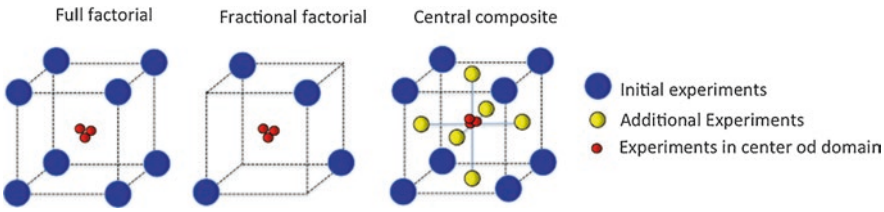


Fig. 10.5 Graphical comparison of three most used DoE, example with three factors, two levels each, and three replicates in the central of experimental domain. (Based on Eriksson et al. (2008))

The most widely DoE used are the full factorial (where all possible combinations are investigated), fractional factorials (only a fraction of all experiments is needed), and central composite (where additional axial experiments are proposed). Figure 10.5 shows a graphical comparison of the experimental domain with three factors where each point represents an experiment.

Full factorial designs consist of all possible combinations of levels for all factors and require a large number of experiments. Central composite designs are far more efficient in terms of number of experiments. As the number of factors and their levels increases, the designs become more complex and the experiments needed are numerous; in these cases other designs become attractive, and they are called the D-optimal designs. D-optimal designs allow to reduce considerably the number of experiments by minimizing the determinant of variance-covariance matrix of the estimated parameter.

10.5 Principal Component Analysis

Principal component analysis (PCA) is one of the most used algorithms to compress and extract information from large amounts of data (Miller 2010). The data set is represented by an **X** matrix where each column is a measured variable, and each row represents an object or sample as shown in Fig. 10.6. Thus, the **X** matrix may describe NIR spectra and also other measurements such as process variables or product properties.

PCA transforms a highly correlated bidimensional matrix **X** (NxM) into a new set of orthogonal axes that explains the maximum variation of the samples. The matrix can be constituted by N observations (spectra of samples) – that occupies the rows – and M number of variables (wavelength or wavenumber), located in the columns (Miller 2010). Thus, the decomposition of the covariance matrix gives us results in the following terms: scores **T** (describing relationship between samples) and loadings **P** (describing relationship between variables) plus information not explained by the model, residuals **E** (Kumar et al. 2014):

$$\mathbf{X} = \mathbf{T}.\mathbf{P}^{\mathbf{T}} + \mathbf{E} \tag{10.6}$$

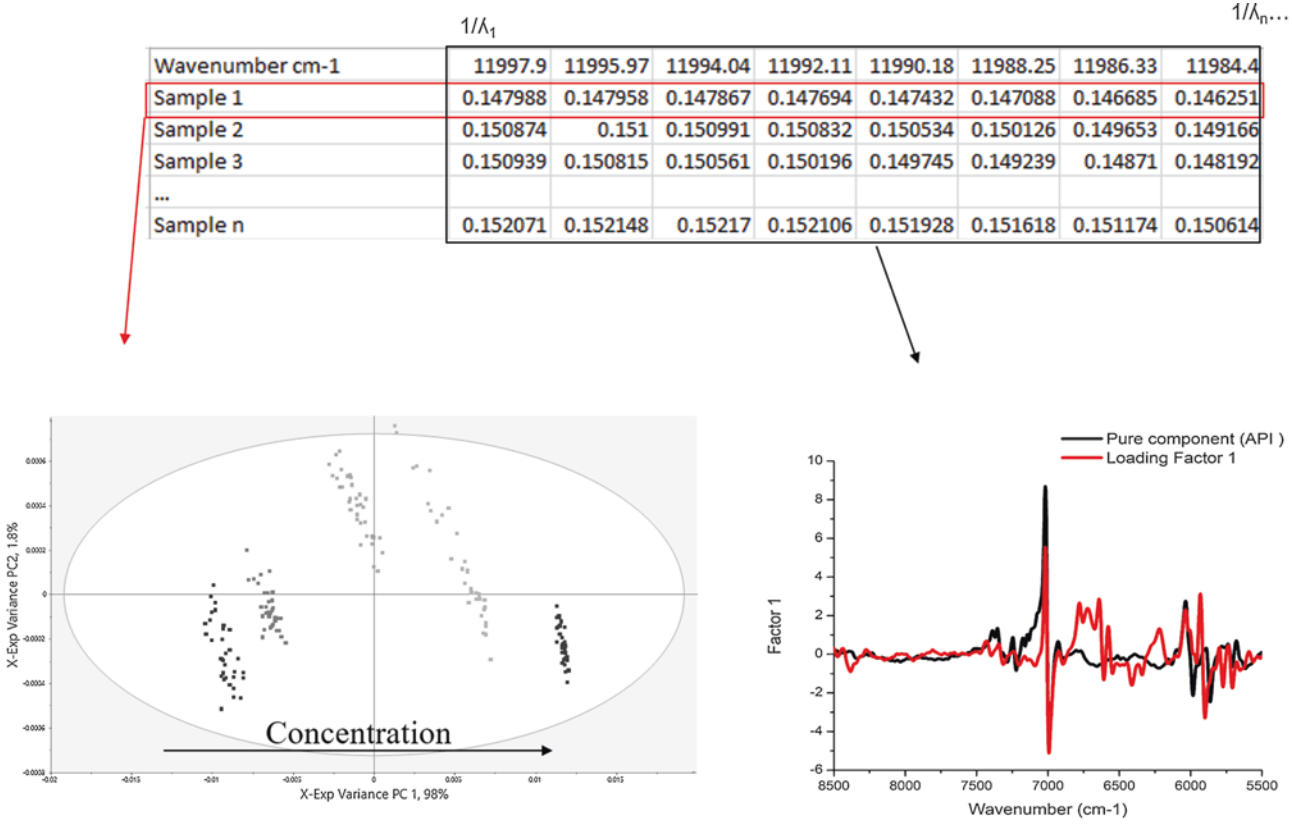


Fig. 10.6 Principal component analysis describing the transformation of data from NIR spectra

The new calculated axes called principal components (PC) are orthogonal to each other, meaning that they are completely uncorrelated but aligned with the principal sources of variation. The first component explains the maximum variance, and as the number of principal components increases, a higher percentage of the total variance is explained (Jackson 2004).

PCA provides a new coordinate system where the score plot can be considered a map of samples and the loading plot a map of the variables as shown in Fig. 10.6. The scores and loadings contain useful information that describes the properties of products in a process (Esbensen 2012; Kumar et al. 2014). Figure 10.6 shows a score plot in which the samples are clustered in the first PC (principal component) according to its concentration. Also the loadings plot show the characteristic peaks of the API and confirm that the most relevant information explained by the first PC is related to the API of interest.

PCA is often used for exploratory data analysis in continuous and batch manufacturing to gain insight into a data set and obtain a preliminary assessment of the variables that may help to monitor a system (Esbensen and Geladi 2009). PCA is also often used in classification and in understanding the changes that occur in a reaction or system (Fontalvo-Gomez et al. 2013). Even though PCA is a qualitative method, the extracted information is essential for quantitative regression methods (Saerens et al. 2012). PCA may be used to determine whether the NIR or Raman spectra obtained are similar to the calibration model spectra. This assessment could indicate that the spectra obtained are very different from those expected. In this case the reason for these unexpected or outlier spectra should be investigated (Miller 2010).

Several studies showed the suitability of PCA for process monitoring. The information obtained with PCA helps to identify, classify, and follow material changes – physical and chemical – due to process-induced changes during manufacturing (Vanarase et al. 2010; Colón et al. 2014; Martinez et al. 2013; Alcalà et al. 2010, 2012; Saerens et al. 2012; Rantanen et al. 2005). PCA has been used to determine process attributes such as force variation during roller compaction, tablet compression, and shear strain during blending (Acevedo et al. 2012; Roman-Ospino et al. 2016). PCA score plots have also pinpointed powder density changes that occur at different paddle wheel speeds (Mateo-Ortiz et al. 2014). Continuous mixing has been monitored by PCA. Figure 10.7 shows the start-up phase, steady state, and emptying of continuous mixers from the continuous manufacturing facility at Rutgers (Martinez et al. 2013).

Several statistics are derived after PCA calculation, which are useful for process monitoring and quality evaluation through manufacturing. Regulatory agencies indicate that a statistical spectral quality test should be performed before the model is applied, to assure that the spectra of a sample fall within a predefined range of variation (European Medicine Agency E 2014; Miller 2010). Q residual and Hotelling's T^2 establish critical limits based on the PCA model at a confidence interval that defines the acceptable range of variation for a given process and product. While Q statistics refers to the amount of sample variation not captured in the model, Hotelling's T^2 refers to the variation with respect to the center of the cluster defined by samples of the desired materials with the expected physical (Rantanen et al. 2005) and chemical properties as shown in Fig. 10.8 (Miller 2010). The outlier

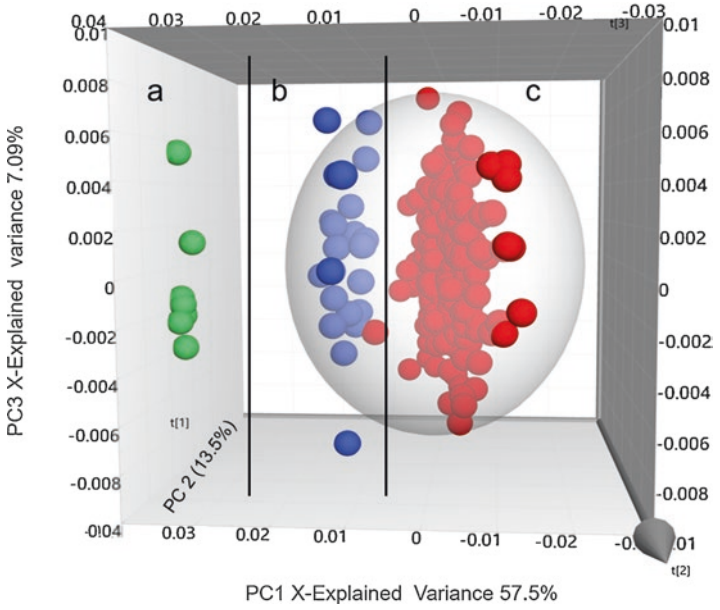


Fig. 10.7 PCA of NIR spectra obtained at Rutgers continuous manufacturing facility in three process stages (**a**) empty chute (green), (**b**) filled chute (blue), and (**c**) steady state (red)

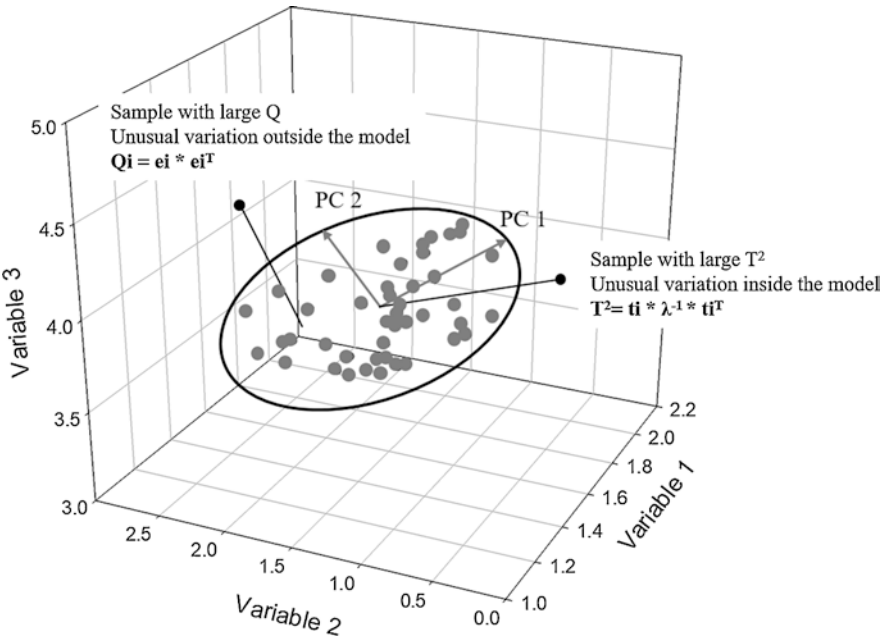


Fig. 10.8 Geometry of Q residuals and T^2 Hotelling

diagnostics have been implemented in commercial continuous manufacturing (Vargas et al. 2018).

PCA scores may be regressed against drug concentration values or other response variables. This may be done in a two-step approach, called principal component regression (PCR) that first calculates the scores and then performs a multilinear regression. This approach was followed recently to develop a nondestructive NIR calibration model to predict the dissolution profiles of tablets prepared by continuous direct compression (Pawar et al. 2016). However, calibration models are more frequently developed with partial least squares (PLS) regression.

10.6 Partial Least Squares (PLS) Regression

PLS regression is considered one of the most important statistical methods in pharmaceutical continuous manufacturing for developing quantitative calibration models (Vargas et al. 2017, 2018; Vanarase et al. 2010; Martinez et al. 2013; Zhang et al. 2014; Saerens et al. 2012, 2014). Calibration is defined as the development of a mathematical relationship between the changes in the NIR or Raman spectroscopic data matrix **X** and the property of interest, matrix **Y** for calibration and data compression, usually drug concentration (Beebe et al. 1998). PLS regression is appropriate to analyze collinear-noisy multivariate data and to model simultaneously the spectroscopic responses and the property of interest (Kramer 1998). PLS is designed to maximize the covariance between the **X** and **Y** data. The decomposition of both matrices **X** and **Y** takes place simultaneously as shown in Fig. 10.9 (Kumar et al. 2014).

Where **T** and **U** represent the scores, **P`** and **Q`** represent the loadings and **E** and F the residuals of the **X** and **Y** matrices, respectively.

PLS finds linear combinations (components or factors) that seek the maximum covariance between the analytical signal (near infrared or Raman spectra) and the property of interest, represented by **X** and **Y,** respectively. PLS calibration models

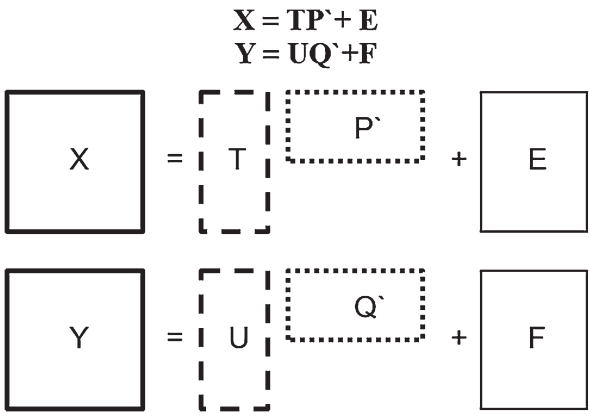


Fig. 10.9 Scheme of matrix decomposition in PLS. (Geladi and Kowalski 1986)

can be developed to determine changes in chemical composition and also to evaluate the physical properties of the material analyzed. These linear combinations are subsequently used in the regression equation for the prediction of drug concentration or other properties of unknown samples. The predictive ability of the calibration model depends on how well the calibration set encompasses the variation observed in the unknown samples (Blanco et al. 2008; Càrdenas et al. 2014).

The development of a calibration model requires planning to prepare the calibration set samples necessary for **X** and **Y** (Romañach et al. 2016). The calibration set samples are the key to the robustness and long-term success of the calibration model. The calibration samples should contain all expected components, span the concentration ranges and conditions of interest, and be as similar as possible to the unknown samples that it will predict (Kramer 1998). Thus, the development of a PLS calibration model requires a projection to the future, a consideration of how future sample will be. Various approaches have been followed to obtain this future variation (Càrdenas et al. 2014), including experimental designs which will be discussed below (Romañach et al. 2016).

The development of a calibration set can be a complex process considering the wide overlapping bands of NIR spectra. The overlapping bands from each component in the formulation make the extraction of the information related to the analyte of interest in the multicomponent system more difficult. The overlapping bands are due to the broad overtones and combination bands presented in NIR spectra. A well-designed calibration set is critical to ensuring accurate predictions of future unknown samples.

Bondi et al. investigated the effect of DoE on prediction performance of PLS models based on NIR spectra of an acetaminophen formulation with four excipients (Bondi Jr. et al. 2012). A two-factor, five-level, full factorial design (5-L FF) was generated where API content and the ratio of two excipients were the factors and the two remaining excipients were kept constant. Several other designs were investigated including a three-level full factorial (3-L FF), a central composite design (CCD), a D-optimal design, and an I-optimal design. I-optimality refers to the integrated prediction variance of the regression model, and the integrated-variance optimality criterion minimizes the prediction variance over the specified region of interest. A total of 170 tablet samples were produced, where 125 were destined to the calibration set and 45 to the validation set. The results based on SEP and bias comparisons of PLS models showed that 5-L FF and I-optimal design provided models with similar statistics and presented lowest RMSEP. This evaluation showed that the performance of a PLS model depends on the DoE used to generate the calibration set. The I-optimal design was superior to all other models, taking into account the performance and efficiency. Compared to 5-L FF design, I-optimal design used 16 fewer design points (80 samples), which could be very useful in drug development stages where the amount of API is limited.

The calibration model is often developed in a laboratory where it is easier to obtain the variation in calibration samples than in the pharmaceutical production area. However, the true objective of the calibration model is not to work in the laboratory but in the pharmaceutical manufacturing site. Calibration models have been

developed by obtaining NIR spectra for calibration blends for powder mixtures flowing in a chute (Vanarase et al. 2010) or over a conveyor belt simulating the conditions obtained after mixing (Colón et al. 2014). Calibration models were also developed using the reference values from the gravimetric feeders of the continuous manufacturing system (Vargas et al. 2018). Once the calibration model is deployed to the production area, it will be thoroughly challenged by the prediction of independent (test set) samples not used in building the calibration model (Esbensen and Geladi 2010). This is the true validation of the calibration model, since it is being challenged by the prediction of test set samples under the conditions of use for which it was designed.

10.7 Performance Evaluation of Calibration Models: Figures of Merit

The prediction of the test set samples is usually evaluated in terms of the root mean square error of prediction (RMSEP) defined as (Naes et al. 2002):

$$\text{RMSEP} = \sqrt{\sum_{N_p}^{i=1} \left(\hat{y}_i - y_i\right)^2 / N_p} \tag{10.7}$$

The RMSEP is calculated with the predicted value $\hat{y}_i$ and y_i as a reference value from a gravimetric or HPLC method and N_p – the number independent (test set) samples analyzed. In the calculation of the RMSEP, the predicted values are compared versus the reference value, while in the calculation of the standard deviation, the predicted values are compared versus the mean of the predicted values as shown below:

$$\text{STD Dev} = \sqrt{\frac{\sum_{i=1}^{n} \left(\hat{y}_i - \hat{y}_{\text{mean}}\right)^2}{\left(N_p - 1\right)}} \tag{10.8}$$

The bias is also calculated and provides a measurement of the accuracy of the method:

$$\text{Bias} = \sum_{N_p}^{i=1} \left(\hat{y}_i - y_i\right) / N_p \tag{10.9}$$

The standard error of prediction (SEP) is also used and defined as the standard deviation of the predictive residuals (Naes et al. 2002):

$$\text{SEP} = \sqrt{\sum_{N_p}^{i=1} \frac{\left(\hat{y}_i - y_i - \text{BIAS}\right)^2}{\left(N_p - 1\right)}} \tag{10.10}$$

These statistical terms provide the basis for a relationship that is important to understand the measurement error:

$$\mathrm{RMSEP}^2 \approx \mathrm{SEP}^2 + \mathrm{BIAS}^2 \tag{10.11}$$

Thus, the RMSEP encompasses the systematic error associated with the measurement (accuracy) and also the random error (precision).

In some cases it may be useful to also calculate the relative standard error of prediction:

$$\mathrm{RSEP}(\%) = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i\right)^2}} \times 100 \tag{10.12}$$

All evaluations of a PAT method used for continuous manufacturing should include what has been termed the essential statistical information for describing a data set (Massart 1997). The statistical information should first include the number of observations or predictions made. This requirement may appear to be common sense, but it is often overlooked. The essential statistical information also requires a parameter for the central tendency of the results, such as the arithmetic mean (or average). The dispersion of results (standard deviation) is needed to indicate method precision. This statistical information is also essential to determine the number of significant figures for reporting results (Cárdenas and Romañach 2018).

10.8 Statistical Methods for Process Monitoring

PLS and PCA have been used to extract information from NIR spectra obtained in real-time monitoring of continuous mixing (Osorio et al. 2015). Vanarase et al. developed a system for real-time quantification of drug in a continuous powder mixing process using NIR spectroscopy. The NIR spectrometer was placed at the exit of the blender obtaining spectra after the blends were emptied to a metal chute. Statistical methods such as PCA were used as exploratory and analysis tool and PLS for the calibration model to determine the drug concentration in the blends. The performance of the proposed NIR method showed its suitability for monitoring continuous mixing (Vanarase et al. 2010).

Martinez et al. also developed a PLS method to quantify the drug concentration after continuous mixing. NIR spectra were used to monitor process start-up, steady state, and emptying phases. The characterization of these process phases is important to determine the optimum operational state and provide adequate mixing. Therefore, several statistical methods were developed for determining the end point of the start-up phase. PCA and statistical calculations were used to establish steady state and PLS for drug concentration (Martinez et al. 2013).

Colón et al. also described a methodology to monitor a continuous mixing process using NIR spectroscopy in which the spectra were taken at the exit of a custom-made tumble mixer while the mixed powder moved over a conveyor belt. This study thoroughly evaluated the sources of error in the drug concentrations predicted by the NIR calibration model (Colón et al. 2014). The variation in the drug concentrations was evaluated by fast Fourier transform analysis, which revealed a cyclic behavior in the drug concentration values. This cyclic behavior was traced to the cyclic behavior of the screw feeders and mixer used in this continuous manufacturing system. Thus, the statistical methods provided information on the performance of the continuous manufacturing hardware and also the critical process parameter (drug concentration).

PCA and PLS may also be used to obtain information from non-spectroscopic responses (Rogers and Ierapetritou 2015, 2016). PLS has been used to extract additional information from a series of simulations performed to further understand continuous powder blending performance (Gao et al. 2013). These simulations included 8 different responses related to 64 simulation designs that covered the operation and design space of a continuous blender. The variations calculated in these simulations were summarized by a PLS model. The results showed that the shaft angle and blade speed had the greatest influence in the blender's performance.

Critical process parameters (CPPs) such as mixing time, shear stress, compaction pressure, and coating characteristics need to be monitored during manufacturing, since these affect substantially the quality of the product. Several studies have shown the use of statistical methods – described above – for monitoring of CPPs and the effectiveness of such methods for the purpose.

PLS has also been used to model the effect of shear stress. Shear stress is a force applied on a section of the powder material in which one material surface is moving over another that remains static; this force is part of the mixing dynamics of many particulate industries (Shinbrot and Muzzio 2013). In the pharmaceutical industry, shear is important for blending cohesive materials and breaking down agglomerates during solid dosage manufacture. In continuous mixing, shear strain is developed within the process when powder particles are constantly moving and can affect the quality attributes of the final product.

Hernández et al. developed a NIR method to predict dissolution of tablets and spectral changes associated with shear strain. A series of experiments were performed in which blends of a pharmaceutical formulation were subjected to different levels of shear strain during mixing. This was followed by tableting through a compaction step at a single pressure of 350 MPa. Once the tablets were obtained, their dissolution profile was analyzed using a USP Apparatus 2 as a reference method. This previous analysis showed changes in the dissolution profile of tablets in which the blends used to prepare the tablets were subjected to four strain levels (Hernandez et al. 2016a). PCA was used to visualize the variation in strain levels between samples and PLS-2 to predict dissolution profile and drug content using a unique simple calibration model. PLS-2 is a least squares method that allows the determination of multiple properties with a single calibration model, correlating the independent variables **X** (spectra) and multiple dependable **Y** variables relating the property of interest. This approach is faster than performing several PLS-1 calculations.

The study confirmed the suitability of the model for the purpose showing low prediction errors (RSEP <7%) and bias close to 0 (0.5–0.7%). The results obtained with the calculation of the similarity factor show similarity between the developed NIR method and the reference method. Spectral changes caused by shear strain were also investigated using NIR chemical imaging and angle-resolved elastic light scattering experiments. As shear reduces the particle size of the material by de-agglomeration, a reduction of the scattering intensity is observed in the NIR spectra (Hernandez et al. 2016b). These studies showed (Hernandez et al. 2016a, b) that multivariate statistical methods are a significant asset in extracting information from both batch and continuous systems. The subtle NIR spectral changes provide useful information when multiple statistical methods are applied.

10.9 Variographic Analysis

The validation of the PLS NIR calibration model is often delayed due to fundamental sampling issues. Pharmaceutical scientists will invariably want to compare the NIR results to the gold standard high-performance liquid chromatography (HPLC) or Karl Fischer (KF) moisture results. The first fundamental issue is that the sample measured by the NIR system is often not the same as that measured in the laboratory (Green et al. 2005). The systematic error that ensues from this situation is usually the principal source of error in NIR measurements (Mark 1991). The mass analyzed by the NIR method may be much smaller than the mass analyzed by the reference method in the laboratory. The sample measured by the reference method may be a composite of a greater number of increments from the lot than the sample measured by the NIR or Raman system, resulting in a systematic error. Sampling issues are receiving increasing attention and becoming an important research area in PAT and pharmaceutical manufacturing (Esbensen and Paasch-Mortensen 2010; Romañach and Esbensen 2015; Romañach 2017).

Variographic analysis has been recently introduced as a means to evaluate the mixing of pharmaceutical powder mixtures with the advantage of providing an estimate of sampling and analytical errors (Esbensen et al. 2016; Romañach et al. 2018). The calculation of a variogram requires knowing the order in which samples were collected and analyzed. Thus, the real-time monitoring of drug concentration in a continuous mixing system is ideal for variographic analysis (Romañach et al. 2018). Drug concentration may be determined by NIR spectroscopy as the powder leaves the mixer and flows down a chute or moves over a conveyor belt (Vargas et al. 2017). In this monitoring of the continuous system, the order in which the spectra were obtained is known making it possible to perform variographic analysis. Variograms are calculated by plotting the sample-pair variances **V(j)** versus the distance (lag) between the extracted samples (Romañach et al. 2018). The variogram function **V(j)** is calculated according to:

$$\mathbf{V(j)} = \frac{\mathbf{1}}{\mathbf{2\left(Q_{total} - j\right)}} \sum_{\mathbf{Q_{total}-j}}^{\mathbf{q=1}} \left(\mathbf{h_{q+j} - h_q}\right)^{\mathbf{2}} \tag{10.13}$$

where $\mathbf{h_q}$ is the drug concentration determined by the NIR method, **j** is the lag (which refers to the distance in time between two samples), and $\mathbf{Q_{total}}$ is the number of analyses for extracted samples (from a lot). A lag of 1 measures the correlation between consecutive NIR predictions, and a lag of 2 measures the correlation between every other NIR prediction. Low **V(j)** values indicate a reduced heterogeneity in the blend and good mixing. The low **V(j)** values are usually observed for samples obtained at contiguous locations or obtained within close intervals of time, when the lag values are low. Higher **V(j)** values are often observed when the lag time is increased and represent the random variation of the process.

Variographic analysis permits the measurement of three key parameters: the nugget effect, the sill, and the range. These terms originated in the mining industry. The nugget effect provides the minimum practical error (MPE) – equivalent to the sum of the sampling and analytical method errors with the sampling method used at the time. The nugget effect is calculated by estimating **V(j)** for a lag value of 0. The measurement of error (**V(j)**) for samples that are spaced 0 lags apart is impossible, and thus it is necessary to back extrapolate this value based on the **V(j)** values for lag = 1– 5 values. The MPE represents the error estimated when there is no difference between the samples (lag = 0). The MPE is a valuable contribution from variographic analysis since the accuracy, precision, range, and linearity of analytical methods are thoroughly studied in method validation, but very few studies have investigated sampling error (Green et al. 2005; Esbensen et al. 2016; Romañach et al. 2018).

Figure 10.10 shows the sill variance level which provides a measurement of the maximum variation (heterogeneity) obtained in the powder mixing process. The sill encompasses the entire process, including the process variation, sampling, and analytical errors. The difference between the sill and the MPE provides an estimate of the process variance. The top variogram in Fig. 10.10 shows a case where the sill is high and the MPE is only slightly lower. In this situation the process (mixing) could be improved only to the level of the MPE. Further improvements of the sill would require a reduction in sampling and analytical errors (a lower MPE). The bottom variograms in Fig. 10.10 show a process where the sill is much lower, indicating less heterogeneity and much better mixing. This mixing process could be further improved to the level of the MPE. Thus, the variograms are also useful as a tool for process improvement providing a means to discern between analytical and sampling errors and the process variance.

The range in the variograms also provides useful information. The range of the variograms is defined as the lag at which there is no autocorrelation. The lag values below the range provide lower **V(j)** values indicating that there is a greater degree of similarity (autocorrelation) in the concentrations that were determined. Thus, the range corresponds to the beginning of the maximum variance (sill). The range can be used to evaluate the sampling rate and to design a composite sampling plan for the process. The range, sill, and MPE are terms that are definitely unfamiliar in the pharmaceutical field but part of variographic analysis – a very promising approach for the evaluation of pharmaceutical processes.

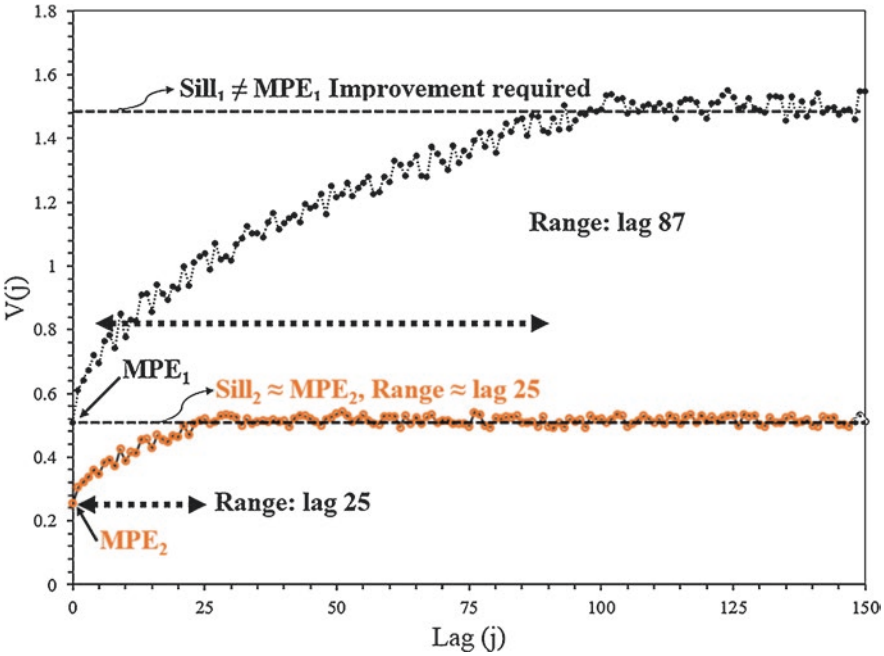


Fig. 10.10 Variogram indicating the information that may be obtained from a continuous manufacturing process

Variographic analysis has been used in a manufacturing setup as part of real-time analysis. Vargas et al. described the use of variograms to estimate sampling and analytical errors. This study represents the first variographic analysis performed for a continuous manufacturing system installed at a commercial pharmaceutical manufacturing facility (Vargas et al. 2017). A subsequent study included variographic analysis for a 28-hour continuous manufacturing production of a commercial product, the largest variographic analysis reported within the pharmaceutical field (Vargas et al. 2018).

10.10 Quality by Design and Multivariate Statistical Process Control: From Batch Manufacturing Process Toward Continuous Manufacturing

Statistical methods are an essential part of a significant number of studies conducted following the quality by design framework. The development of a comprehensive design space shows that the manufacturing process has been well described and critical sources of risk to quality have been mitigated (Bondi and Drennen 2011). ICH Q8 provides the framework for implementation of quality by design (QbD) in

pharmaceutical development and emphasizes the different analytical tools and applications that can be used to achieve QbD. Thus, the QbD initiative also includes "ICH Q9, Quality Risk management" and "ICH Q10 Pharmaceutical quality system (Q9: Quality Risk Management 2006; Q10: Pharmaceutical Quality System 2009). Quality by design (QbD) approach promotes the key idea that "quality cannot be tested or inspected into a finished product, but rather that quality, safety and effectiveness must be designed and built into a product and its manufacturing process" (Yu 2008; Yu et al. 2014).

Tabasi et al. addressed a complete study to better understand a tablet manufacturing process by using various statistical methods (Tabasi et al. 2008a, b, c). In the study NIR spectroscopy was used to predict tablet compression force, crushing strength, and content uniformity in core tablets (Tabasi et al. 2008a). NIR spectroscopy was also used to predict coating thickness and drug release from coating process and to predict and monitor the extent of film coat curing (Tabasi et al. 2008b, c). Calibration models were developed using PLS regression. One of the key points was the use of PCA, ensuring that calibration samples encompassed the spectral variability from chemical and physical changes. The PCA algorithm was able to differentiate between capped and normal tablets. Spectra in absorbance mode of tablets revealed that higher compression force is correlated to a higher absorbance. In addition, when the compression force increases the occurrence of a capping-like breakage pattern increased as well. Tablets made individually in small batch and tablets manufactured in lab-scale production were used in the study. Finally, PLS model was validated according to the guidelines evaluating parameters as specificity, linearity, accuracy, precision, and robustness.

Additional PLS calibrations based on NIR spectra were developed to predict coating thickness and determine the end point in coating process (Tabasi et al. 2008b) as well as to predict the dissolution profile and drug release. Finally, NIR spectroscopy was used to assess film coat curing for coated tablets, especially its ability to monitor film coat curing in comparison to conventional methods such as differential scanning calorimetry (DSC) and hot-stage microscopy. Results revealed that variation in the curing temperature and duration affected the NIR spectra. In addition, the effect of curing duration and temperature on cast films, uncoated tablets, coated tablets, and coated glass beads was investigated by using PCA, in particular scatter plots of scores showed that curing duration and temperature affected coated glass beads, uncoated tablets, and coated tablets significantly.

The same QbD principles and methodologies may be incorporated in the continuous manufacturing framework currently being developed. The QbD approach and the associated design space concept bring the opportunity to progress from the traditional approach of changing a fixed formulation to a more flexible operating region which allows to continuously improve the process (Reklaitis et al. 2010).

QbD is also facilitated through the use of multivariate statistical process control (MSPC) for process monitoring (Nomikos and MacGregor 1995). MPSC consists of monitoring temporal changes in properties of a batch being processed with others obtained under normal operating conditions (NOC). The batches can be compared by plotting their scores over process time. The line connecting different scores is

considered as the batch trajectory, and based in NOC batches, specific control limits can be established (Rosas et al. 2011a). The data for a batch can be considered as a three-way array, a data cube where one dimension corresponds to process variables or wavelengths in relation to spectra, another to the samples withdrawn during the process or spectra collected, and the third dimension to the different batches.

Nomikos and MacGregor described an unfolding procedure for the data cube that allows to compare batches being manufactured to NOC batches (Nomikos and MacGregor 1995). This approach, however, requires that batches be completely finished for comparison; also, all batches should have an identical process duration, which is quite difficult to achieve in production. This procedure was modified by Wold et al., where the data cube with process variables "X" is unfolded into a two-way data array from which a response vector "Y" can be built (Wold et al. 1998). In batch process control, Y can be a property of the product or a timestamp value. PCA or PLS scores and loadings are monitored via control charts such as Hotelling's T^2 and Q residuals (Rosas et al. 2011a).

In one of the first applications of PAT technologies to pharmaceutical process, Cogdill et al. described the development and implementation of quantitative NIR spectroscopy and MSPC method as part of a RTR strategy for immediate-release tablet (Cogdill et al. 2005). Methods were based on indicators of high-flux noise (noise factor level) and wavelength uncertainty. These measurements, in combination with Hotelling's T^2 and Q residual, were used to continuously monitor instrument performance and model relevance.

Skibsted et al. (2007) used NIR spectroscopy and MSPC to implement a RTR strategy applied to a typical tablet manufacturing process containing mixing, granulation, drying, and tableting. Four models were generated, one MSPC model and three regression models to predict quality indicators. The MSPC model utilized NIR spectra collected subsequent to granulation to detect particle size deviation that may potentially exacerbate future processing. The study demonstrated that water content could provide real-time control of the drying time.

Rosas et al. (2011a, b) applied the QbD approach to identify the design space of a pharmaceutical gel process and proposed multivariate PLS models by using NIR spectroscopy to monitor the process in the predefined design space. This work combined the use of classical statistical (i.e., DoE, response surface analysis) and multivariate (i.e., PLS, MSPC) approaches to understand the production process and identify the design space where PAT tool can be applied to monitor the process. PLS models were developed based in NIR spectra for determining the composition of the pharmaceutical gel and assess the temporal changes in major physical factors, which affect the quality of product. Applied multivariate statistical process control (MSPC) was used to compare the batches included in the experimental design with batches manufactured under normal operating conditions (NOC).

The data cube X was unfolded into a two-way data array (Fig. 10.11) from which a vector Y can be constructed. In batch process control, Y can be a property of the product or a time value in the sample collection period. PCA or PLS scores and loadings are monitored via control charts (D-statistic and Q-statistic) and assumed to represent the variables of a standard process.

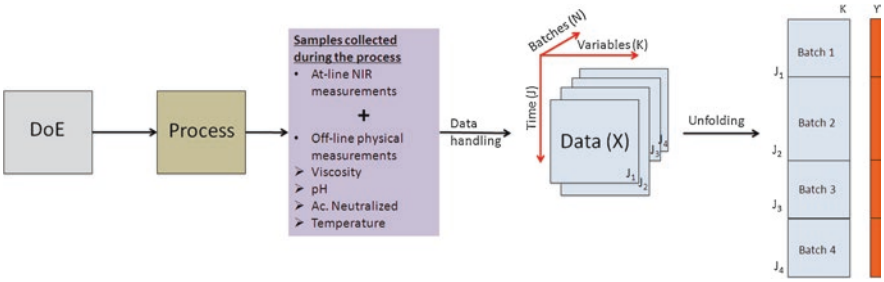


Fig. 10.11 Multivariate unfolding of process data prior to the construction of MSPC models. (Rosas et al. 2011b)

10.11 Real-Time Release Testing

This chapter has presented a number of statistical methods that contribute to quality by design and PAT in both batch and continuous processes and which may be used to enable real-time quality control and eventually real-time release testing. Real-time release starts with the design and development of a manufacturing process, where the design of experiments methodology is essential. Real-time release also requires process monitoring and control as defined within PAT. There are now multiple studies that address drug concentration in blends and real-time moisture measurements. The available tools continue to increase with the recent development of methods for monitoring of powder density by both NIR and microwave dielectric spectroscopy providing a means for possible control of powder density. PAT usually relies on multivariate methods to extract and interpret the information from process measurements. This chapter has provided clear evidence that there are now multiple methods for extracting information from PAT measurements in pharmaceutical processes. However, end product testing is also important, especially dissolution testing.

Two methods have been published recently describing the prediction of dissolution profiles of tablets through nondestructive NIR spectroscopy. There are now many possibilities available for companies interested in enabling real-time release testing.

Acknowledgments The authors thank Dr. Andres Roman Ospino and Dr. James Scicolone for their help with figures in this chapter. This work was funded by the National Science Foundation Engineering Research Center on Structured Organic Particulate Systems, through Grant NSF-ECC 0540855, and the Puerto Rico Science, Technology and Research Trust.

References

Acevedo D, Muliadi A, Giridhar A, Litster JD, Romanach RJ. Evaluation of three approaches for real-time monitoring of roller compaction with near-infrared spectroscopy. AAPS PharmSciTech. 2012;13(3):1005–12.

Agaian S. Hadamard matrices and their applications. Heidelberg: Springer; 1985.

Aksu B, Beer TD, Folestad S, Ketolainen J, Lindén H, Lopes JA, et al. Strategic funding priorities in the pharmaceutical sciences allied to Quality by Design (QbD) and Process Analytical Technology (PAT). Eur J Pharm Sci. 2012;47(2):402–5.

Alcala M, Blanco M, Bautista M, Gonzalez JM. On-line monitoring of A granulation process by NIR spectroscopy. J Pharm Sci. 2010;99(1):336–45.

Alcalà M, Blanco M, Menezes JC, Felizardo PM, Garrido A, Pérez D, et al. Near-infrared spectroscopy in laboratory and process analysis. Encyclopedia of analytical chemistry: Wiley; Hoboken, New Jersey, USA; 2012.

Barajas MJ, Cassiani AR, Vargas W, Conce C, Ropero J, Figueroa J, et al. Near-infrared spectroscopic method for real-time monitoring of pharmaceutical powders during voiding. Appl Spectrosc. 2007;61(5):490–6.

Beach L, Ropero J, Mujumdar A, Alcalà M, Romañach RJ, Davé RN. Near-infrared spectroscopy for the in-line characterization of powder voiding part II: quantification of enhanced flow properties of surface modified active pharmaceutical ingredients. J Pharm Innov. 2010;5(1–2):1–13.

Beebe KR, Pell RJ, Seasholtz MB. Chemometrics: a practical guide: John Wiley & Sons, Hoboken, New Jersey, USA; 1998. 360 p.

Berntsson O, Danielsson LG, Folestad S. Estimation of effective sample size when analysing powders with diffuse reflectance near-infrared spectrometry. Anal Chim Acta. 1998;364:243–51.

Blanco M, Bautista M, Alcala M. Preparing calibration sets for use in pharmaceutical analysis by NIR spectroscopy. J Pharm Sci. 2008;97(3):1236–45.

Bondi RW, Drennen JK. Quality by design and the importance of PAT in QbD. Sep Sci Technol. 2011;10:195–224.

Bondi RW Jr, Igne B, Drennen JK 3rd, Anderson CA. Effect of experimental design on the prediction performance of calibration models based on near-infrared spectroscopy for pharmaceutical applications. Appl Spectrosc. 2012;66(12):1442–53.

Boodoosingh GL, Lopez GE. An efficient algorithm in the grand canonical ensemble: constructing adsorption isotherms. Mol Simul. 2002;28(3):273–85.

Box G, Draper N. Practical choice of a response design. Response surfaces, mixtures, and ridge analysis: John Wiley & Sons, Hoboken, New Jerse; 2006. p. 483–508.

Box G, Hunter J, Hunter W. Statistics for experimenters: design, innovation, and discovery. 2nd ed. New Jersey: Wiley-Interscience; 2005.

Callis J, Illman D, Kowalski B. Process analytical chemistry. Anal Chem. 1987;59(9):624A–37A.

Cárdenas V, Romañach RJ. Determining the number of significant figures for reporting NIR results. NIR news. 2018;29(4):15–17.

Càrdenas V, Blanco M, Alcalà M. Strategies for selecting the calibration set in pharmaceutical near infrared spectroscopy analysis. A Comparative Study. J Pharm Innov. 2014;9(4):272–81.

Cogdill RP, Delgado-Lopez M, Molseed D, Chisholm R, Bolton R, Herkert T, Afnán A, Drennen J. Process analytical technology case study part I: feasibility studies for quantitative near-infrared method development. AAPS PharmSciTech. 2005;6(2):E363–272.

Colón YM, Florian MA, Acevedo D, Méndez R, Romañach RJ. Near infrared method development for a continuous manufacturing blending process. J Pharm Innov. 2014;9(4):291–301.

Dahm D, Dahm K. Interpreting diffuse reflectance and transmittance: a theoretical introduction to absorption spectroscopy of scattering materials: NIR Publications; Chichester, West Sussex, United Kingdom; 2007.

de Matas M, De Beer T, Folestad S, Ketolainen J, Lindén H, Lopes JA, et al. Strategic framework for education and training in Quality by Design (QbD) and process analytical technology (PAT). Eur J Pharm Sci. 2016;90:2–7.

Deng X, Scicolone J, Han X, Davé RN. Discrete element method simulation of a conical screen mill: a continuous dry coating device. Chem Eng Sci. 2015;125:58–74.

Doehlert D. Uniform shell design. Applied Stat. 1970;19:231–9.

El Hagrasy A, Chang SY, Desai D, Kiang S. Raman spectroscopy for the determination of coating uniformity of tablets- assessment of product quality and coating pan mixing efficiency during scale-up. J Pharm Innov. 2006;1(1):37–42.

Engisch W, Muzzio F. Using Residence Time Distributions (RTDs) to address the traceability of raw materials in continuous pharmaceutical manufacturing. J Pharm Innov. 2016;11:64–81.

Eriksson LJE, Kettaneh-Wold N, Wikstrom C, Wold S. Design of experiments-Principles and applications. 3nd ed. Sweden: Umea; 2008.
Esbensen K, Geladi P. Principal component analysis: concept, geometrical interpretation, mathematical background, algorithm, history, practice. In: SDBT W, editor. Comprehensive chemometrics. Oxford: Elsevier; 2009. p. 211–26.
Esbensen KH, Geladi P. Principles of proper validation: use and abuse of re-sampling for validation. J Chemom. 2010;24(3–4):168–87.
Esbensen KH, Paasch-Mortensen P. Process Sampling: Theory of Sampling – the Missing Link in Process Analytical Technologies (PAT). In: Bakeev K, editor. Process Analytical Technology. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd; 2010. p. 37–80.
Esbensen KE, Swarbrick B. Multivariate Data Analysis – in practice. An Introduction Multivariate Analysis, Process Analytical Technology and Quality by Design. 6th ed. Oslo, Norway: CAMO Software AS; 2018. 480 p.
Esbensen KH, Roman-Ospino AD, Sanchez A, Romanach RJ. Adequacy and verifiability of pharmaceutical mixtures and dose units by variographic analysis (Theory of Sampling) – a call for a regulatory paradigm shift. Int J Pharm. 2016;499(1–2):156–74.
European Medicine Agency E. Guideline on the use of near infrared spectroscopyby the pharmaceutical industry and the data requirements for new submissions and variations; 2014.
Fontalvo-Gomez M, Colucci JA, Velez N, Romanach RJ. In-line near-infrared (NIR) and Raman spectroscopy coupled with principal component analysis (PCA) for in situ evaluation of the transesterification reaction. Appl Spectrosc. 2013;67(10):1142–9.
Fonteyne M, Vercruysse J, De Leersnyder F, Van Snick B, Vervaet C, Remon JP, et al. Process analytical technology for continuous manufacturing of solid-dosage forms. TrAC Trends Anal Chem. 2015;67:159–66.
Gao Y, Boukouvala F, Engisch W, Meng W, Muzzio FJ, Ierapetritou MG. Improving continuous powder blending performance using projection to latent structures regression. J Pharm Innov. 2013;8(2):99–110.
Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. Anal Chim Acta. 1986;185:1–17.
Goupy J, Creighton L. Introduction to design of experiments with JMP examples: SAS press, Houghton, Johannesburg; 2007.
Green RL, Thurau G, Pixley NC, Mateos A, Reed RA, JP H. In-line monitoring of moisture content in fluid bed dryers using near-IR spectroscopy with consideration of sampling effects on method accuracy. Anal Chem. 2005;77(14):4515–22.
Griffiths PR, de Haseth JA. Fourier transform infrared spectrometry. 2nd ed; 2007. 560 p.
Hernandez E, Pawar P, Keyvan G, Wang Y, Velez N, Callegari G, et al. Prediction of dissolution profiles by non-destructive near infrared spectroscopy in tablets subjected to different levels of strain. J Pharm Biomed Anal. 2016a;117:568–76.
Hernandez E, Pawar P, Rodriguez S, Lysenko S, Muzzio FJ, Romanach RJ. Effect of shear applied during a pharmaceutical process on near infrared spectra. Appl Spectrosc. 2016b;70(3):455–66.
Hu Y, Liang JK, Myerson AS, Taylor LS. Crystallization monitoring by Raman spectroscopy-simultaneous measurement of desupersaturation profile and polymorphic form in flufenamic acid systems. Ind Eng Chem Res. 2005;44(5):1233–40.
Huang Z, Scicolone JV, Han X, Dave RN. Improved blend and tablet properties of fine pharmaceutical powders via dry particle coating. Int J Pharm. 2015;478(2):447–55.
Hussain A. interviewed by Rodolfo Romañach [Personal communication] Greenville, NC; 2015.
Jackson J. A User's guide to principal components: Wiley Interscience, Hoboken, New Jersey, USA; 2004.
Joglekar GS, Giridhar A, Reklaitis G. A workflow modeling system for capturing data provenance. Comput Chem Eng. 2014;67:148–58.
Kramer R. Chemometric techniques for quantitative analysis: Taylor & Francis, New York, NY; 1998.
Kumar N, Bansal A, Sarma GS, Rawal RK. Chemometrics tools used in analytical chemistry: an overview. Talanta. 2014;123:186–99.
Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10(3):191–9.

Mark H. Principles and Practice of Spectroscopic Calibration. New York, NY: Wiley; 1991.
Markl D, Wahl PR, Menezes JC, Koller DM, Kavsek B, Francois K, et al. Supervisory control system for monitoring a pharmaceutical hot melt extrusion process. AAPS PharmSciTech. 2013;14(3):1034–44.
Martinez L, Peinado A, Liesum L, Betz G. Use of near-infrared spectroscopy to quantify drug content on a continuous blending process: influence of mass flow and rotation speed variations. Eur J Pharm Biopharm. 2013;84(3):606–15.
Massart DL. In: Massart DL, editor. Handbook of Chemometrics and Qualimetrics Part A; Elsevier Science, Amsterdam, The Netherlands; 1997. p. 26.
Mateo-Ortiz D, Colon Y, Romanach RJ, Mendez R. Analysis of powder phenomena inside a Fette 3090 feed frame using in-line NIR spectroscopy. J Pharm Biomed Anal. 2014;100:40–9.
Mercado J, Alcalà M, Karry KM, Ríos-Steiner JL, Romañach RJ. Design and in-line Raman spectroscopic monitoring of a protein batch crystallization process. J Pharm Innov. 2008;3(4):271–9.
Miller CE. Chemometrics for on-line spectroscopy applications—theory and practice. J Chemom. 2000;14(5–6):513–28.
Miller CE. Chemometrics in process analytical technology (PAT). In: Bakeev KA, editor. Process Analytical Technology. Second ed. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd; 2010. p. 353–438.
Moghtadernejad S, Escotet-Espinoza MS, Oka S, Singh R, Liu Z, Román-Ospino AD, et al. A training on: continuous manufacturing (direct compaction) of solid dose pharmaceutical products. J Pharm Innov. 2018;13(2):155–187.
Næs T, Isaksson T, Fearn T, Davies T. A User-Friendly Guide to Multivariate Calibration and Classification. Chichester, West Sussex: NIR Publications; 2002. 344 p.
Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. Technometrics. 1995;37(1):41–59.
Osorio JG, Vanarase AU, Romañach RJ, Muzzio FJ. Continuous powder mixing. Pharmaceutical blending and mixing: John Wiley & Sons, Chichester, West Sussex, United Kingdom; 2015. p. 101–27.
Pawar P, Wang Y, Keyvan G, Callegari G, Cuitino A, Muzzio F. Enabling real time release testing by NIR prediction of dissolution of tablets made by continuous direct compression (CDC). Int J Pharm. 2016;512(1):96–107.
Pell RJ, Seasholtz MB, Beebe KR, Koch MV. Process analytical chemistry and chemometrics, Bruce Kowalski's legacy at The Dow Chemical Company. J Chemom. 2014;28(5):321–31.
Petersen L, Esbensen KH. Representative process sampling for reliable data analysis—a tutorial. J Chemometr. 2005;19(11–12):625–47.
Pharmaceutical Quality System. Technical requirements for registration of pharmaceuticals for human use. 2009.
Quality Risk Management. Technical requirements for registration of pharmaceuticals for human use; 2006.
Rantanen J, Khinast J. The future of pharmaceutical manufacturing sciences. J Pharm Sci. 2015;104(11):3612–38.
Rantanen J, Wikström H, Turner R, Taylor LS. Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes. Anal Chem. 2005;77(2):556–63.
Reklaitis GV, Khinast J, Muzzio F. Pharmaceutical engineering science—new approaches to pharmaceutical development and manufacturing. Chem Eng Sci. 2010;65(21):iv–vii.
Rogers A, Ierapetritou M. Challenges and opportunities in modeling pharmaceutical manufacturing processes. Comput Chem Eng. 2015;81:32–9.
Rogers A, Ierapetritou MG. Mathematical tools for the quantitative definition of a design space. In: Ierapetritou GM, Ramachandran R, editors. Process simulation and data modeling in solid oral drug development and manufacture. New York, NY: Springer; 2016. p. 225–79.
Romañach RJ. Theory of sampling - from missing link to key enabler for process analytical technology (PAT). In: Dominy SC, Esbensen KH, editors. World conference on sampling and blending; May 9–11, 2017. 8th ed. Perth: Australian Institute of Mining and Metallurgy; 2017. p. 63–8.

Romañach R, Esbensen K. Sampling in pharmaceutical manufacturing—many opportunities to improve today's practice through the theory of sampling (TOS). TOS Forum. 2015;2015(4):5.

Romañach RJ, Hernández Torres E, Roman Ospino A, Pastrana I, Semidei F. NIR and Raman spectroscopic measurements to train the next generation of PAT scientists. Am Pharm Rev. 2014;17(6):82–7.

Romañach RJ, Román-Ospino AD, Alcalà M. A procedure for developing quantitative near infrared (NIR) methods for pharmaceutical products. In: Ierapetritou MG, Ramachandran R, editors. Process simulation and data modeling in solid oral drug development and manufacture. Methods in pharmacology and toxicology. New York: Springer; 2016. p. 133–58.

Romañach RJ, Sanchez-Paternina A, Esbensen KH. Variographic analysis of 1-D lots in pharmaceutical manufacturing (powder mixing). Am Pharm Rev. 2018;21(1):22–6.

Roman-Ospino AD, Singh R, Ierapetritou M, Ramachandran R, Mendez R, Ortega-Zuniga C, et al. Near infrared spectroscopic calibration models for real time monitoring of powder density. Int J Pharm. 2016;512(1):61–74.

Romero-Torres S, Perez-Ramos JD, Morris KR, Grant ER. Raman spectroscopic measurement of tablet-to-tablet coating variability. J Pharm Biomed Anal. 2005;38(2):270–4.

Romero-Torres S, Perez-Ramos JD, Morris KR, Grant ER. Raman spectroscopy for tablet coating thickness quantification and coating characterization in the presence of strong fluorescent interference. J Pharm Biomed Anal. 2006;41(3):811–9.

Ropero J, Beach L, Alcalà M, Rentas R, Davé RN, Romañach RJ. Near-infrared spectroscopy for the in-line characterization of powder voiding part I: development of the methodology. J Pharm Innov. 2009;4(4):187–97.

Rosas JG, Blanco M, Gonzalez JM, Alcala M. Quality by design approach of a pharmaceutical gel manufacturing process, part 2: near infrared monitoring of composition and physical parameters. J Pharm Sci. 2011a;100(10):4442–51.

Rosas JG, Blanco M, Gonzalez JM, Alcala M. Quality by design approach of a pharmaceutical gel manufacturing process, part 1: determination of the design space. J Pharm Sci. 2011b;100(10):4432–41.

Rozo JI, Zarow A, Zhou B, Pinal R, Iqbal Z, Romanach RJ. Complementary near-infrared and Raman chemical imaging of pharmaceutical thin films. J Pharm Sci. 2011;100(11):4888–95.

Saerens L, Dierickx L, Quinten T, Adriaensens P, Carleer R, Vervaet C, et al. In-line NIR spectroscopy for the understanding of polymer–drug interaction during pharmaceutical hot-melt extrusion. Eur J Pharm Biopharm. 2012;81(1):230–7.

Saerens L, Segher N, Vervaet C, Remon JP, De Beer T. Validation of an in-line Raman spectroscopic method for continuous active pharmaceutical ingredient quantification during pharmaceutical hot-melt extrusion. Anal Chim Acta. 2014;806:180–7.

Seasholtz MB. Making money with chemometrics. Chemometr Intell Lab Syst. 1999;45(1–2):55–63.

Shinbrot T, Muzzio FJ. Mixing and segregation in tumbling blenders. Encyclopedia of pharmaceutical science and technology. 4th ed: CRC Press, Raton, FL; 2013. p. 2208–21.

Sierra-Vega NO, Sánchez-Paternina A, Maldonado N, Cárdenas V, Romañach RJ, Méndez R. In line monitoring of the powder flow behavior and drug content in a Fette 3090 feed frame at different operating conditions using Near Infrared spectroscopy. J Pharm Biomed Anal. 2018;154:384–96.

Singh R, Ierapetritou M, Ramachandran R. System-wide hybrid MPC-PID control of a continuous pharmaceutical tablet manufacturing process via direct compaction. Eur J Pharm Biopharm. 2013;85(3 Pt B):1164–82.

Singh R, Sahay A, Karry KM, Muzzio F, Ierapetritou M, Ramachandran R. Implementation of an advanced hybrid MPC-PID control system using PAT tools into a direct compaction continuous pharmaceutical tablet manufacturing pilot plant. Int J Pharm. 2014;473(1–2):38–54.

Singh R, Sahay A, Muzzio F, Ierapetritou M, Ramachandran R. A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. Comput Chem Eng. 2014;66:186–200.

Skibsted ET, Westerhuis JA, Smilde AK, Witte DT. Examples of NIR based real time release in tablet manufacturing. J Pharm Biomed Anal. 2007;43(4):1297–305.

Smith W. Experimental design for formulation. New York: Pittsford; 2005.

Smith E, Dent G. Modern Raman spectroscopy – a practical approach: John Wiley & Sons, Ltd, Chichester, West Sussex, United Kingdom; 2005. p. 1–21.

Tabasi SH, Fahmy R, Bensley D, O'Brien C, Hoag SW. Quality by design, part I: application of NIR spectroscopy to monitor tablet manufacturing process. J Pharm Sci. 2008a;97(9):4040–51.

Tabasi SH, Fahmy R, Bensley D, O'Brien C, Hoag SW. Quality by design, part II: application of NIR spectroscopy to monitor the coating process for a pharmaceutical sustained release product. J Pharm Sci. 2008b;97(9):4052–66.

Tabasi SH, Fahmy R, Bensley D, O'Brien C, Hoag SW. Quality by design, part III: study of curing process of sustained release coated products using NIR spectroscopy. J Pharm Sci. 2008c;97(9):4067–86.

U.S. Department of Health and Human Services FDA. Guidance for industry – PAT a framework for innovative pharmaceutical development, manufacturing, and quality assurance; 2004. pp. 1–19.

Vanarase AU, Alcalà M, Jerez Rozo JI, Muzzio FJ, Romañach RJ. Real-time monitoring of drug concentration in a continuous powder mixing process using NIR spectroscopy. Chem Eng Sci. 2010;65(21):5728–33.

Vargas JM, Roman-Ospino AD, Sanchez E, Romañach RJ. Evaluation of analytical and sampling errors in the prediction of the active pharmaceutical ingredient concentration in blends from a continuous manufacturing process. J Pharm Innov. 2017;12:155–67.

Vargas JM, Nielsen S, Cardenas V, Gonzalez A, Aymat EY, Almodovar E, et al. Process analytical technology in continuous manufacturing of a commercial pharmaceutical product. Int J Pharm. 2018;538(1–2):167–78.

Wang F, Wachter JA, Antosz FJ, Berglund KA. An investigation of solvent-mediated polymorphic transformation of progesterone using in situ Raman spectroscopy. Org Process Res Dev. 2000;4(5):391–5.

Wold S, Kettaneh N, Fridén H, Holmberg A. Modelling and diagnostics of batch processes and analogous kinetic experiments. Chemometr Intell Lab Syst. 1998;44:331–40.

Yu LX. Pharmaceutical quality by design: product and process development, understanding, and control. Pharm Res. 2008;25(10):2463.

Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, et al. Understanding pharmaceutical quality by design. AAPS J. 2014;16(4):771–83.

Żarów A, Zhou B, Wang X, Pinal R, Iqbal Z. Spectroscopic and X-ray diffraction study of structural disorder in cryomilled and amorphous griseofulvin. Appl Spectrosc. 2011;65(2):135–43.

Zhang J, Ying Y, Pielecha-Safira B, Bilgili E, Ramachandran R, Romanach RJ, et al. Raman spectroscopy for in-line and off-line quantification of poorly soluble drugs in strip films. Int J Pharm. 2014;475(1–2):428–37.

Zhou X, Hines P, Borer MW. Moisture determination in hygroscopic drug substances by near infrared spectroscopy. J Pharm Biomed Anal. 1998;17(2):219–25.

Chapter 11
Active Process Control in Pharmaceutical Continuous Manufacturing – The Quality by Control (QbC) Paradigm

Qinglin Su, Sudarshan Ganesh, Gintaras V. Reklaitis, and Zoltan K. Nagy

Abstract Pharmaceutical continuous manufacturing is essentially in a steady state, or in a state of control, in process operation, by which variations in critical material/product properties and process parameters can be monitored and controlled in real time within an acceptable range that enables the comprehensive implementation of the Quality by Design (QbD) principles. This advantage, facilitated by the implementation of active process control, has recently been evolving and nurturing a new paradigm of Quality by Control (QbC) in pharmaceutical continuous manufacturing. The concept of QbC has acquired acknowledgement in recent applications in pharmaceutical continuous manufacturing and bioprocessing.

Keywords Quality by Design · Quality by Control · Continuous manufacturing · Process control

11.1 Introduction

Pharmaceutical manufacturing is traditionally operated in batch mode by processing a certain amount of raw material, defined as a batch or lot, through various steps of unit operations, such as reaction, crystallization, filtration, drying, blending, tableting, etc., to obtain final drug products, for example, oral solid dosages. Quality attributes of an in-process material or final product are tested batch wise at the end of each processing step, or known as the *Quality by Testing* (QbT) (Hubert et al. 2014). Take the crystallization step for instance; quality inspection on crystal-size distribution, polymorphic purity, etc., of the active pharmaceutical ingredient (API) is con-

Q. Su · S. Ganesh · G. V. Reklaitis ·
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Z. K. Nagy (✉)
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA

Department of Chemical Engineering, Loughborough University, Loughborough, UK
e-mail: zknagy@purdue.edu

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_11

ducted at the end of batch crystallization. Only on-spec API crystalline products are allowed to proceed to the following downstream manufacturing. Another example is in the blending step, where powders of active ingredients and excipients are mixed together in a rotating Y- or V-shaped vessel before the mixing uniformity end point is detected at the batch end (Igne et al. 2014). Although various process monitoring and control strategies are proposed for batch manufacturing in pharmaceutical or other manufacturing industries to achieve the batch-end product quality, for example, the multiway partial least squares for batch process monitoring (Nomikos and MacGregor 1995; Su and Chiu 2016), a shrinking horizon nonlinear model predictive control (NMPC) for batch-end product quality control (Su et al. 2016a), etc., it should be noted that many of them are relied on the known priori optimal or nominal process operation trajectories (or recipes), such as a cooling temperature profile in batch crystallization, it is the optimizing and tracking of these operation trajectories during the batch that lead to the batch end product quality. Rarely is a rework or a recycle of the material at the batch end, particularly at the secondary manufacturing process, allowed in the highly regulated pharmaceutical manufacturing industry. Hence, the loss of a whole batch is possible after the testing, and remedial control actions can only start from the next batch. In other words, there is a large time delay of the entire batch time in controlling the batch-end product qualities by mere end testing, which is usually known as the batch-to-batch control strategy (Mockus et al. 2015). Over the years, the US Food and Drug Administration (FDA) recognized that increased testing does not necessarily improve product quality and quality must be built into the product (Yu et al. 2014).

Quality by Design (QbD) is a concept first developed by quality pioneer Dr. Joseph M. Juran, who suggests that quality should be designed into a product and that most quality crises and problems relate to the way in which a product was designed in the first place (Juran 1992). Over the years, pharmaceutical QbD has evolved with the issuance of ICH Q8 (R2) (Pharmaceutical Development), ICH Q9 (Quality Risk Management), ICH Q10 (Pharmaceutical Quality System), and ICH Q11 (Development and Manufacture of Drug Substance). These documents provide high-level directions with respect to the scope and definitions of QbD as it applies to the pharmaceutical industry (Yu et al. 2014). QbD elements include the following: (1) a quality target product profile (QTPP) that identifies the critical quality attributes (CQAs) of a drug product; (2) product design and understanding, including identification of critical material attributes (CMAs); (3) process design and understanding, including identification of critical process parameters (CPPs) and linking CMAs and CPPs to CQAs; (4) a control strategy that includes specifications for drug substances, excipients, and drug products, as well as controls for each step of the manufacturing process; and (5) process capability and continual improvement. QbD tools and studies include priori knowledge, risk assessment, mechanistic models, design of experiments (DoE) and data analysis, and process analytical technology (PAT).

Though QbD principles are equally implemented in the conventional pharmaceutical batch manufacturing (Potter 2009), in the meantime, in QbD issuance, there have been significant advancements in science and engineering to support the

implementation of pharmaceutical continuous manufacturing over the past decade (Ierapetritou et al. 2016). Conventional batch manufacturing faces many challenges in scaling up, manufacturing cost, product quality variance, etc., while continuous manufacturing, defined as the processing of raw materials without interruption and with continuity of production over a sustained period of time, offers many benefits in the minimization of uncertainty in scale-up, reduced space and capital requirement, improved quality with consistent operations, etc. (Pernenkil 2008). A great deal of potentials for improved agility, flexibility, and robustness in the continuous manufacture of pharmaceuticals have also been demonstrated. These investments, along with the adoption of the QbD paradigm for pharmaceutical development and the advancement of PAT for designing, analyzing, and controlling manufacturing, have progressed scientific and regulatory readiness for continuous manufacturing (Lee et al. 2015). For example, pharmaceutical continuous manufacturing has already been identified as an emerging technology by the FDA, and it is now staging from conceptual designs to pilot or production processes. Recently, a few continuous manufacturing facilities and drugs have also been approved by the FDA, for example, Orkambi (lumacaftor/ivacaftor) from Vertex in 2015 and darunavir from Janssen in 2016 (Yu 2016).

Besides the flexibility and cost-saving features of pharmaceutical continuous manufacturing, a unique advantage of continuous manufacturing compared to batch manufacturing when both are under QbD guidance is that the identified CQAs and CPPs can be continuously monitored and controlled in real time, which therefore paves the way for a superior real-time product quality assurance, viz., real-time release (RTR) strategy. For example, in batch crystallization, the CQA variable of final polymorphic purity can only be available at the end of each batch crystallization step, even though this CQA variable is also continuously measured during the batch process (Su et al. 2014), whereas this final polymorphic purity can be continuously measured at the outlet of a continuous crystallizer and sent back to the control system as feedback signals to support a closed-loop control of this CQA variable by adjusting CPP variables, for example, supersaturation (Lai et al. 2014). As such, it is acknowledged that continuous manufacturing, integrated with online/inline PAT tools and efficient control systems, can accelerate the comprehensive implementation of the QbD principles for the next generation of pharmaceutical products (Singh et al. 2013; Yu et al. 2014). It is also worth mentioning that active process control strategies are not new in other continuous manufacturing processes, for example, in bulk chemicals and petrochemical industries. However, despite the feasibility of implementing active process control, which has been demonstrated for a number of pharmaceutical continuous manufacturing facilities, for example, continuous crystallization, continuous direct compaction, etc., there remains hesitation in the widespread adoption of active process control strategies due to existing investments in established and mature batch technologies (Diab and Gerogiorgis 2018). Besides, most of the pharmaceutical continuous unit operations are still operated locally with very limited flexibility to implement the much more advanced active process control system.

Hence, owing to the substantial change in the control strategy for QbD implementation in pharmaceutical batch and continuous manufacturing on the one hand and the rare application of active process control system in the pharmaceutical industry on the other hand, a *Quality by Control* (QbC) paradigm has since been evolving and nurturing the implementation of efficient advanced active process control of the CPPs and CQAs in continuous manufacturing. It has also been shown recently that QbC is indispensable to QbD implementation in pharmaceutical continuous manufacturing and that it ensures more robustness and efficiency than the techniques used in conventional batch manufacturing, providing the key ingredient toward the comprehensive implementation of QbD (Koswara and Nagy 2017).

In this chapter, the concept of QbC is further discussed in detail in the next section, including a definition of QbC, a brief review of the recent developments in research and industrial applications toward the QbC paradigm, and a perspective on the challenges of QbC implementation. This will be followed by a demonstration of the implementation of QbC: Case I, in continuous crystallization, and Case II, in continuous rotary tablet press. A concluding remark will be given at the end of this chapter.

11.2 Quality by Control

11.2.1 Definition

Ever since the release of QbD guidance, systematic process knowledge gained through pharmaceutical development by identification of critical material properties and process parameters has inherently requested the development of an appropriate control strategy to maintain process operation by controlling those critical variables. In a recent perspective on understanding the pharmaceutical QbD guidance, a three-level control strategy was discussed, in which the idea of active process control in pharmaceutical manufacturing was highlighted (Yu et al. 2014), as shown in Fig. 11.1. This was then further elaborated in modernizing the pharmaceutical manufacturing from batch to continuous production (Lee et al. 2015), laying the cornerstone for the idea of Quality by Control.

It was proposed that the control strategy for a continuous manufacturing process should be designed to maintain the quality of the product in response to potential variations or disturbances in the process, equipment conditions, incoming raw materials, or environmental factors over time. For example, as shown in Fig. 11.1, a Level 3 control imposes tight constraints on material attributes and process parameters that affect product quality and relies on extensive end-product testing at each processing step to ensure final product quality. This level of control is commonly used in batch manufacturing, viz., the Quality by Testing, by strictly tracking a recipe of operation of those parameters within constraints. This level of control requires a limited understanding of QbD with regard to how raw material and pro-

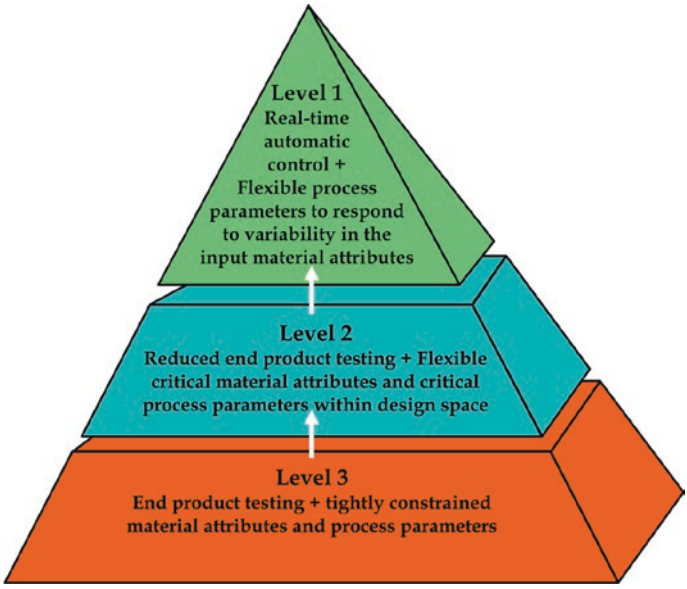


Fig. 11.1 Control strategy implementation options (Reprint permission obtained from Yu et al. 2014. Copyright © 2014 Springer Nature)

cess variability affect product quality. It is therefore too conservative and is neither feasible to be implemented in continuous manufacturing processes nor adaptable to the merits of continuous manufacturing.

The end-product testing can be reduced at a Level 2 control when raw material attributes and process parameters are within a design space, which characterizes the multidimensional combination and interaction of input variables and process parameters that have been demonstrated to provide assurance on quality (ICHQ8 2009). The design space established under the QbD guidance requires the identification of potential sources of raw material and process variability that can impact product quality, as well as the understanding of the impact that variability from these sources has on in-process materials, downstream processing, and drug product quality. Hence, manufacturing within a design space allows some flexibility in raw material and process parameters and reduces reliance on end-product testing. The intrinsically steady state operation in continuous manufacturing endeavors the process operation continuously and constantly within a design space. As a result, a Level 2 control featuring an established design space is mostly adopted in many reported continuous manufacturing facilities.

However, the operation within an often-limited design space established during product and process development could result in lack of efficiency in responding to process disturbances or variations that are commonly seen in a continuous manufacturing process. For example, when a process disturbance leads to an offset of a CQA variable from its targeted set point or acceptable range, the adjustment of a CPP variable within the design space could take a long time to bring the CQA variable

back to acceptable criteria, resulting in a long period of diversion of nonconforming materials, while a more aggressive manipulation in CPP variable, for example, with an overshoot and maybe out of the design space, is more likely to turn the CQA variable within the design space as soon as possible. Furthermore, the concept of design space is also too rigid to be adaptive to any mismatches in product and process understanding.

To this regard, a Level 1 control features an active process control system to monitor and control the quality attributes of materials in real time. In response to a disturbance, process parameters can be automatically adjusted to ensure that these quality attributes consistently conform to the established acceptance criteria. This level of control represents a high degree of product and process understanding that can also be identified under the QbD guidelines. Process design and understanding, including identification of dynamic relationships linking CMAs and CPPs to CQAs, help the design of an engineering control system in a quantitative and predictive manner. In such a way, the impact of disturbance at upstream can be minimized at the downstream processing with optimal control adjustments, and any mismatches in product and process understanding can be mitigated with an adaptive and predictive control strategy. It is the design of such a quantitative and predictive control system on the basis of the QbD guidance upon which the risk of producing off-spec product is minimized and a real-time release strategy is enabled in continuous manufacturing, is emphasized to form the foundation of Quality by Control. Herein, the definition of QbC is therefore restated as the design and implementation of an active process control system, based on high-level quantitative and predictive product and process understanding of QbD, that enables a real-time release strategy in pharmaceutical continuous manufacturing.

11.2.2 Recent Development Toward QbC

Systematic approaches to the design and implementation of active process control systems for continuous manufacturing processes have been proposed and progressively improved in the past few years by the Engineering Research Center for Structured Organic Particulate Systems (ERC-SOPS) at Rutgers University and Purdue University (Singh et al. 2009, 2014; Ierapetritou et al. 2016), in particular the integration of PAT sensor network to the control system (Singh 2018a, b), as shown in Fig. 11.2. The monitoring PAT sensors, control hardware and software required for the implementation of the active process control system, are identified and connected through step 2 to step 5. For example, it was demonstrated that feeder, comill, blender, and tablet press unit operations involved in direct compaction processes can be integrated with commercially available control platforms using process field bus (PROFIBUS), serial ports, and Open Platform Communications (OPC) protocols (Singh 2018a).

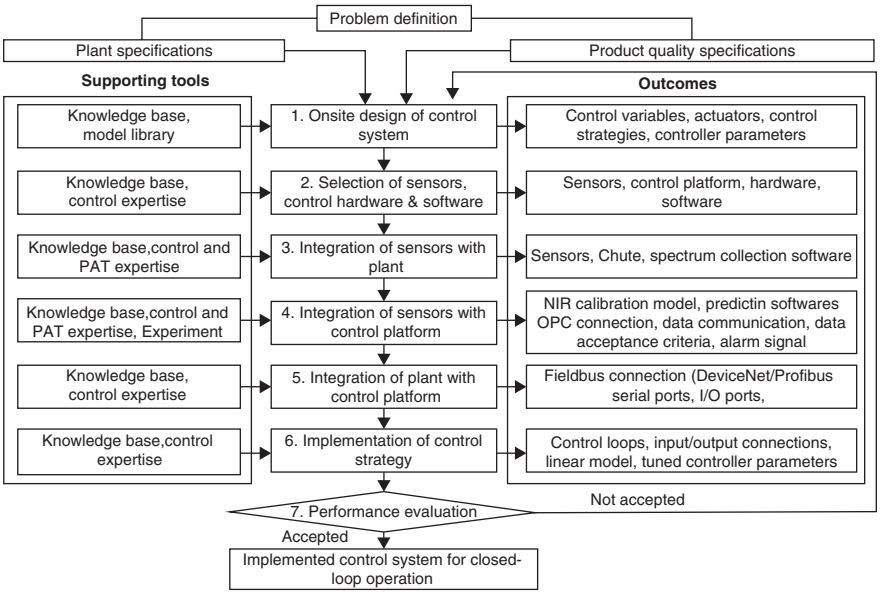


Fig. 11.2 Systematic integration methodology for PAT sensors with control system (Reprint permission obtained from Singh et al. 2014. Copyright © 2014 Elsevier)

In addition, a variety of control techniques, ranging from heuristic control algorithm, such as concentration control for crystallization (Su et al. 2014, 2015) and simple proportional-integral-derivative (PID) controllers (Yang and Nagy 2015b), to advanced model-based control and real-time optimization for pharmaceutical continuous crystallization and manufacturing processes via direct compaction, roller compaction, and wet granulation (Singh et al. 2015a, b; Ramachandran et al. 2011; Yang and Nagy 2015a, b; Singh 2018c) have been tested. Performance in set point tracking and disturbance rejection was often evaluated to finalize the control system design. The Novartis-MIT center proposed another plant-wide control strategy, for example, involving a bottom-up and hierarchical top-down approach, which stabilizes the underlying process control layers upward and prioritizes the control objectives downward (Lakerveld et al. 2013; Lakerveld et al. 2015). This work has also encouraged attempts to categorize control techniques and systematically evaluate their comparative performance with respect to aspects that are of concern to regulators. It is also recognized that process analytical technology (PAT) sensors/controls are indispensable to achieving superior control performance and to ensuring the consistent production of quality solid dosage under the range of process disturbances encountered in practice (Gupta et al. 2013).

However, to rigorously establish the continuous manufacturing processes that will consistently produce quality products, the efforts of regulators, academia, and industry are now focused on issues of risk assessment and management of quality

control, for example, ICH Q9 Quality Risk Management and Q 10 Pharmaceutical Quality System (ICH-Q9 2006; ICH-Q10 2007). The risk-based assessments of the degree of process automation and control system design, which would ensure robust operation and real-time release in continuous manufacturing, have not been thoroughly developed (Ierapetritou et al. 2016; Singh et al. 2012). Such assessments might also include the implementation of ISA-88 Batch Control Standard, ISA-95 Enterprise-Control System Integration Standard, or ISPE GAMP 5.0 Good Automation Manufacturing Practice, with an aim to achieve a resilient and fault-tolerant plant-wide active process control system design for pharmaceutical continuous manufacturing processes.

There are existing frameworks for fault-tolerant control in other continuous manufacturing industries, which have been applied over the last three decades, that can accommodate faults among system components automatically while maintaining system stability, along with a desired level of overall performance (Blanke et al. 1997; Jiang and Yu 2012). Generally, there are two main approaches to deal with faults. One is to respond to the failure by reorganizing the remaining system elements in real time to carry out necessary control functions. The other is to make the system failure proof for a certain well-defined risk/fault sets at the design stage (Jiang and Yu 2012). However, due to the rapidly emerging pharmaceutical continuous manufacturing technology, the practical implementation of fault-tolerant control, or even the basic process control, is largely uninvestigated for pharmaceutical processes. In addition to system stability, product quality under system component failures is even more important and relevant under the ICH Q9 Quality Risk Management guidance. Hence, these features provide the motivation for a systematic framework to combine existing fault-tolerant control practice with product quality concern for process control design and risk analysis in pharmaceutical continuous manufacturing processes (Singh et al. 2014; Lakerveld et al. 2013).

Recently, a systematic framework, employing appropriate process system engineering tools, to develop and evaluate feasible control strategies is briefly discussed (Su et al. 2017a). The proposed systematic framework is general and thus can be applied to the design and analysis of process control systems for any continuous pharmaceutical process, as shown in Fig. 11.3.

The framework interfaces with additional supporting knowledge and tools that facilitate the integration of software and hardware for control strategy design and implementation, as shown in Fig. 11.2. After proper product design (D2) based on the knowledge drawn from the material library (D1), an appropriate formulation for the solid dosage will be generated, requiring a specific continuous manufacturing technology at the process design (D3) level. A pilot plant or manufacturing process (P2) will then be configured using modularized unit operations (feeding, blending, tableting, etc.) provided by equipment vendors with associated integration of control platform and PAT sensors in the P1 development step. Process modeling tools consisting of a model library (S1) and flowsheet simulation (S2) software play an important role in efficient plant-wide control strategy development by accelerating process design optimization and achieving desired control objectives. For example, global sensitivity analysis and system identification (C1) based on flowsheet simu-

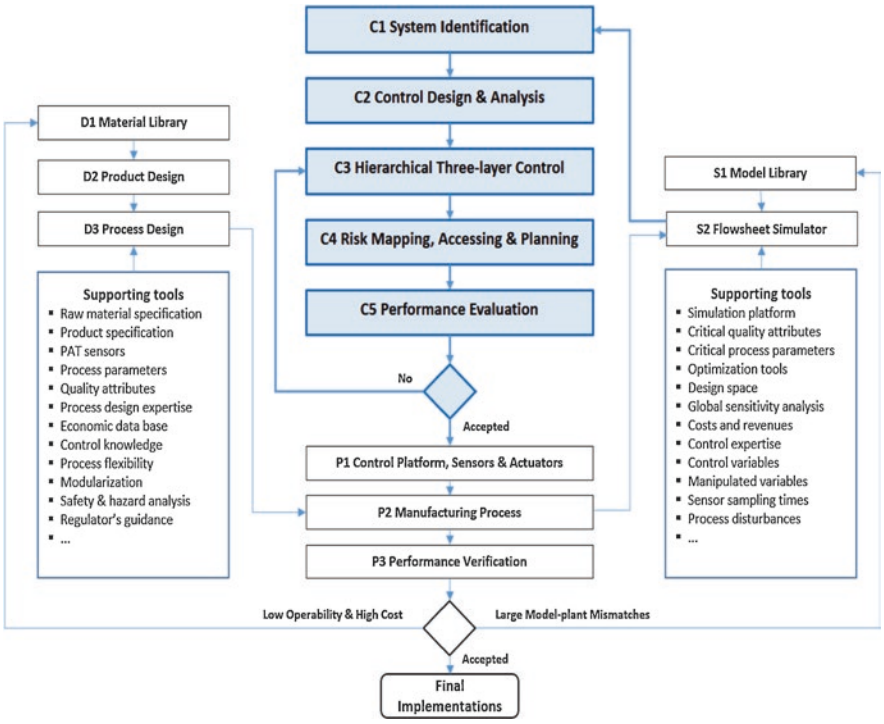


Fig. 11.3 Systematic framework for fault-tolerant process control design and risk analysis (Reprint permission obtained from Su et al. 2017a. Copyright © 2017 Springer Nature)

lation (S2), like gSOLIDS of Process Systems Enterprise (Singh et al. 2014; Rogers et al. 2013; Su et al. 2016b), SIMULINK of MathWorks (Hsu et al. 2010a, b), and JACOBIAN of RES Group (Lakerveld et al. 2013; Benyahia et al. 2012), can help to identify and address the potential challenges or risks in process control design, that is, decentralization, pairing, stability, resilience, etc., through rational control design metrics in the control design and analysis (C2), as shown in Fig. 11.4. System transfer function matrix G(s) can be first identified with the process input and output variables, followed by the three major steps of (1) selection of input/output variables to be included in the control design from all available critical process parameters (CPPs) and critical quality attributes (CQAs) based on condition number (CN), Morari's resilience index (MRI), etc.; (2) pairing and interaction analysis of the selected input/output variables using control metrics such as relative gain array (RGA) or dynamic interaction measure (DIM); and (3) closed-loop control system evaluation using dynamic flowsheet simulations, etc.

A hierarchical three-level process control design (C3) and risk analysis (C4) must be followed by rigorous performance evaluation (C5), as shown in Figs. 11.5 and 11.6. Specifically, the hierarchical control structure, in accordance to the ISA-95

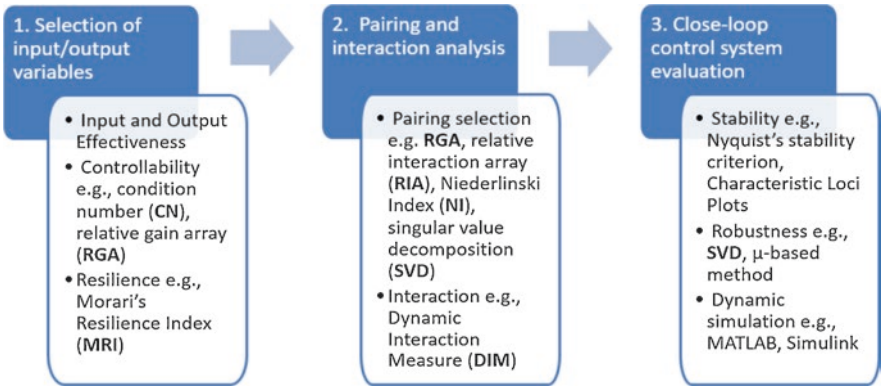


Fig. 11.4 Three-step procedure for control design and interaction analysis (Reprint permission obtained from Su et al. 2017a. Copyright © 2017 Springer Nature)

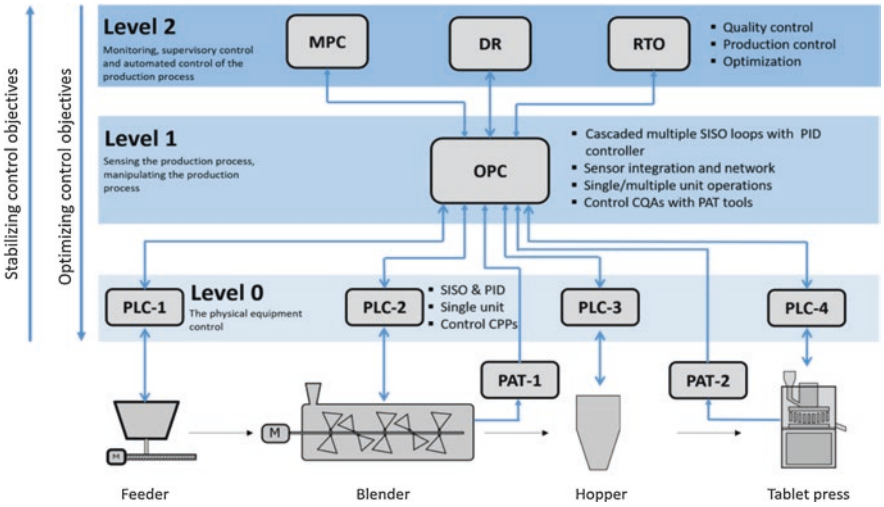


Fig. 11.5 General three-level classification of control approaches for a continuous direct compaction process (Reprint permission obtained from Su et al. 2017a. Copyright © 2017 Springer Nature)

Enterprise-Control System Integration Standard, is focused more on implementation with the levels classified according to the scale of their control objectives, the process understanding needed in QbD development, and the potential capabilities in handling process disturbances and risks. For example, in a typical continuous direct compaction process, the Level 0 control includes single/multiple-loop single input single output (SISO) control, which is often implemented via a programmable logic control (PLC) panel built into the equipment. Algorithms of simple feedback proportional, integral, and derivative (PID) control, feedforward ratio control, or

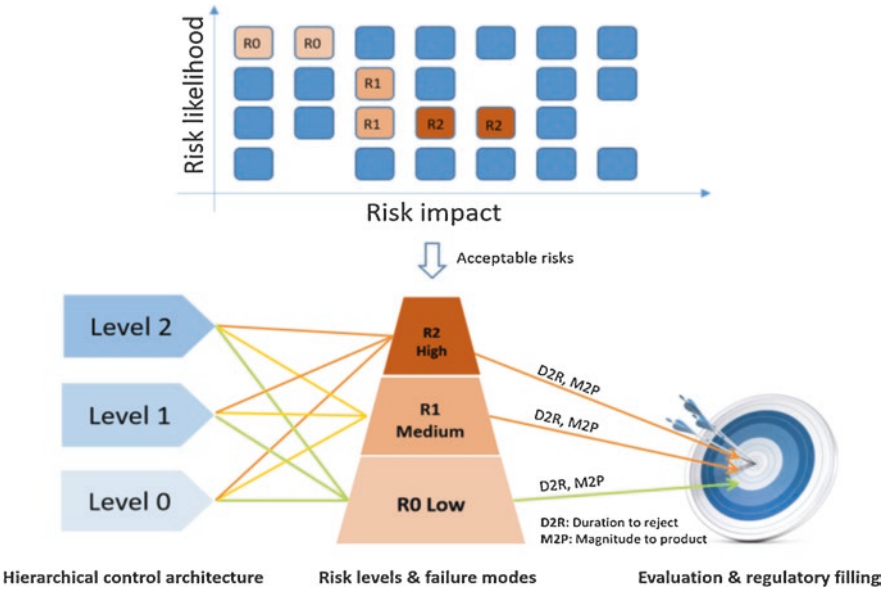


Fig. 11.6 Framework for risk-based feedback control evaluation (Reprint permission obtained from Su et al. 2017a. Copyright © 2017 Springer Nature)

advanced proprietary MPC control in case of highly nonlinear process dynamics can be implemented within the PLC of the equipment. This level of control is often designed and provided by the equipment vendor to control single/multiple critical process parameters, hence maintaining the nominal operating conditions of that equipment. The Level 1 control also involves single or multiple SISO controllers. However, in most cases, these loops rely on the use of PAT tools to measure and directly control CQAs and may encompass multiple unit operations. These control loops are commonly executed via supervisory control and data acquisition (SCADA) systems, for example, the commercial OPC servers of distributed control systems (DCSs). An OPC server retrieves data from, or sends commands to, all unit operation equipment or sensors in the field. Hence, the Level 1 control supervises the Level 0 control typically using cascaded loops with the aim of achieving desired set points for CQAs, which are measured in situ by PAT sensors. Level 1 control systems often span across unit operations and are designed using efficient feedback/feedforward control algorithms to reduce the impact of disturbance that otherwise may propagate downstream. In light of the QbD guidance, the distinguishing feature of the more advanced control approaches applied at Level 2 is the use of mathematical models/data-driven models for understanding/predicting the effects of disturbances and changes in the CPPs on the CQAs, either measured or unmeasured. Level 2 typically employs a series of advanced process control techniques to manage multiple functions, such as plant monitoring, quality control, real-time optimization (RTO), and data reconciliation (DR), among others. Level 2 control

approaches are often implemented using an independent software package that links through OPC servers to communicate with and supervise Level 1 or Level 0 controls.

A risk map for the manufacturing process can be presented in the form of a matrix that characterizes the likelihood that a risk event will occur and describes its impact on production. Only the nominal risks that are acceptable to continuous manufacturing are investigated. An acceptable risk is a risk that is understood and tolerated usually because the cost or difficulty of implementing an effective countermeasure exceeds the expected impact of the risk event on process operations. For instance, reduced flowability of powders may occur due to increased humidity in the environment during the rainy seasons of the year, and this could have an adverse effect on the mixing uniformity of the API and excipient blend. However, it might be too costly to monitor moisture content in feed materials and to add a unit operation for reducing water content to ensure specified flowability measures. The acceptable risk scenarios identified from risk mapping are further classified into three categories according to their frequency and severity: R0, low risk; R1, medium risk; and R2, high risk. Iterations between steps C3 to C5 will continue until a controller design that meets real-time release requirements is established. It is expected that by implementing an active process control strategy at different levels using the risk assessment framework, these acceptable risks can be further evaluated and accommodated.

Besides the commonly used performance indicators, such as integral of time absolute error (ITAE) for control tuning and evaluation (Seborg et al. 2004), three additional performance indicators are suggested for C5 performance evaluation in Fig. 11.3 to explicitly relate the control performance to the consistent production of on-spec products, as shown in Fig. 11.7. The first one is *time-to-product* (T2P). T2P describes the amount of time the disturbance takes to propagate from its source to the product, and it is detected as a deviation in one or more CQAs. T2P is of practical importance to determine when to start/prepare diverting the nonconforming product in case unacceptable disturbances are detected in CPPs upstream in a continuous manufacturing line. The second indicator, *duration-to-reject* (D2R), is the length of time that the process with or without a control system requires to smooth out the process disturbance or to reach a new set point for the CQA. A smaller D2R is desirable so that when a process disturbance occurs in the system, a smaller quantity of nonconforming product will be generated. A third indicator, *magnitude-to-product* (M2P), describes the maximum deviation in the CQAs from the target set point when the process is under manipulation. It is important to restrict the control actions to avoid large M2P such that the CQAs largely remain within the design space. Figure 11.7 shows the definition of these performance indicators and the potential advantages of implementing active process control that shows better performance, as defined by these criteria, when compared to the open-loop control of the CQAs, viz., manual operation.

The resulting control design will then be implemented on the manufacturing process (P2) and verified experimentally in step P3. Continuous improvements in the manufacturing process (P2) will be pursued either by improving model prediction accuracy or by enhancing product formulation and process designs (D1 & D2).

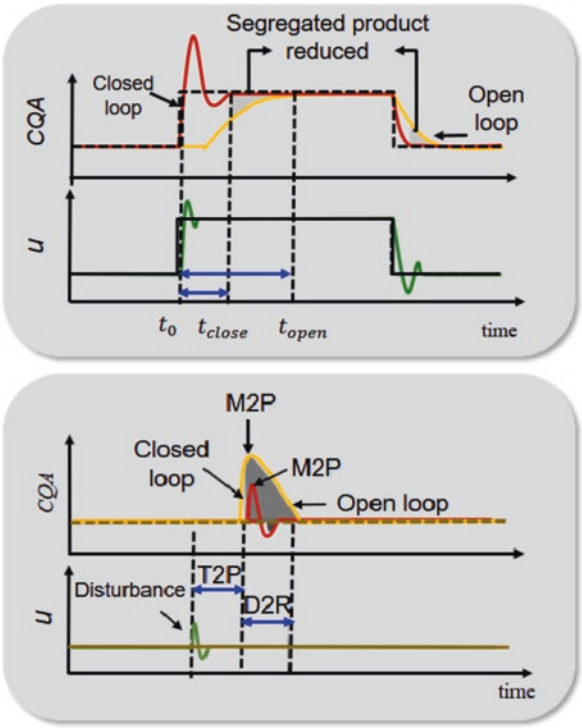


Fig. 11.7 Qualitative indicators for control performance: set point change (top) and disturbance rejection (bottom) (Reprint permission obtained from Su et al. 2017a. Copyright © 2017 Springer Nature)

For performance verification (P3) of the control design (C2) in the manufacturing process (P2), a state of statistical control in relation to the established acceptance criteria of CQAs should be attained for a stable process, whose inherent variability due to common causes could be measured with process capability (C_p and C_{pk}) indices (Yu et al. 2014). The process capability index is a widely used statistical measure of the ability of a process to produce output within specification limits, indicating how much "natural variation" a process experiences relative to its specification limits. It can be used as an indicator of how well the process is under control. Furthermore, cause-and-effect relationships can also be determined using statistical data analysis tools to seek out the root cause of inherent variability from process operation conditions and raw material quality to continuously improve and optimize the current process.

11.2.3 The Challenges of QbC

Firstly, classical process control engineering theories and experience for continuous processes have been matured and are extensively employed in fluid-based petroleum and chemical industries (Darby and Nikolaou 2012). In these applications, the

process dynamics is often driven by chemical reactions, which have response times measured in minutes to hours. However, this process control experience may or may not be directly transferable to the more challenging solid-phase unit operations typically employed in pharmaceutical secondary manufacturing, where physical changes usually occur within seconds or minutes and therefore faster response may be required of the control system (Singh et al. 2015a). Furthermore, in pharmaceutical continuous manufacturing there may be limited holdup in each unit operation, and thus the buffering provided by a material inventory is limited. Additionally, stream recycling or substantial back mixing in the process must be avoided in the highly regulated pharmaceutical secondary manufacturing process due to the necessities of material tracking (Lee et al. 2015). Thus, variability in raw materials upstream has a rapid and direct impact on downstream processes, which affect the in-process materials and final drug product qualities. In this regard, a control system with QbC design should be able to respond rapidly to the disturbance in a predictive or combined feedback and feedforward way (Singh et al. 2015a), rather than the classical feedback control design, making the consistent production of quality solid dosage challenging.

Secondly, it is worth noting that the robust mechanical design of traditional manufacturing equipment (such as a rotary tablet press, a roller compactor, etc.) has resulted in a minimum variation of CPPs and CQAs during operation and thus allowed batch pharmaceutical manufacturing to assess quality using postbatch statistical quality control (SQC) methods. A QbC active process control system, by contrast, can use possibly noisy and biased CQA measurements to effectively supervise the control of CPPs and thus minimize the need for batch-end SQC (Su et al. 2018). An important aspect to investigate is the extent to which this integration imposes additional dynamics on the process and how this could potentially amplify variations in CPPs and thus in CQAs. Data reconciliation strategy has been recently shown to be able to detect the measurement gross errors in CQA variables and reduce the measurement noise in a continuous feeding-blending system (Liu et al. 2018; Moreno et al. 2017). Further studies on using data reconciliation with joint state estimation and parameter estimation are of significance to QbC implementation to address the issue on the often uncertainties in CQA measurements with many spectroscopic PAT tools (Bagajewicz and Cabrera 2003; Câmara et al. 2017; Wu et al. 2016; Weiss et al. 1996; Rafiee and Behrouzshad 2016).

Thirdly, continuous improvement is pursued in most manufacturing sectors as a consequence of the deeper understanding of the manufacturing system and its components, which naturally develop as manufacturing experiences with a product and process are accumulated, as also identified in the QbD guidelines. Despite its potential, continuous improvement has not been pursued too aggressively in pharmaceutical manufacturing, given the real and perceived regulatory burden of approvals required for changes. Hence, the advent of continuous pharmaceutical manufacturing with the proposed QbC paradigm opens the door to continuous improvement at multiple levels, including predictive maintenance, control performance monitoring, control structure reorganizing, etc., since such improvements can be targeted to achieve tighter tracking of CQA and more robust plant-wide control, which will

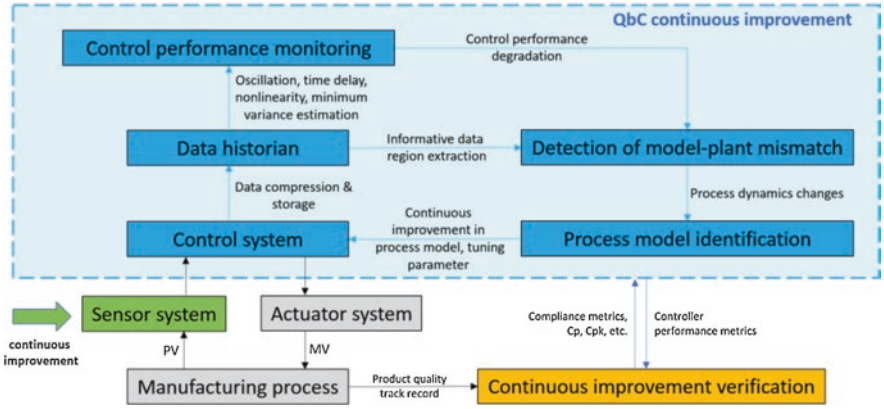


Fig. 11.8 Continuous improvement in Quality by Control system for pharmaceutical continuous manufacturing (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2019 Elsevier)

maintain the process within its design operating space. The direct impact is to allow longer continuous runs without forced interruption, reduced frequency and duration of periods during which nonconforming materials are generated, and reduced risk that a product lot released may actually include nonconforming materials. Herein, three research challenges to be addressed under QbC are centered on the use of high frequency sampled data, reduction of process-plant mismatch, and closed-loop model identification, which is also sketched in Fig. 11.8.

1. The large quantities of high-frequency-sampled data that are collected from the sensor network are often down-sampled or even compressed before being stored in a data historian. This compressed data is then recovered to track process dynamics, assess control performance, and monitor product quality. However, little is known about the effects of information loss resulting from these data storage processes, especially in continuous manufacturing lines with fast dynamics. Hence, complementing the developments in sensor network management in pharmaceutical continuous manufacturing (Cao et al. 2018), the effect of sampling time, and the delay on the use of routine closed-loop operating data for process identification should be investigated. For instance, using the condition number of the Fisher information matrix for a desired model structure (Shardt and Huang 2013), the sensitivity of the model identification method to changes in sampling time, model structure, and controller type should be evaluated as well. Several control performance monitoring schemes that have been extensively studied in the cost performance index (CPI), such as the likelihood methods (Tyler and Morari 1996) and model-invalidation mechanisms (Dumont et al. 2002), can be assessed. These studies will provide guidance on the best practices to adopt for pharmaceutical continuous manufacturing.
2. As the use of model-based control systems becomes more prevalent in pharmaceutical continuous manufacturing for Quality by Control (Muzzio et al. 2013), the issue of process model-plant mismatch will become an increasingly impor-

tant factor requiring careful monitoring and continuous improvement. Several alternative detection methodologies for model-plant mismatch in closed-loop MPC-based control systems should be implemented and evaluated for their performance (Ling et al. 2017). Particular attention should be given to the isolation of changes in plant dynamics and changes in disturbance dynamics through the use of methods such as the two-model divergence method (Jiang et al. 2009).

3. When large model-plant mismatch indicates room for control system improvement, then it is necessary to extract appropriate data from the historian to improve the model. The issue arises as to whether those data are sufficient or whether additional excitation signals need to be imposed in order to achieve improvement. Previous work has shown the possibilities of identifying process models using closed-loop routine data even if the reference signal was not being excited (Shardt and Huang 2011); this is exactly the case in pharmaceutical continuous manufacturing with priority in steady-state operation. However, this approach requires the process to have a sufficiently large time delay or the sensor network to have sufficiently fast sampling time. As a closely related issue, the minimal required excitation for close-loop identification should also be investigated (Shardt et al. 2015).

11.3 Case I: Continuous Crystallization

11.3.1 A Mindset Change in Crystallization Process Control

Crystallization is an important unit operation for separation and purification in the pharmaceutical manufacturing industry, in which a solid crystalline product with desired purity, size, and shape may be obtained from an impure feed solution. Conventionally, crystallization processes are operated in batch mode to allow flexibility to respond to varying customized design requirements and changing market demands; however, this approach can also lead to increased manufacturing costs and batch-to-batch variations in product quality, as discussed in the *Introduction* section. The advent of patent expiration of many bulk drugs (Aksu et al. 2012) has generated an increasing need to become more competitive through advanced manufacturing technologies (Plumb 2005). Hence, the development of continuous crystallization techniques has received increasing interest and research effort from academic and industrial sectors in the past decade due to the potential advantages of reduced costs, shorter development times, more robust scale-up, and improved control performance (Leuenberger 2001; Buchholz 2010).

This paradigm shift from batch to continuous crystallization in pharmaceutical manufacturing also demands a mindset change in product quality control strategies. For example, in batch crystallization process, the product quality is robustly controlled by following an optimal concentration trajectory, thereby containing the supersaturation within a predefined modest range, or widely known as the concen-

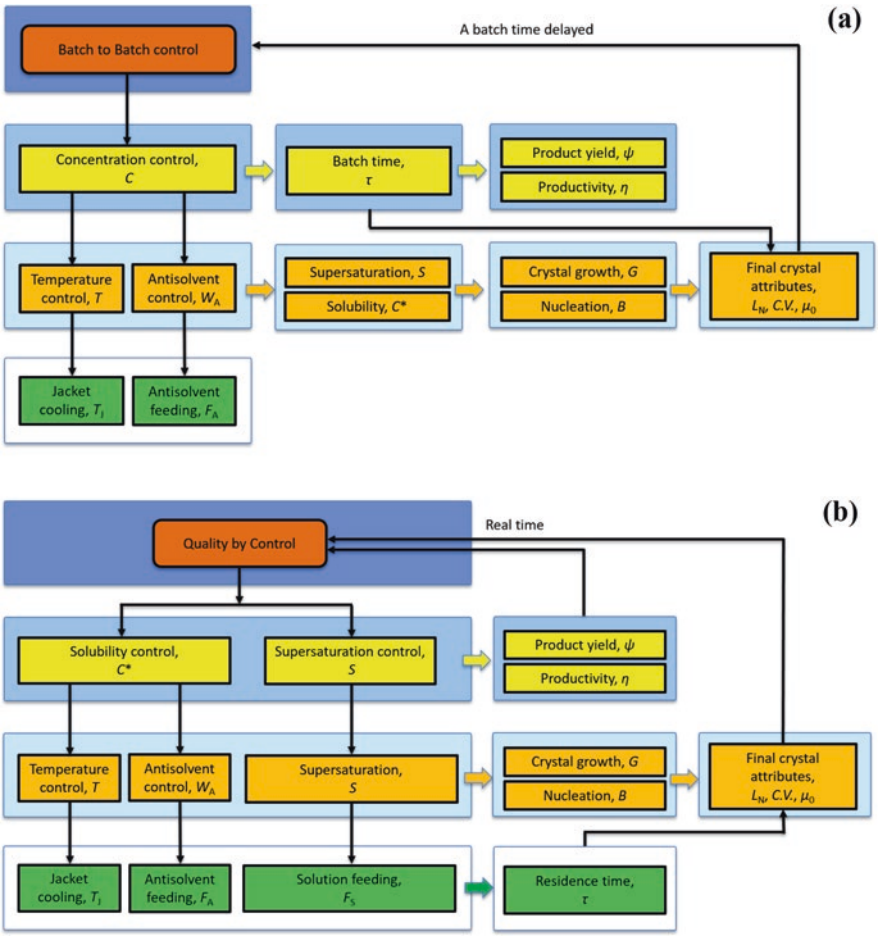


Fig. 11.9 Comparison of batch (**a**) and continuous (**b**) crystallization process control

tration control, as schemed in Fig. 11.9a. Herein, supersaturation is the difference between solute concentration and solubility and is generally regarded as the driving force for crystal nucleation and growth. Concentration control strategy has been well acknowledged for its robustness against the uncertainties in crystallization kinetics for batch crystallization control. It is pointed out that, under concentration control, only solubility is primarily manipulated by cooling and/or adding antisolvent (the addition of antisolvent, to some extent, can also affect solute concentration due to dilution effect), and batch time is extendable to adjust to the uncertainties in the crystallization kinetics (Su et al. 2014). Therefore, under the QbD guidance, only solubility could be considered as an independently manipulated critical process parameter therein.

However, in recent developments on continuous crystallization process control, research efforts often took it for granted by manipulating solubility directly to maintain a steady-state supersaturation in continuous crystallization process, for example, using the mixed suspension mixed product removal (MSMPR) crystallizers. It is highlighted in this study that feeding solution or slurry adds one more independently manipulated CPP variable to the continuous crystallization process control, resulting in the possibility of decoupling the conventional concentration control into the supersaturation and solubility controls, as demonstrated in Fig. 11.9b. Solubility is controlled by manipulating the cooling jacket temperature or the antisolvent feeding flow rate, while supersaturation is controlled by adjusting the feeding solution flow rate. This is significant for product yield and subsequent filtrate recycle since the solubility of the exiting slurry can be constantly maintained. Moreover, residence time can play an important role in addressing the uncertain crystallization kinetics, which is a risk scenario commonly identified in the crystallization process. For example, when unexpected slow crystal growth occurs, it would require a lower supersaturation and a smaller feeding flow rate, thus a longer residence time for crystal growth, and vice versa.

This mindset changes in process understanding of system dynamics in continuous crystallization based on the QbD guidance have promoted the implementation of QbC in continuous crystallization process to achieve real-time active process control of the final crystalline product quality attributes.

11.3.2 Active Process Control in Continuous MSMPR Crystallization

One of the simplest ways to demonstrate QbC control in continuous crystallization is thus to implement concentration control (C-control) with a decoupled control structure, as shown in Fig. 11.9b, which was demonstrated in a two-stage MSMPR antisolvent crystallization process of paracetamol (Su et al. 2015). Therein, at each crystallizer stage, the antisolvent feeding flow rate was adjusted simply by a PID controller to maintain constant solubility, while the feeding slurry flow rate was regulated also by a PID controller to achieve constant supersaturation. It was found that the decoupling control was robust in handling uncertainties in crystallization kinetics.

Here we define “Flow rate control” as a system in which the additional flow rates of antisolvent and fresh seeded solutions are controlled at their steady-state value, which is the common sense of operating CPPs within a design space, viz., the Level 2 control strategy in Fig. 11.1. In contrast, concentration control with decoupled control structure actively maintains solubility and supersaturation at their respective set points by regulating the antisolvent addition and fresh slurry, respectively, viz., the Level 1 control strategy in Fig. 11.1.

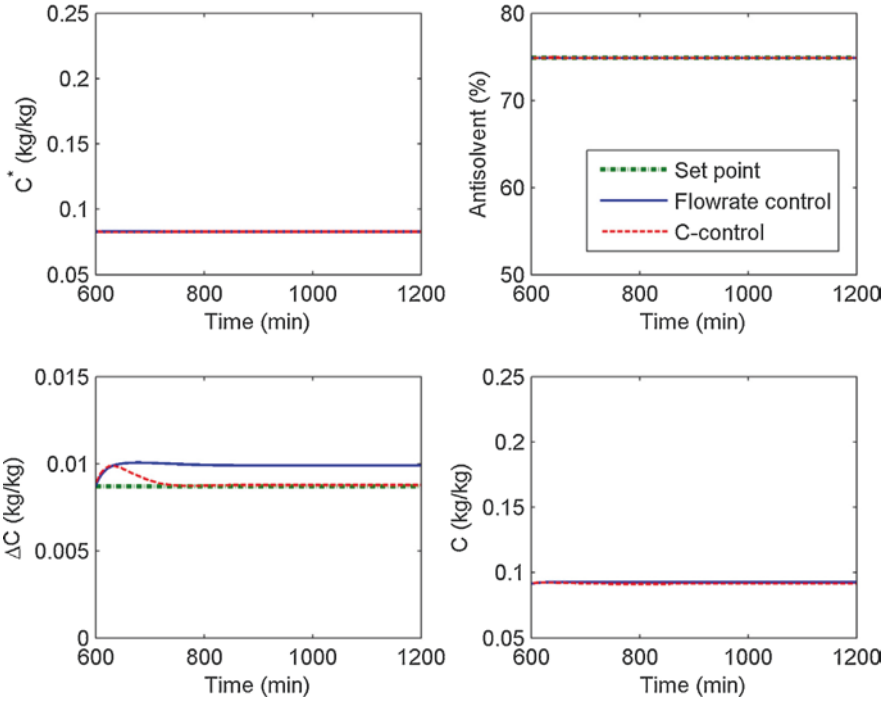


Fig. 11.10 The decoupling of solubility and supersaturation control for S2 crystallizer in a continuous two-stage MSMPR antisolvent crystallization process (Reprint permission obtained from Su et al. 2015. Copyright © 2015 Elsevier)

In Figs. 11.10 and 11.11, the crystal growth rate is reduced by 20% from its nominal value and sluggish crystal growth occurs from 600 min onward. The performances of flow rate control and C-control for the second-stage crystallizer S2 in the cascaded two-stage MSMPR crystallizers are shown. By maintaining steady-state antisolvent and fresh slurry flow rates (see Fig. 11.11), the flow rate control can hold antisolvent mass fraction, and thus solubility, at its steady-state value. However, concentration and supersaturation are disturbed from their steady-state values when uncertainty in crystal growth rate occurs, as shown in Fig. 11.10. On the contrary, the C-control procedure brings back supersaturation at its steady-state set point while maintaining antisolvent mass fraction by reducing the antisolvent feed rate in Fig. 11.11. Interestingly, the fresh slurry feed rate is again kept constant at zero. Both the antisolvent and fresh slurry feed rates were reduced in the first-stage crystallizer (results not shown here) to account for the reduced crystal growth rate for paracetamol. Comparisons of the final number-based mean crystal size for both control strategies are shown in Fig. 11.12, where the mean crystal size is nearly retained at its previous steady-state values by C-control under this risk scenario of reduced crystal growth rate.

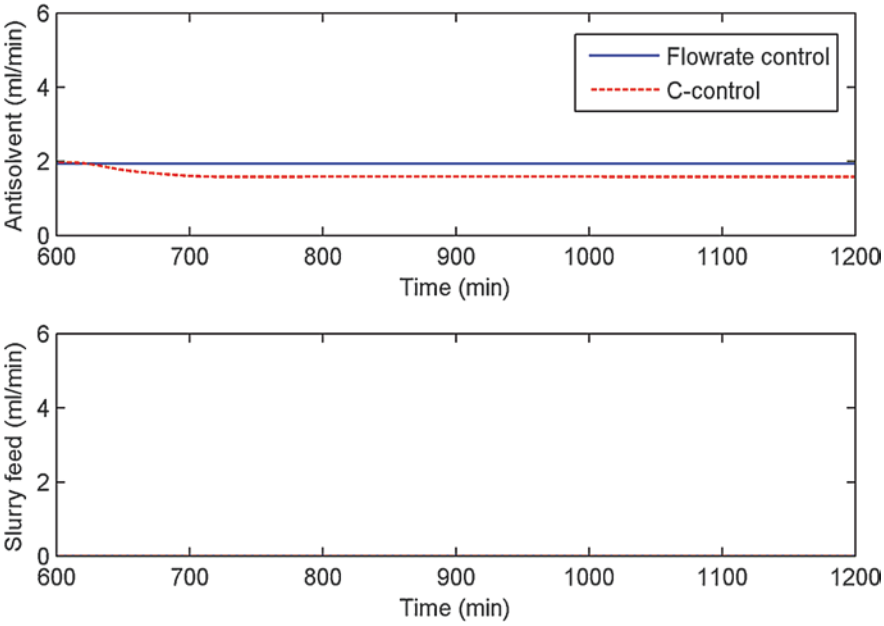


Fig. 11.11 Feeding flow rates to crystallizer S2 by flow rate control and C-control under growth rate uncertainty (Reprint permission obtained from Su et al. 2015. Copyright © 2015 Elsevier)

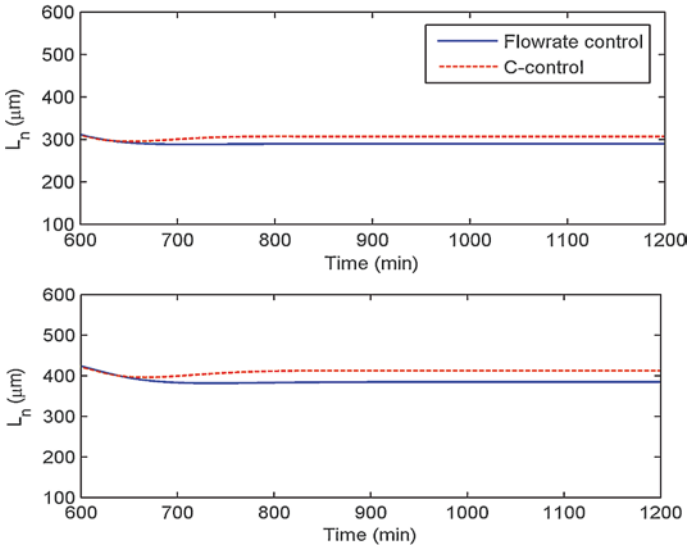


Fig. 11.12 Number-based mean crystal sizes in a cascaded two-stage crystallizer under growth uncertainty (top: crystallizer S1, bottom: crystallizer S2) (Reprint permission obtained from Su et al. 2015. Copyright © 2015 Elsevier)

Similar to the drawback of C-control in batch crystallization, which extends the batch time to tackle the sluggish crystal growth, the reduced addition flow rates by C-control for continuous operation also lead to lower productivity at steady state, that is, from nominal 14.88 g/hr to 11.94 g/hr, and longer mean residence time from nominal 89.91 minus to 105.27 minus. For moderate kinetic uncertainties (>80% reduction) in crystal growth rate, the robustness of C-control over the flow rate control is also demonstrated, although an even longer mean residence time is required. Additionally, the effect of mean residence time on crystal size has been experimentally confirmed in an innovative periodic continuous crystallization process (Su et al. 2017b). Nevertheless, with QbC design, the cost of uncertain crystal growth kinetics lies with productivity but not with final product quality attributes, for example, the mean crystal size here.

To summarize accordingly, a high-level understanding in the product and process dynamics of continuous crystallization process would ensure the design of QbC strategy, even with simple PID controllers, to achieve a robust real-time control of final product quality.

11.4 Case II: Continuous Tableting

11.4.1 Continuous Rotary Tablet Press

Tablets are the most common oral solid dosage form that can be generally manufactured by direct compression. Common processing steps for direct compression include feeding, blending, and tableting. The pilot continuous direct compression line at Purdue University was investigated in this chapter, wherein a Natoli BLP-16 rotary tablet press was considered as a case study. The tablet press is a multistage process in which each station undergoes the following major steps: die filling, metering, precompression, main compression, tablet ejection, and take-off from lower punch, as shown in Fig. 11.13. After the die filling in the feeder, the metering

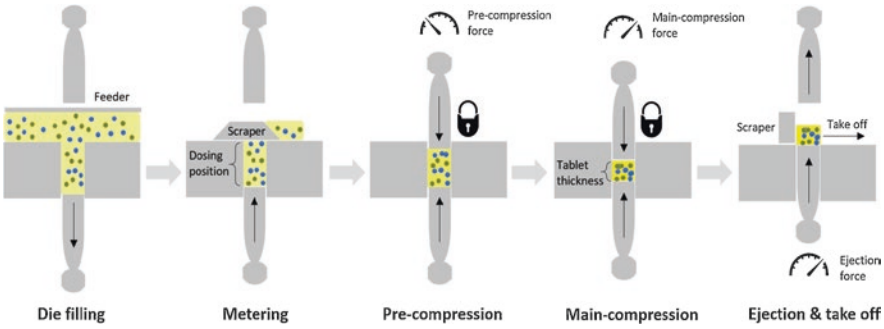


Fig. 11.13 Major steps in Natoli BLP-16 rotary tablet press (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2019 Elsevier)

stage manipulates the dosing position to adjust the volume of powder within the die, which is then locked between upper and lower punches during precompression and main compression until tablet ejection and take-off. The purpose of precompression is to remove the air trapped in the die and to rearrange the particle packing, while the purpose of main compression is to compact and solidify powder into tablet. Obviously, the tablet's weight can be controlled by changing the dosing position subject to the variation of powder bulk density or filling time due to change in turret speed. In-die tablet thickness is determined by punch displacement, which is manually set before tableting operation with the current tablet press. Hence, the maximum main-compression force is causally dependent on the amount of powder in the die or, equivalently, the tablet weight.

11.4.2 QbD Implementation

The continuous direct compaction process was integrated with state-of-the-art process analytical technologies and process control strategies to monitor the process operation with the design space and maintain consistent product quality. For example, API mass fraction was measured *in situ* using a near-infrared spectrometer (Control Development, Inc.) at the exit of the continuous blender (Vanarase et al. 2010). The powder flow rate was measured using an X-ray-based mass flow meter (SETXvue XP-300, Enurga, Inc.) (Ganesh et al. 2017). These CQA and CPP measurement data were collected using Emerson DeltaV OPC system to support a hierarchical three-level process control system design according to the ISA 95 standard (Su et al. 2017a), as shown in Fig. 11.14.

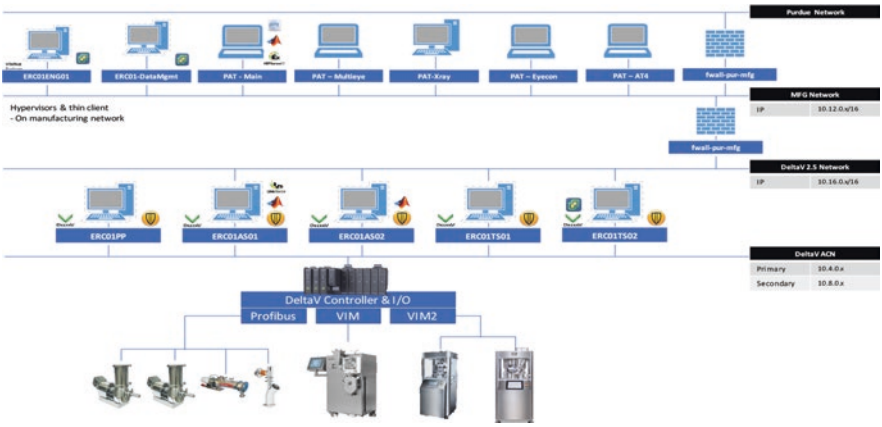


Fig. 11.14 A hierarchical three-level process control for direct compaction (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2015 Elsevier)

Specifically, the critical-to-quality variables in the tablet press were identified as tablet weight, relative density, strength, and main compression force (Su et al. 2018). The weight of the tablet ultimately determines the API potency within a dose. It also determines the main compression force at the preset punch displacement and thus the relative density and tensile strength of the tablet, which in turn affect the final product attributes, such as dissolution rate. Commercial at-line instruments, such as Sotax Auto Test 4, are often employed to measure tablet weight, as well as strength, every couple minutes. However, destructive and slow measurements cannot be efficiently integrated with the existing process control system to maintain consistent quality production in real time.

11.4.3 Data Reconciliation for Tablet Press

An in-house design for real-time tablet weight measurement based on a Mettler Toledo ME 4001E mass balance was considered in this study, which is schematically shown in Fig. 11.15. Tablets exciting the tablet press chute is collected in a bucket placed on top of the electronic mass balance. The balance is connected to a computer with an RS232 cable to transfer the measurement data of the total tablet weight via the MathWorks MATLAB instrument connection toolbox. The corre-

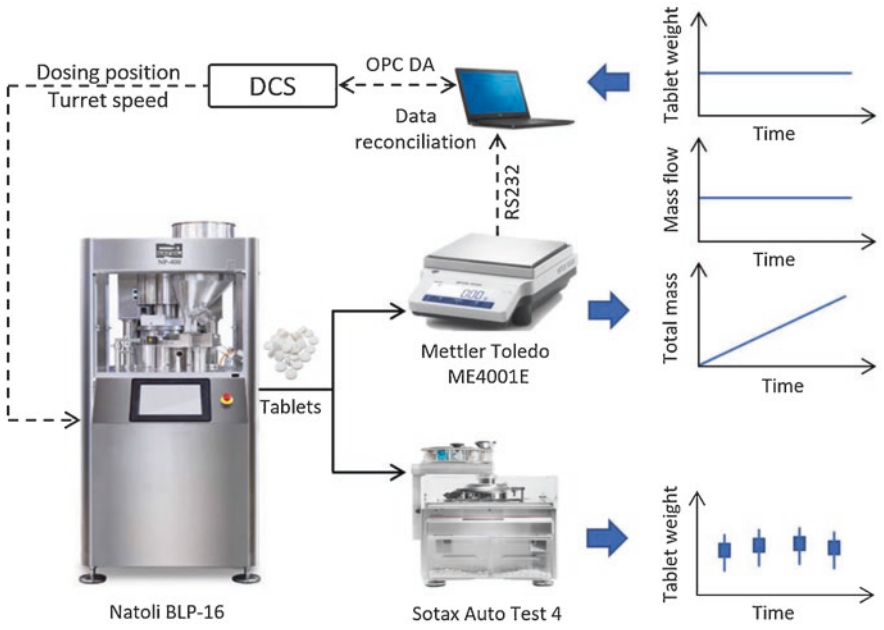


Fig. 11.15 Tablet weight measurement real-time monitoring and control (Reprint permission otained from Su et al. 2019a, b. Copyright © 2019 Elsevier)

sponding tablet production rate can be obtained by taking the first-order derivate of the total tablet weight measurement. A tablet weight measurement (W_t) is then calculated by dividing the tablet production rate with turret speed and station number, as below:

$$W_t = \frac{\text{Tablet production rate}}{\text{Turret speed} \times \text{Station number}} \tag{11.1}$$

where the turret speed determines the total counts of tablet produced per time unit. The proposed in-house design is efficient and accurate but may lack enough precision due to nonstatic measurement with the mass balance. Hence, a flexible moving window time to collect roughly 100 tablets was adopted when taking the first-order derivative to calculate the tablet production rate and tablet weight. However, gross errors can also occur when tablet samples are pneumatically diverted to the at-line Sotax Auto Test 4 device or when an empty bucket is reloaded. Therefore, a data reconciliation strategy is desired to reduce measurement uncertainty and eliminate gross errors with this tablet weight measuring device.

First, powder compressibility is commonly characterized by the relationship between the main compression force and the resulting in-die tablet relative density. In this study, the Kawakita model was employed to characterize the main compression force at high tablet relative density (Ludde 1966), as given below:

$$\frac{\text{CF}}{1-\rho_c/\rho_r} = \frac{\text{CF}}{a} + \frac{\pi D^2/4}{ab} \tag{11.2}$$

$$\rho_r = \frac{4W_t}{\pi D^2 \rho_t T} \tag{11.3}$$

where CF is the main compression force, ρ_c is the critical relative density, ρ_r is the calculated in-die relative density from tablet weight W_t, and ρ_t is the known a priori true density of the powder; parameters a and b are interpreted as the maximum degree of compression and the reciprocal of the pressure applied to attain the maximum degree of compression, respectively; D is the diameter of the die, and T is the in-die tablet thickness preset by the punch displacement for B tooling punches with a flat head and bottom. It is worth mentioning that the critical relative density ρ_c is an important material property that can be used to characterize the material and compression process, which should be real-time monitored and estimated.

A general data reconciliation framework was implemented for the tablet weight measurement to correct the imperfect measurement data and estimate the critical relative density based on the process knowledge of material compressibility:

$$\begin{aligned} &\min_{z} \sum_{t\in T} w_z^T \rho\left(z - z_m\right) + \sum_{t\in T} w_m^T \left(z_t - z_{t-1}\right) \\ &\text{subject to } z \subseteq \left[y,x,u,\theta\right] \\ &f\left(y,x,u,\theta,t\right) = 0 \\ &g\left(y,x,u,\theta,t\right) \le 0 \end{aligned} \tag{11.4}$$

where a robust estimator $\rho(\cdot)$ is preferably considered, w_z is the weight vector for the measurement error $e = (z - z_m)$, and w_m is the penalty weight vector for the successive move of the reconciled measurement or estimated process variables and model parameters; t is the process time when $f(\cdot)$ and $g(\cdot)$ are process model, and T is a moving window within which the process data set are used for dynamic data reconciliation (Liu et al. 2018).

Since it is hard to ensure normality for the measurement error with the in-house tablet weight measurement, for example, the impact of dropping tablets on the uncertainty of a nonstatic mass balance measurement changes as the total tablet mass builds up in the bucket, as well the occurrence of gross errors, a Welsch robust estimator ρ_w was used in this study to accommodate the more variety of measurement error distribution, as well to handle the nonlinearity of Eq. (11.2) (Dennis Jr and Welsch 1978; Prata et al. 2010), as shown below:

$$\rho_w\left(\varepsilon\right) = \frac{c_w^2}{2}\left\{1 - \exp\left[-\left(\frac{\varepsilon}{c_w}\right)^2\right]\right\} \tag{11.5}$$

$$\varepsilon = \frac{e}{\sigma} \tag{11.6}$$

where ε is the standardized residual, c_w is a tuning parameter, $e = (z - z_m)$ is the measurement error, and σ is the standard deviation of measurement error e, which can be estimated from the historical data for the main compression force and tablet weight measurement (Wu et al. 2016). To jointly estimate and update the Kawakita model parameter, the critical density ρ_c is included in the reconciled vector variable $z = [\mathrm{CF}, W_t, \rho_c]$. Instead of a measured parameter, a reference critical density estimated from the historical distribution of ρ_c can be considered herein (Camara et al. 2016). As such, both the uncertainty with the tablet press measurements and the model-plant mismatch in powder compressibility can be tackled in a simple, but not simpler, steady-state real-time optimization problem.

11.4.4 Active Process Control of Continuous Table Press

The reconciled tablet weight measurement was then integrated into the tablet press control system, so is the tablet production rate calculated by Eq. (11.1) with a reconciled tablet weight and turret speed. The Natoli BLP-16 tablet press has a built-in programmable logic control (PLC) panel to manipulate the process parameters of dosing position and turret speed, which is regarded as the Level 0 control in this context, as shown in Fig. 11.14. A Level 1 control with decoupled proportional integral derivative (PID) control loops was designed for a cascaded control of tablet weight, tablet production rate, and main compression force by manipulating the set points of dosing position and turret speed at the Level 0 control. A Level 2 model predictive control (MPC) was also designed, in which the main compression force was constrained and monitored as it closely related to the tablet CQAs of hardness, tensile strength, and dissolution rate (Su et al. 2018). The Emersion DeltaV Control Studio and DeltaV Predict toolbox were utilized for Level 1 and 2 control development and implementation; details of the control loops can be found in Fig. 11.16. Continuous tabletting experiment was run first with open-loop control, then Level 1 control, and Level 2 control to validate the online data reconciliation, as well as to compare control system performances, as shown in Fig. 11.17. The at-line Sotax AT4 tester was also sampling the tablets during the run to independently verify the final tablet quality.

It should be pointed out that a total of 16 tablets were analyzed at each sample time with the Sotax AT4 tester, viz., collecting all the tablets within a round of turret. In this way, variation among punch stations and variation with processing time can both be characterized; see the box plot of sampling data points and the zoomed-in inset in Fig. 11.17a. The open-loop control operation at the beginning of the tablet press run in Fig. 11.17a was to confirm that the reconciled tablet weight measurement agreed well with the at-line measurement or to allow the data reconciliation framework to automatically update the relative critical density in order to reduce a possible model-plant mismatch due to material variation; see the reconciled tablet weight measurement starting to match the Sotax AT4 measurement at the third sample and the updated relative critical density at the beginning of the operation in Fig. 11.17c. Furthermore, during the control closed-loop operation, the data reconciliation continued updating the model parameter and reached a plateau under Level 1 and Level 2 control set-point changes. Even after a reinitialization of the data reconciliation at time 3600 s, offset between reconciled tablet weight measurement and at-line Sotax AT4 measurement was observed but was then gradually reduced with the update of the relative critical density. Hence, the proposed data reconciliation framework demonstrated an important feature of online automatic calibration for tablet weight measurement and was not interfering with the control system design.

The control performance of the tablet press was promisingly good both at Level 1 and Level 2, viz., the tablet weight reached the set points steadily with both control strategies except that a Level 2 MPC control showed a more aggressive and promis-

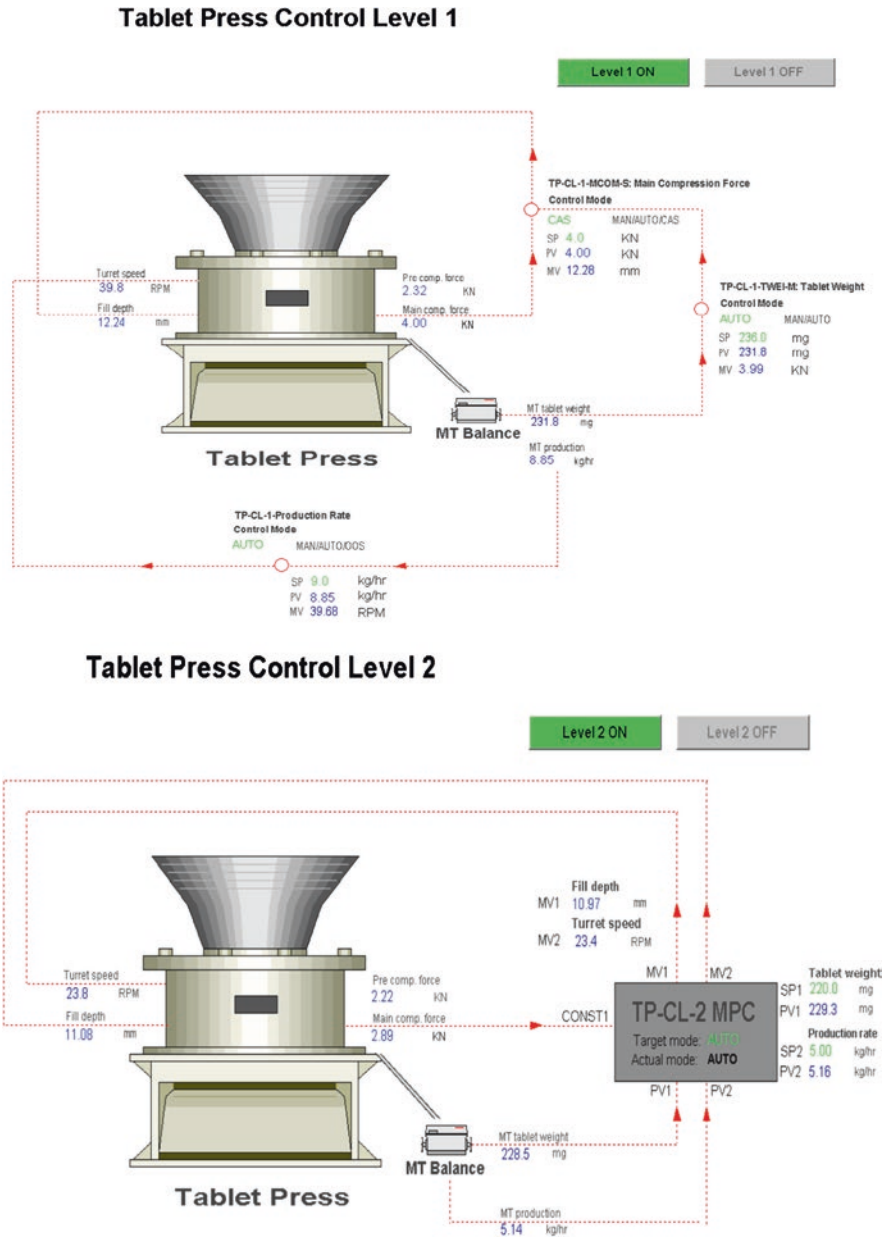


Fig. 11.16 The hierarchical level 1 (top) and level 2 (bottom) control for continuous tablet press (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2015 Elsevier)

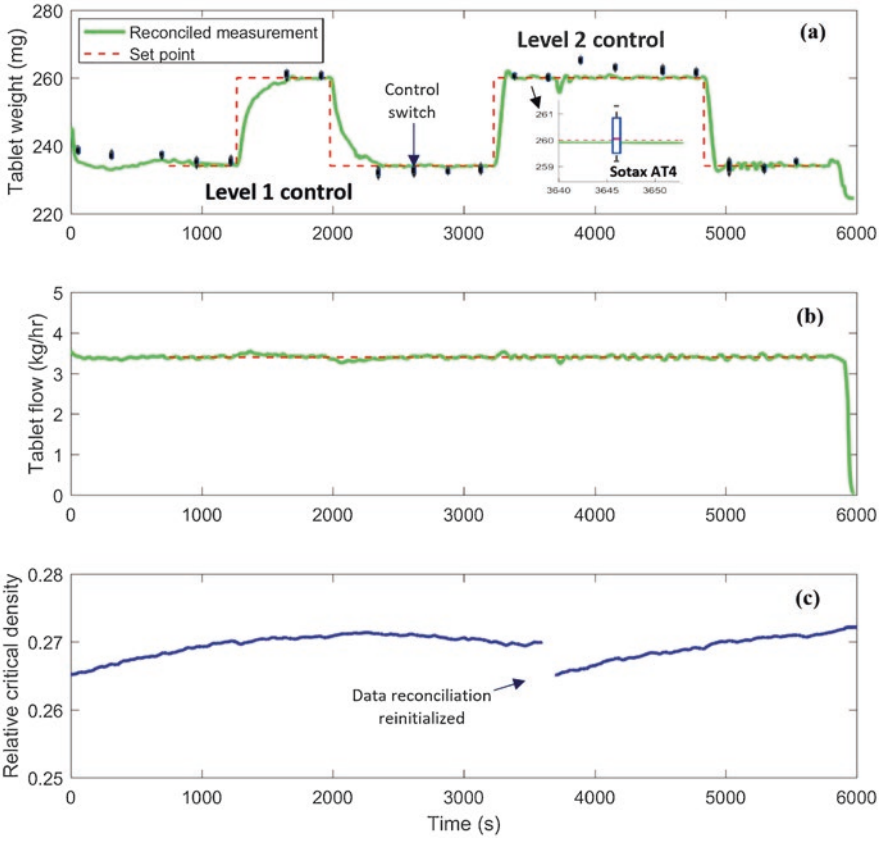


Fig. 11.17 Online integrated data reconciliation and process control for continuous tablet press (Reprint permission obtained from Su et al. 2019a, b. Copyright © 2015 Elsevier)

ing control performance. During the set point changes of tablet weight, the tablet production rate was maintained the same to adjust to the campaign production or processing capability upstream, for example, feeding and blending; see Fig. 11.17b. More importantly, variations of the tablet weight among stations at steady state remained the same as the control open-loop operation with current experiment runs; see the box plots of each sampling point in Fig. 11.17a. Overall, the control system design has shown capability in achieving process automation to reach the targeted tablet weight set point automatically and steadily, which is important during process startup or product switch. Another benefit of the process control system is the ability to maintain the tablet weight when under process disturbances or material property variations, as long as the true variations in product quality is correctly understood and not biased by measurement uncertainty.

11.5 Conclusions

Based on high-level product and process understanding gained through the Quality by Design guideline, active process control through the proposed Quality by Control can be attained for real-time release strategy in pharmaceutical continuous manufacturing. Compared to a rigid process operation within a predefined design space, active control response to common risk scenarios of process variation, disturbance, or uncertainty can be automatically and optimally calculated in a quantitative and predictive way. Two case studies are given in this chapter. Case I compares the process system dynamics in batch and continuous crystallization process and proposes a mindset in the process control of continuous crystallization process by decoupling the concentration control into solubility and supersaturation control. The risk scenario of kinetic uncertainties in the crystal growth can be efficiently handled in order to obtain consistent final mean crystal size by reducing/increasing supersaturation and mean residence time. Case II demonstrates QbC implementation in continuous rotary tablet press, where risk scenarios of measurement uncertainty with tablet weight measurements and material uncertainty in powder compressibility due to bulk density changes can be tackled in a simple, but not simpler, steady-state real-time data reconciliation problem. The implementation of an integrated model-predictive control and data reconciliation in continuous tablet press was demonstrated to be robust and efficient in process automation and real-time quality control of table weight.

Last, but not least, it is envisioned that a new paradigm in pharmaceutical continuous manufacturing is evolving in which not only quality should be originally designed by product and process understanding based on QbD but also quality can be ultimately controlled in a real-time release way, viz., Quality by Control. More research efforts toward sensor network maintenance and control performance monitoring and continuous improvement are highlighted for future work in QbC implementation.

References

Aksu B, De Beer T, Folestad S, Ketolainen J, Linden H, Lopes J, et al. Strategic funding priorities in the pharmaceutical sciences allied to quality by design (QbD) and process analytical technology (PAT). Eur J Pharm Sci. 2012;47:402–5.

Bagajewicz MJ, Cabrera E. Data reconciliation in gas pipeline systems. Ind Eng Chem Res. 2003;42(22):5596–606.

Benyahia B, Lakerveld R, Barton PI. A plant-wide dynamic model of a continuous pharmaceutical process. Ind Eng Chem Res. 2012;51(47):15393–412.

Blanke M, Izadi-Zamanabadi R, Bogh R, Lunau ZP. Fault-tolerant control systems-a holistic view. Control Eng Pract. 1997;5(5):693–702.

Buchholz S. Future manufacturing approaches in the chemical and pharmaceutical industry. Chem Eng Process. 2010;49:993–5.

Camara MM, Quelhas AD, Pito JC. Performance evaluation of real industrial RTO systems. Process. 2016;4:44.

Câmara MM, Soares RM, Feital T, Anzai TK, Diehl FC, Thompson PH, Pinto JC. Numerical aspects of data reconciliation in industrial applications. Processes. 2017;5(4):56.

Cao H, Mushnoori S, Higgins B, Kollipara C, Fermier A, Hausner D, et al. A systematic framework for data management and integration in a continuous pharmaceutical manufacturing processing line. Process. 2018; In review

Darby ML, Nikolaou M. MPC: current practice and challenges. Control Eng Pract. 2012;20:328–42.

Dennis JE Jr, Welsch RE. Techniques for nonlinear least squares and robust regression. Commun Stat Simul Comput. 1978;7(4):345–59.

Diab S, Gerogiorgis DI. Process modelling, simulation and technoeconomic evaluation of crystallization antisolvents for the continuous pharmaceutical manufacturing of rufinamide. Comput Chem Eng. 2018;111:102–14.

Dumont G, Kammer L, Alison B, Ettaleb L, Roche A. Control system performance monitoring: new developments and practical issues. The IFAC 15th Triennial World Congress, Barcelona, Spain, 2002, pp. 387–392.

Ganesh S, Troscinski R, Schmall N, Lim J, Nagy Z, Reklaitis G. Application of x-ray sensors for in-line and non-invasive monitoring of mass flow rate in continuous tablet manufacturing. J Pharm Sci. 2017;106(12):3591–603.

Gupta A, Giridhar A, Venkatasubramanian V, Reklaitis GV. Intelligent alarm management applied to continuous pharmaceutical tablet manufacturing: an integrated approach. Ind Eng Chem Res. 2013;52:12357–68.

Hsu S-H, Reklaitis GV, Venkatasubramanian V. Modeling and control of roller compaction for pharmaceutical manufacturing. Part I: process dynamics and control framework. J Pharm Innov. 2010a;5:14–23.

Hsu S-H, Reklaitis GV, Venkatasubramania V. Modeling and control of roller compaction for pharmaceutical manufacturing part II: control system design. J Pharm Innov. 2010b;5:24–36.

Hubert C, Lebrun P, Houari S, Ziemons E, Rozet E, Hubert P. Improvement of a stability-indicating method by Quality-by-Design versus Quality-by-Testing: a case of a learning process. J Pharm Biomed Anal. 2014;88:401–9.

ICH-Q10. Q10: pharmaceutical quality system. ICH Tripartite guidelines. International conference on harmonization of technical requirements for registration of pharmaceuticals for human use, 2007.

ICHQ8. ICH harmonized Tripartite guideline: pharmaceutical development Q8 (r2), current step 4 Version. The international conference on harmonisation. Geneva, Switzerland, 2009.

ICH-Q9. Q9: quality risk management. ICH harmonized Tripartite guidelines. International conference on harmonization of technical requirements for registration of pharmaceuticals for human use, 2006.

Ierapetritou M, Muzzio F, Reklaitis G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AICHE J. 2016;62(6):1846–62.

Igne B, Juan A, Jaumot J, Lallemand J, Preys S, Drennen J, Anderson C. Modeling strategies for pharmaceutical blend monitoring and end-point determination by near-infrared spectroscopy. Int J Pharm. 2014;473(1–2):219–31.

Jiang J, Yu X. Fault-tolerant control systems: a comparative study between active and passive approaches. Ann Rev Control. 2012;36(1):60–72.

Jiang H, Huang B, Shah S. Closed-loop model validation based on the two-model divergence method. J Process Control. 2009;19(4):644–55.

Juran J. Juran on quality by design: the new steps for planning quality into goods and services. New York: The Free Press; 1992.

Koswara A, Nagy ZK. On-off feedback control of plug-flow crystallization: a case of Quality-by-Control in continuous manufacturing. IEEE Life Sci Lett. 2017;3(1):1–4.

Lai T, Ferguson S, Palmer L, Trout B, Myerson A. Continuous crystallization and polymorph dynamics in the L-Glutamic acid system. Org Process Res Dev. 2014;18:1382–90.

Lakerveld R, Benyahia B, Braatz RD, Barton PI. Model-based design of a plant-wide control strategy for a continuous pharmaceutical plant. AICHE J. 2013;59(10):3671–85.
Lakerveld R, Benyahia B, Heider PL, Zhang H, Wolfe A, Testa CJ, et al. The application of an automated control strategy for an integrated continuous pharmaceutical pilot plant. Org Process Res Dev. 2015;19:1088–100.
Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, et al. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10(3):191–9.
Leuenberger H. New trends in the production of pharmaceutical granules: batch versus continuous processing. Eur J Pharm Biopharm. 2001;52:289–96.
Ling D, Zheng Y, Zhang H, Yang W, Tao B. Detection of model-plant mismatch in closed-loop control system. J Process Control. 2017;57:66–79.
Liu J, Su Q, Moreno M, Laird C, Nagy Z, Reklaitis G. Robust state estimation of feeding-blending systems in continuous pharmaceutical manufacturing. Chem Eng Res Des, Accepted, 2018.
Ludde KH. Die Pulverkompression. Pharmazie. 1966;21:393–403.
Mockus L, Peterson J, Lainez J, Reklaitis G. Batch-to-Batch variations: a key compoenet for modeling chemical manufacturing processes. Org Process Res Des. 2015;19:908–14.
Moreno M, Liu J, Ganesh S, Su Q, Yazdanpanah N, O'Connor T, et al. Steady-state data reconciliation of a direct compression tableting line. AIChE Annual Meeting. Minneapolis: AIChE Annual Meeting, 2017.
Muzzio F, Singh R, Chaudhury A, Rogers A, Ramachandran R, Ierapetritou M. Model-predictive design, control, and optimization: applying model-predictive methods and a continuous process-control framework to continuous tablet manufacturing processes. Pharm Technol. 2013;37(6):40–1.
Nomikos P, MacGregor J. Multi-way partial least squares in monitoring batch processes. Chemom Intell Lab Syst. 1995;30:97–108.
Pernenkil L. Continuous blending of dry pharmaceutical powders. Boston: Massachusetts Institute of Technology; 2008.
Plumb K. Continuous processing in the pharmaceutical industry changing the mindset. Chem Eng Res Des. 2005;83(A6):730–8.
Potter C. PQLI application of science- and risk-based approaches (ICH Q8, Q9, and Q10) to exisiting products. J Pharm Innov. 2009;4(1):4–23.
Prata DM, Schwaab M, Lima EL, Pinto JC. Simultaneous robust data reconciliation and gross error detection through particle swarm optimization for an industrial polypropylene reactor. Chem Eng Sci. 2010;65:4943–54.
Rafiee A, Behrouzshad F. Data reconciliation with application to a natural gas processing plant. J Nat Gas Sci Eng. 2016;31:538–45.
Ramachandran R, Arjunan J, Chaudhury A, Ierapetritou MG. Model-based control-loop performance of a continuous direct compaction process. J Pharm Innov. 2011;6:249–63.
Rogers AJ, Hashemi A, Ierapetritou MG. Modeling of particular processes for the continuous manufacture of solid-based pharmaceutical dosage forms. Processes. 2013;1:67–127.
Seborg DE, Edgar TF, Mellichamp DA. Process dynamics and control: Hoboken NJ. Wiley; 2004.
Shardt Y, Huang B. Closed-loop identification with routine operating data: effect of time delay and sampling time. J Process Control. 2011;21(7):997–1010.
Shardt Y, Huang B. Data quality assessment of routine operating data for process identification. Comput Chem Eng. 2013;55:19–27.
Shardt Y, Huang B, Ding S. Minimal required excitation for closed-loop identification: some implications for data-driven, system identification. J Process Control. 2015;27:22–35.
Singh R. Chapter 17 Automation of continuous pharmaceutical manufacturing process. Comput Aided Chem Eng. 2018a;41:431–46.
Singh R. Chapter 18 Implementation of control system into continuous pharmaceutical manufacturing pilot plant (powder to tablet). Comput Aided Chem Eng. 2018b;41:447–69.

Singh R. Chapter 13 Model-based control system design and evaluation for continuous tablet manufacturing processes (via direct compaction, via roller compaction, via wet granulation). Comput Aided Chem Eng. 2018c;41:317–51.
Singh R, Gernaey KV, Gani R. Model-based computer-aided framework for design of process monitoring and analysis systems. Comput Chem Eng. 2009;33(1):22–42.
Singh R, Ierapetritou M, Ramachandran R. An engineering study on the enhanced control and operation of continuous manufacturing of pharmaceutical tablets via roller compaction. Int J Pharm. 2012;438:307–26.
Singh R, Godfrey A, Gregertsen B, Muller F, Gernaey K, Gani R, Woodley J. Systematic substrate adoption methodology (SAM) for future flexible, generic pharmaceutical production process. Comput Chem Eng. 2013;58:344–68.
Singh R, Sahay A, Muzzio F, Ierapetritou M, Rohit R. A systematic framework for onsite design and implementation of a control system in a continuous tablet manufacturing process. Comput Chem Eng. 2014;66:186–200.
Singh R, Muzzio FJ, Ierapetritou M, Rohit R. A combined feed-forward/feed-back control system for a QbD-based continuous tablet manufacturing process. Processes. 2015a;3:339–56.
Singh R, Sen M, Ierapetritou M, Ramachandra R. Integrated moving horizon-based dynamic real-time optimization and hybrid MPC-PID control of a direct compaction continuous tablet manufacturing process. J Pharm Innov. 2015b;10:233–53.
Su Q-L, Chiu M-S. Monitoring pH-shift reactive crystallization of L-Glutamic acid using moving window MPCA. J Chem Eng Jpn. 2016;49(7):680–8.
Su Q-L, Braatz RD, Chiu M-S. JITL-based concentration control for semi-batch pH-shift reactive crystallization of L-glutamic acid. J Process Control. 2014;24:415–21.
Su Q, Nagy ZK, Rielly CD. Pharmaceutical crystallisation processes from batch to continuous operation using MSMPR stages: modelling, design, and control. Chem Eng Process Process Intensif. 2015;89:41–53.
Su Q, Hermanto M, Braatz R, Chiu M. Just-in-Time-Learning based extended prediction self-adaptive control for batch processes. J Process Control. 2016a;43:1–9.
Su Q, Schiano S, Wu CY, Nagy ZK, Rielly CD. Dynamic impact milling model with a particle-scale breakage kernel. Comput Aided Chem Eng. 2016b;38:475–80.
Su Q, Moreno M, Giridhar A, Reklaitis GV, Nagy ZK. A systematic framework for process control design and risk analysis in continuous pharmaceutical solid-dosage manufacturing. J Pharm Innov. 2017a;12:327–46.
Su Q, Rielly CD, Powell KA, Nagy ZK. Mathematical modelling and experimental validation of a novel periodic flow crystallisation processes using MSMPR crystallizers. AICHE J. 2017b;63(4):1313–27.
Su Q, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK. Variation and risk analysis in tablet press control for continuous manufacturing of solid dosage via direct compaction. The 13th international symposium on process systems engineering PSE 2018. San Diego, 2018.
Su Q, Bommireddy Y, Shah Y, Ganesh S, Moreno M, Liu J, et al. Data reconciliation in the Quality-by-Design (QbD) implementation of pharmaceutical continuous tablet manufacturing. Int J Pharm. 2019a;563:259–72.
Su Q, Ganesh S, Moreno M, Bommireddy Y, Gonzalez M, Reklaitis GV, Nagy ZK. A perspective on Quality-by-Control (QbC) in pharmaceutical continuous manufacturing. Comput Chem Eng. 2019b;125:216–31.
Tyler M, Morari M. Performance monitoring of control systems using likelihood methods. Automatica. 1996;32(8):1145–62.
Vanarase AU, Alcalà M, Jerez Rozo JI, Muzzio FJ, Romañach RJ. Real-time monitoring of drug concentration in a continuous powder mixing process using NIR spectroscopy. Chem Eng Sci. 2010;65:5728–33.
Weiss GH, Romagnoli JA, Islam KA. Data reconciliation-an industrial case study. Comput Chem Eng. 1996;20(12):1441–9.

Wu SY, Chen C, Gu X. Research on data reconciliation based on generalized T distribution with historical data. Neurocomputing. 2016;175:808–15.

Yang Y, Nagy Z. Combined cooling and antisolvent crystallization in continuous mixed suspension, mixed product removal cascade crystallizers: steady-state and startup optimization. Ind Eng Chem Res. 2015a;54:5673–82.

Yang Y, Nagy ZK. Advanced control approaches for combined cooling/antisolvent crystallization in continuous mixed suspension mixed product removal cascaded crystallizers. Chem Eng Sci. 2015b;127:362–73.

Yu L. Continuous manufacturing has a strong impact on drug quality. 2016, April 12. Retrieved April 11, 2017, from FDA Voice: https://blogs.fda.gov/fdavoice/index.php/2016/04/continuous-manufacturing-has-a-strong-impact-on-drug-quality/.

Yu LX, Amidon G, Khan MA, Hoag SW, Polli J, Raju GK, Woodcock J. Understanding pharmaceutical quality by design. AAPS J. 2014;16(4):771–83.

Chapter 12
Real-Time Optimization: How to Change Setpoints in Pharmaceutical Manufacturing

Arun Giridhar and Gintaras V. Reklaitis

Abstract Real-time optimization (RTO) is an important requirement of pharmaceutical production processes. Production processes are operated by adjusting the setpoints of one or more individual unit operations, either to optimize a given objective function or to satisfy externally imposed constraints. Most commonly, setpoints are changed based on inputs from other components of a larger real-time process management (RTPM) system. Since setpoint changes may have knock-on effects elsewhere in the process, it is important to treat setpoint changes not as an isolated activity but as part of a more comprehensive process management suite. This chapter describes common manufacturing scenarios that require setpoints to be changed or optimized, discusses the trade-offs between static optimization based on a priori flowsheet simulation and dynamic optimization based on current process data, and describes good practices in such setpoint transitions.

Keywords Real time · Optimization · Process management · Fault diagnosis · Production · Operations · Process engineering · Pharmaceutical manufacturing

12.1 Introduction

Real-time optimization (RTO) is an important requirement of pharmaceutical production processes. Production processes are operated by adjusting the setpoints of one or more individual unit operations, either to optimize a given objective function or to satisfy externally imposed constraints.

The functionality of real-time optimization is better understood as part of a larger suite of functionalities providing real-time process management (RTPM), which will serve to put RTO in context from the control perspective as well. A real-time process management system is a full-fledged automatic control system for running

A. Giridhar (✉) · G. V. Reklaitis
Davidson School of Chemical Engineering, Purdue University, West Lafayette, IN, USA
e-mail: agiridha@purdue.edu; reklaiti@purdue.edu

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_12

a pharmaceutical production plant continuously while maintaining product quality; such a system would include hardware for process control and communication, process management algorithms, and communication protocols. A description of hardware and communication protocols is beyond the scope of this chapter, beyond noting that commercial distributed control systems (DCS) and supervisory control and data acquisition (SCADA) systems address nearly all hardware and communication needs. We henceforth restrict ourselves to process management algorithms. Such algorithms broadly implement three distinct layers of process management: setpoint changing, setpoint enforcing, and fault correction. In the following sections, the scope and use of real-time optimization in pharmaceutical manufacturing will be examined in more detail.

12.2 Materials

The methods described in this chapter are applicable to many different materials, processes, and unit operations in pharmaceutical manufacturing. They have been developed and tested based on the following processes:

- A dry granulation line making placebo tablets (microcrystalline cellulose) or tablets with an active ingredient (acetaminophen, microcrystalline cellulose, magnesium stearate, silicon dioxide)
- A drop-on-demand/hot-melt extrusion line making solid oral drug products for high potency drugs and clinical trial applications. Materials used were naproxen, polyethylene glycol 3350, Pluronic F35, polyvinylpyrrolidone, solvents, and solubilizers

12.3 Methods

Real-time process management systems implement three distinct layers of process management: setpoint changing, setpoint enforcing, and fault correction. Each of these will be described in turn.

12.3.1 Changing Setpoints

Changing setpoints requires knowledge of two distinct pieces of information: the desired value of each setpoint of interest and a desired or acceptable trajectory to transition from the current setpoint to the desired setpoints. Calculating the desired value of each setpoint in general requires a mathematical model in some form, and its solution may take the form of either satisfying many simultaneous linear or

nonlinear equations or optimizing an objective function such as cost or production rate. Predicting the optimal trajectory requires knowledge of the process dynamics, including the effect of unit operation setpoints on product quality. The primary functionality provided by real-time optimization is the calculation of new setpoints and the best way to transit to them. More applications and methods are described by Cutler and Perry (1983), Marlin and Hrymak (1997), Dua et al. (2006), Moritz Diehl et al. (2002), and Ariyur and Krstic (2003).

The practical need to change process setpoints in pharmaceutical manufacturing is best understood by considering a few scenarios across the entire process. There are many such scenarios for why process setpoints may change in pharmaceutical manufacturing.

A common scenario in continuous manufacturing is to increase or decrease production rate to match market demand and meet planned production targets; such simple changes on a process-wide scale may be more complicated at individual equipment scale. Considering a tableting unit operation, if the throughput needs to double, it is fairly obvious that the turret revolution per minute (RPM) should double; but how should the compression force and turret RPM relate to each other? Should the compression force double as well to compensate for the halved dwell time of compression, or should it stay constant, or should it scale differently? Will tablet quality (weight, hardness, density, friability, etc.) be achieved easily across the entire range of production rates? Similar considerations apply to other unit operations across pharmaceutical manufacturing, such as roller compaction, wet granulation, continuous blending, tablet coating, and continuous milling. Thus, scaling the throughput requires more careful setpoint calculation than simply scaling everything by the same amount.

A second scenario is to handle materials with varying physical properties. For example, as powders are refilled in hoppers from storage, their cohesiveness may vary significantly, often owing to storage conditions. If a wet granulation process is supposed to granulate powders whose stickiness varies, should the granulation setpoints (RPM, liquid:solid ratio, etc.) be the same or should they change? Similar considerations apply to other unit operations as well. Knowing what setpoints to use for raw materials of different properties also requires an understanding of how raw material properties and equipment settings jointly affect product quality.

A third scenario is when throughput and product quality are both being maintained, but there exist several different combinations of process parameters that can each give the same product quality and throughput. In such a case, it is possible to optimize process parameters to reduce cost or to optimize another merit function. Real-time optimization provides the functionality to calculate what setpoints ought to be and how to achieve them based on external constraints or on given objective functions.

For each of the scenarios described above, changing the setpoint may be static or dynamic. In static RTO, only the desired setpoint is calculated for every unit operation, and the transition from the current steady state to the new steady state is implicitly assumed to be a (possibly large) step change. Such a transition is simple to implement but may result in a large quantity of off-spec material during the

transient or a difficult-to-control transition depending on the process. In dynamic RTO, the desired setpoint as well as the trajectory of the process from current setpoints to the new setpoints are both calculated. Dynamic RTO is useful for processes that exhibit control difficulties like inverse responses or significant overshoots/undershoots; a special application of dynamic RTO is to start up or shut down a process automatically.

Static Setpoint Optimization

It is best to examine a simple case study in detail to see how a static setpoint optimization exercise can be carried out. Consider a continuous dry granulation production line (Fig. 12.1). A necessary first step in running such a line continuously is to ensure that all individual unit operation setpoints are matched with each other such that production rates are balanced (there is no overflow or underflow on any of the intermediate hoppers) and the product quality is acceptable. Such a steady-state mass balance is easily modeled for any process, as shown below for a static RTO application. Let a dry granulation process be modeled as having the following parameters:

Overall process throughput (kg/hour)

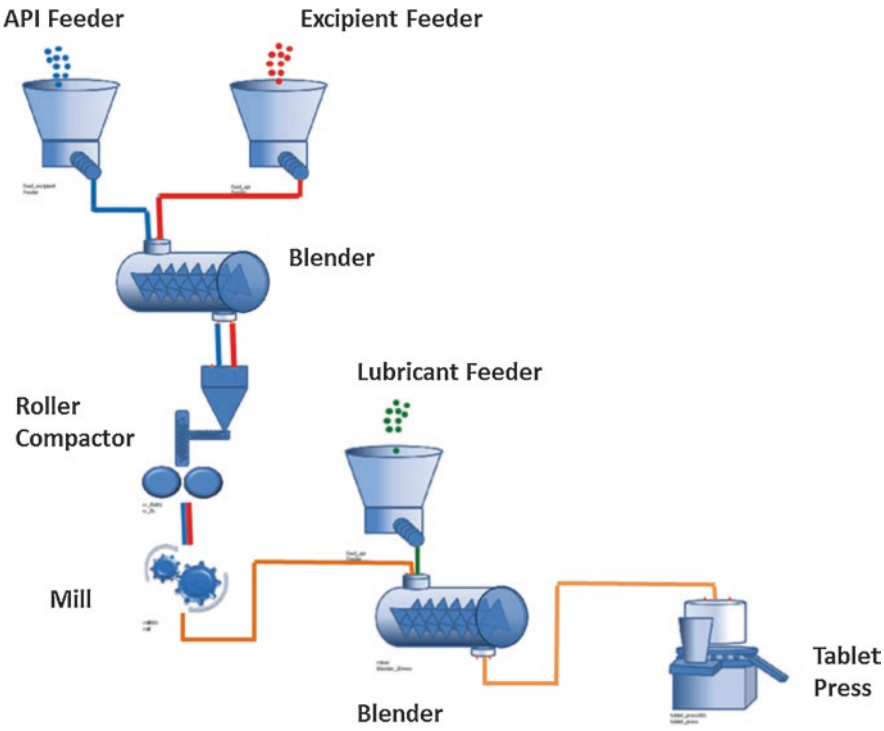


Fig. 12.1 Schematic of a dry granulation process

Fraction API
Desired/known safe ribbon properties (thickness, density, etc.)
Desired/known safe tablet settings (tablet weight, density, etc.)

Then a mass balance around each unit yields the following equations:

Required API feed rate = fraction API × overall process throughput
Required excipient feed rate = fraction excipient × overall process throughput
Blended powder rate = API rate + excipient rate
Ribbon rate = ribbon thickness × roll width × ribbon linear speed × ribbon density
Ribbon rate = blended powder rate = granule rate
Tablet rate = tablet weight × turret speed × number of stations = granule rate

Based on such a set of equations, it is clear what needs to be measured or determined. For example, if the feeders are loss-in-weight feeders with their own internal control logic, they only need a kg/hr setpoint, which mass balance provides. The tablet equation reduces to

$$\text{turret speed} \times \text{desired tablet weight} = \text{constant} \times \text{production rate}.$$

The roller compaction equation is more complicated, and knowing the roller geometry parameters reduces the equation to

$$\text{ribbon thickness} \times \text{ribbon density} = \text{constant} \times \text{production rate}.$$

with two unknowns. A further model equation is then required linking ribbon thickness or ribbon density or both to one of the machine parameters. Such a model may be proposed as

$$\text{ribbon density} = \text{some function of roll pressure and throughput}.$$

Note that the RTO exercise can be carried out with any appropriate model, whether data driven or mechanistic. Once all such equations are proposed in this manner, it remains to find the values of all the "constants" in the reduced model equations. This can be done by operating the corresponding unit such as the roller compactor or the tablet press and measuring throughput, adjusting individual setpoints manually to achieve a target flow rate with acceptable quality, and such a single data point can be used to scale all such flow ratio constants (*see* Note[1]).

The above method is an instance of static RTO with a dummy objective function; in other words, the result is feasible for throughput without necessarily being

[1] The goal of this exercise is to simply match production rates. It is not necessary to use the exact materials that are used for production unless flow characteristics are widely different; any cheaper substitute will serve perfectly well for such flow experiments. It is also not necessary to repeat this experiment at different flow rates unless the relationship between equipment settings and production rate (the transfer function) is significantly nonlinear. Most equipment is designed to operate linearly over a wide range.

optimal. The same principle can be used for any process of interest, using any models that are available. The key point is that nearly everything in the setpoint optimization can be derived from modeling except specific parameter values, which may either need individual unit operation experiments or simpler analysis of past data.

With such a method, as described above, it is possible to ensure that processes stay synchronized in production rate across setpoint changes. However, such an approach only cares about production quantity and not product quality. In particular, as a process is transitioning from one setpoint to another, a significant amount of material may be wasted in the transition. It is possible using dynamic setpoint optimization to minimize wasted material, ensuring that the product is of good quality even during transitions.

Dynamic Setpoint Optimization

Dynamic setpoint optimization requires more sophisticated mathematical models than static optimization does. In particular, the models used by dynamic optimization must be able to predict the behavior of the process over time, not just the final steady-state value. Mathematically, they are differential rather than purely algebraic. A wide variety of mathematical models can be used as long as they provide dynamic predictions.

Dynamic optimization can be modeled as an optimization problem wherein a performance measure like transition time is optimized subject to known initial conditions, known mass and energy balances over the transition, path constraints on the process, and the desired final process state. Such a problem can be addressed through both control vector parametrization and collocation.

Using the same dry granulation case study as earlier, we will show here in some detail how to model the amount of material involved in a setpoint transition, as well as how to minimize it. The amount of off-spec material lost in a transition can be quantified and minimized as follows:

$$\min \varphi = \int_0^{t_f} \delta(t) * m_{\text{ribbon}}\, dt \qquad \delta(t) = \begin{cases} 1; \sum_p \left(\dfrac{p(t) - p_{\text{d}}}{p_{\text{d}}} \right)^2 \geq \varepsilon \\ 0; \sum_p \left(\dfrac{p(t) - p_{\text{d}}}{p_{\text{d}}} \right)^2 \leq \varepsilon \end{cases}$$

where

m_{ribbon} = ribbon mass flow rate (kg/hour)
p = performance parameter
p_{d} = desired parameter specification
ε = relative tolerance

The performance parameters in the dry granulation process are then ribbon density and the concentration of API in the blended powder, with decision variables being roller compaction setpoints (feed screw speed, roll speed, and roll pressure) and the intervals of the setpoints. Constraints imposed on the optimization reflect

equipment limits (e.g., maximum and minimum RPMs) and operational limits (e.g., hopper levels, allowable range of ribbon properties).

Each unit operation can be modeled with any appropriate choice of model; the practical requirement is that the entire flowsheet must be simulated and optimized, so either the entire flowsheet must be created in a single program as a monolithic entity or the multiple models must be able to interchange data with each other to achieve convergence together. The result of the optimization would be the same, but the path may be very different.

The result indicates that an optimized start-up procedure starts upstream units early and at their final steady-state values and starts downstream units after appropriate delays (100 seconds for the roller compactor) and with gradual ramps to their steady-state values instead of abrupt changes. The consequence is that production is initially slow, but all the product made is on spec and can be used. Similarly, an optimized shutdown procedure times and ramps production rates such that intermediate hoppers are emptied steadily while final product quality is still maintained. Such an optimized strategy can then be implemented online.

Such a dynamic setpoint-changing exercise can also be carried out online, with actual measurements to validate models and make decisions. The trade-off between online and offline optimization is computational complexity: offline optimization will only be as good as its models and not necessarily reflect the latest state of the process (e.g., a new grade of material), but it is easy to build and run, and its results can be implemented in any RTPM system without any special requirement. Online optimization is much more adept at using actual and current plant data, but will need to retrieve data and communicate setpoints from the DCS or SCADA in real time, and makes validation much more challenging. In the context of pharmaceutical manufacturing, it is practical to use offline optimization over online optimization unless validation can be easily established, notwithstanding the notable admonition of Cutler and Perry (1983), who advocate the use of online setpoint optimization for its benefits.

12.3.2 Enforcing Setpoints

Enforcing setpoints is the domain of traditional controllers such as on/off controllers, servo controllers, proportional-integral-derivative (PID) controllers, advanced PID controllers with automatic tuning, and model-predictive control (MPC) schemes of widely varying complexity. They typically operate at a lower level than RTO systems; in other words, RTO systems calculate the setpoint that ought to be achieved at a given time, and PID or MPC controllers go about achieving it. The design of such a hierarchical system ensures that the PID layer is completely independent of the methods used by the RTO layer to calculate the desired setpoints, which greatly simplifies the design and maintenance of the RTPM system as a whole.

Specific guidelines based on experience with pharmaceutical manufacturing are as follows:

- Noise handling for solid streams: in most cases, process streams have significant noise compared to signal due to the nature of solid flow. Note that this is not sensor noise or electrical noise, which can be filtered away with low-pass or Kalman filters, but a fundamental noise from powder flowing as relatively large particles compared to fluids, which can be approximated as a continuum. Due to the often heavy-tailed nature of noise distributions in such solid streams, distinguishing signal from noise requires some thought. One way to handle it is to calculate finite differences at a high sampling rate, then smooth the derivative over a sliding window. Such high-frequency sampling followed by smoothing seeks to spot trends in signal quickly but damp out noise by smoothing.
- In general, on-off controllers are not preferred for solid handling except to actuate accept/reject gates in product streams. On-off controllers by their very nature are unsuitable for use in making small and gentle changes to product streams and are generally not worth the cost savings.
- PID is preferred where possible. PID can be natively implemented in the DCS level and does not require external software. With the specific case of solid handling, PI is recommended over PID. The advantage of using PI control over pure P control is that it can stabilize the system and eliminate offset, at the expense of a longer time to correct disturbances. PID controllers can act more quickly, acting on the derivative (forward looking) in addition to the area under the curve (backward looking), but finding the derivative on a property of a solid stream is difficult to do without introducing spurious noise or introducing a long delay to smooth out finite difference noise. In both cases, the rapid-action advantages of PID over PI are nullified.
- MPC and other advanced control schemes can be used where appropriate. Depending on the complexity of the specific model used in MPC, it can be implemented natively in the DCS layer or may require external software to implement. A good use for model-predictive control in pharmaceutical manufacturing is to resolve conflicts between interacting control loops. Such interacting control loops occur in complex unit operations like wet granulation, dry granulation, and hot-melt extrusion. A good example of MPC in dry granulation is given by Hsu et al. (2010a, b).
- Data reconciliation and gross error detection get more challenging due to the difficulty in distinguishing signal from noise. There is a wide body of literature in methods for data reconciliation and gross error detection, with work on mathematical methods by Narasimhan and Jordache (2000), Bagajewicz (2000, 1996), Bagajewicz and Jiang (1998), Crowe et al. (1983), Crowe (1986), Özyurt and Pike (2004), Narasimhan and Harikumar (1993), Mah et al. (1976), Johnston and Kramer (1995), Romagnoli and Stephanopoulos (1980), Kelly (2004), Liebman et al. (1992), and Albuquerque and Biegler (1996). Maquin et al. (2000) describe data reconciliation when the underlying models have uncertainty in them. A series of papers by Himmelblau and coworkers (Terry and Himmelblau

1993; Karjala and Himmelblau 1994; Himmelblau and Karjala 1996) introduce data-driven data reconciliation through the appropriate use of neural nets. Gross error detection has been described by Mah and Tamhane (1982), Narasimhan and Mah (1987, 1988), Chen et al. (1998), Chen and Romagnoli (1998), Tong and Crowe (1995), Rollins and Davis (1992), Romagnoli and Stephanopoulos (1981), Jiang et al. (1999), and Yang et al. (1995). Data-driven data reconciliation is entirely appropriate for pharmaceutical and particulate systems, as are data reconciliation approaches with significant model uncertainty. Simple polling or tie-breaking between multiple sensors is often enough in practice, but as confidence grows with particular instruments and sensors, more sophisticated techniques may be adopted.

In pharmaceutical manufacturing, detailed closed-loop control schemes have been proposed and described for a variety of processes, including direct compression, roller compaction, and wet granulation (Icten et al. 2015; Singh et al. 2013; Ramachandran et al. 2011; Giridhar et al. 2014; Singh et al. 2012, 2014).

12.3.3 Fault Detection, Diagnosis, and Mitigation

A third functionality of RTPM systems is fault diagnosis and correction, requiring first the ability to detect process faults by their measurable symptoms on the process, then to diagnose their root causes, and finally to mitigate the effect of such faults on product quality to the extent possible. There are a wide variety of mathematical and statistical techniques available to detect and/or diagnose faults. In all nontrivial cases, the mitigation strategy will involve changing the setpoints of one or more unit operations for the process. Thus, RTO serves as a setpoint-changing mechanism, as the execution end of fault mitigation.

Many methods have been developed for fault detection, diagnosis, and mitigation (Giridhar et al. 2014; Gupta et al. 2013; Hamdan et al. 2012; MacGregor et al. 2005). A detailed description of fault diagnosis methods is outside the scope of this chapter as it is a very large field. However, fault mitigation strategies all rely ultimately on changing setpoints in the process. The mechanism of RTO is the same, but the context changes: instead of seeking to, for example, maximize production rate, the objective in RTO driven by fault mitigation is often to eliminate the effects of the fault on downstream units, or reject material to prevent final product contamination, or maintain process safety. Some examples will serve to illustrate the points.

Consider a continuous direct compression line, with powders being blended continuously and fed to a tablet press. Let this process have online real-time measurement of flows and stream compositions. Upon detecting that the powders are not well blended going into the tablet press, two broad strategies suggest themselves to mitigate the effects of the fault: one is to divert the poorly blended material before it goes to the tablet press hopper, using a reject chute arrangement, and the other is to let the material flow as before but reject the tablets that correspond to the poorly

blended material, accounting for dispersion and mixing effects in the tablet press hopper. Neither strategy dominates over the other. Rejecting the off-spec material requires accept/reject chutes and pathways to be built in and triggered by the control system, and rejecting tablets corresponding to a certain time window can lead to larger waste of material, not to mention mechanical problems caused by the off-spec material itself. Diverting off-spec material from an upstream unit to a reject stream also means that flow rates to downstream units is reduced; hence, downstream units can operate only based on their hopper inventory. A suitable RTO strategy in this context could involve reducing the production rate of downstream units to make their hopper inventory last longer and therefore buy more time to fix the upstream problem. In a direct compression scenario, which is usually done with free-flowing material, it may simply be preferable to let everything flow to the tablet press and reject tablets based on suitable time windows. But if the same sort of fault were to happen in a dry granulation line, the trade-off calculations change. Granulation is normally done with sticky material that cause problems with direct compression; hence, sending known off-spec material downstream may cause bigger problems because of powder sticking to metal surfaces. In such a context, it may be worth examining an upstream reject stream more closely. The specific solution chosen will depend on practical aspects, including the relative costs of the material, the ease of operation, and simply the company culture in preferring a process management philosophy.

References

Albuquerque JS, Biegler LT. Data reconciliation and gross-error detection for dynamic systems. AICHE J. 1996;42:2841–56.

Ariyur KB, Krstic M. Real-time optimization by extremum-seeking control: Wiley-Interscience, Hoboken, New Jersey, 2003.

Bagajewicz MJ. On the probability distribution and reconciliation of process plant data. Comput Chem Eng. 1996;20:813–9.

Bagajewicz MJ. A brief review of recent developments in data reconciliation and gross error detection/estimation. Lat Am Appl Res. 2000;30:335–42.

Bagajewicz MJ, Jiang Q. Gross error modeling and detection in plant linear dynamic reconciliation. Comput Chem Eng. 1998;22:1789–809.

Chen J, Romagnoli JA. A strategy for simultaneous dynamic data reconciliation and outlier detection. Comput Chem Eng. 1998;22:559–62.

Chen J, Bandoni A, Romagnoli JA. Outlier detection in process plant data. Comput Chem Eng. 1998;22:641–6.

Crowe CM. Reconciliation of process flow rates by matrix projection part II: the nonlinear case. AICHE J. 1986;32:616–23.

Crowe CM, Garcia Campos YA, Hrymak A. Reconciliation of process flow rates by matrix projection part I: linear case. AICHE J. 1983;29:881–8.

Cutler CR, Perry RT. Real time optimization with multivariable control is required to maximize profits. Comput Chem Eng. 1983;7(5):663–7, ISSN 0098-1354. https://doi.org/10.1016/0098-1354(83)80010-6.

Dua V, Perkins JD, Pistikopoulos EN. In: Puigjaner L, Heyen G, editors. Real time optimization, in computer aided process and product engineering. Weinheim: Wiley-VCH Verlag GmbH; 2006. https://doi.org/10.1002/9783527619856.ch18.

Giridhar A, Gupta A, Louvier M, Joglekar G, Nagy Z, Reklaitis GV. Intelligent process management for continuous operations in pharmaceutical manufacturing. Comput Aided Chem Eng. 2014;33:391–6.

Gupta A, Giridhar A, Venkatasubramanian V, Reklaitis GV. Intelligent alarm management applied to continuous pharmaceutical tablet manufacturing: an integrated approach. Ind Eng Chem Res. 2013;52(35):12357–68. https://doi.org/10.1021/ie3035042.

Hamdan IM, Reklaitis GV, Venkatasubramanian V. Real-time exceptional events management for a partial continuous dry granulation line. J Pharm Innov. 2012;7:95. https://doi.org/10.1007/s12247-012-9138-6.

Himmelblau DM, Karjala TW. Rectification of data in a dynamic process using artificial neural networks. Comput Chem Eng. 1996;20:805–12.

Hsu S-H, Reklaitis GV, Venkatasubramanian V. Modeling and control of roller compaction for pharmaceutical manufacturing. Part I: process dynamics and control framework. J Pharm Innov. 2010a;5(1–2):14–23.

Hsu S-H, Reklaitis GV, Venkatasubramania V. Modeling and control of roller compaction for pharmaceutical manufacturing. Part II: control system design. J Pharm Innov. 2010b;5(1–2):24–36.

Içten E, Giridhar A, Taylor LS, Nagy ZK, Reklaitis GV. Dropwise additive manufacturing of pharmaceutical products for melt-based dosage forms. J Pharm Sci. 2015;104(5):1641–9. https://doi.org/10.1002/jps.24367.

Jiang Q, Sanchez M, Bagajewicz M. On the performance of principal component analysis in multiple gross error identification. Ind Eng Chem Res. 1999;38:2005–12.

Johnston LPM, Kramer MA. Maximum likelihood data rectification: steady-state systems. AICHE J. 1995;41:2415–26.

Karjala TW, Himmelblau DM. Dynamic data rectification by recurrent neural networks vs. traditional methods. AICHE J. 1994;40:1865–75.

Kelly JD. Techniques for solving industrial nonlinear data reconciliation problems. Comput Chem Eng. 2004;28:2837–43.

Liebman MJ, Edgar TF, Lasdon LS. Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques. Comput Chem Eng. 1992;16:963–86.

MacGregor JF, Honglu Yu, Muñoz SG, Flores-Cerrillo J. Data-based latent variable methods for process analysis, monitoring and control. Comput Chem Eng. 2005;29(6):1217–23. https://doi.org/10.1016/j.compchemeng.2005.02.007.

Mah RSH, Tamhane AC. Detection of gross errors in process data. AICHE J. 1982;28:828–30.

Mah RS, Stanley GM, Downing DM. Reconciliation and rectification of process flow and inventory data. Ind Eng Chem Proc Des Dev. 1976;15:175–83.

Maquin D, Adrot O, Ragot J. Data reconciliation with uncertain models. ISA Trans. 2000;39:35–45.

Marlin TE, Hrymak AN. Real-time operations optimization of continuous processes. In: AIChE symposium series. Vol 93. No. 316. New York: American Institute of Chemical Engineers, 1971-c2002; 1997.

Moritz Diehl H, Bock G, Schlöder JP, Findeisen R, Nagy Z, Allgöwer F. Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. J Process Control. 2002;12(4):577–85, ISSN 0959-1524. https://doi.org/10.1016/S0959-1524(01)00023-3.

Narasimhan S, Harikumar P. A method to incorporate bounds in data reconciliation and gross error detection - I. The bounded data reconciliation problem. Comput Chem Eng. 1993;17:1115–20.

Narasimhan S, Jordache C. Data reconciliation & gross error detection. Houston: Gulf Publishing Company; 2000.

Narasimhan S, Mah RSH. Generalized likelihood ratio method for gross error identification. AICHE J. 1987;33:1514–21.

Narasimhan S, Mah RSH. Generalized likelihood ratios for gross error identification in dynamic processes. AICHE J. 1988;34:1321–31.

Özyurt DB, Pike RW. Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes. Comput Chem Eng. 2004;28:381–402.

Ramachandran R, Arjunan J, Chaudhury A, Ierapetritou MG. Model-based control-loop performance of a continuous direct compaction process. J Pharm Innov. 2011;6(4):249–63.

Rollins DK, Davis JF. Unbiased estimation of gross errors in process measurements. AICHE J. 1992;38:563–72.

Romagnoli JA, Stephanopoulos G. On the rectification of measurement errors for complex chemical plants. Chem Eng Sci. 1980;35:1067–81.

Romagnoli JA, Stephanopoulos G. Rectification of process measurement data in the presence of gross errors. Chem Eng Sci. 1981;36:1849.

Singh R, Ierapetritou M, Ramachandran R. An engineering study on the enhanced control and operation of continuous manufacturing of pharmaceutical tablets via roller compaction. Int J Pharm. 2012;438(1):307–26.

Singh R, Ierapetritou M, Ramachandran R. System-wide hybrid MPC–PID control of a continuous pharmaceutical tablet manufacturing process via direct compaction. Eur J Pharm Biopharm. 2013;85(3, Part B):1164–82, ISSN 0939-6411. https://doi.org/10.1016/j.ejpb.2013.02.019.

Singh R, Barrasso D, Chaudhury A, Sen M, Ierapetritou M, Ramachandran R. Closed-loop feedback control of a continuous pharmaceutical tablet manufacturing process via wet granulation. J Pharm Innov. 2014;9(1):16–37.

Terry PA, Himmelblau DM. Data rectification and gross error detection in a steady-state process via artificial neural networks. Ind Eng Chem Res. 1993;32:3020–8.

Tong H, Crowe CM. Detection of gross errors in data reconciliation by principal component analysis. AICHE J. 1995;41:1712–22.

Yang Y, Ten R, Jao L. A study of gross error detection and data reconciliation in process industries. Comput Chem Eng. 1995;19:S217–22.

Chapter 13
Safety Guidelines for Continuous Processing

Martin D. Johnson and Jeffry Niemeier

Abstract This chapter gives safety guidelines for continuous chemistry experimental work in the laboratory. It also applies to continuous manufacturing of small-volume products using laboratory fume hoods. Pressure relief devices are needed on the discharge of positive displacement pumps. They are needed immediately downstream from back pressure regulators, at the inlet and outlet of plug flow reactors (PFRs), and on all pressure vessels that are not always open to an atmospheric pressure vent. Venting should be sufficient so vessels do not pressure up when filling, with vent paths free of block valves. Vent knockout vessels should be present and appropriately sized to catch process materials in case vessels overfill and also catch bubbler liquid in case of suck back. Plugging and fouling can be minimized by using a large enough tubing size, monitored with pressure transmitters, and managed with strategically placed valves and tees. Inerting is a primary line of defense against fire. Grounding and bonding is used when flowing from one vessel to another to prevent static charge buildup and sparking. Nonconductive liquid flowing through nonconductive tubing at high velocities should be avoided. Heat exchange rate should be sufficient for exotherm removal and chemical reaction safety, and chemical reaction safety analysis should be done, including calorimetry and thermal stability. Secondary containment should be provided, as well as sensors, to detect leaks and hazardous concentrations. Materials of construction must be compatible with the chemistry under reaction conditions. Automated alarms, interlocks, and/or autoshutoffs should be included based on temperatures, pressures, and fill levels. Emergency stops should be installed. Reactant accumulation or total mass accumulation; overfilling product collection vessels; undesired phase changes; backflow into feed vessels, vents, or utilities; and static electricity buildup should all be prevented.

Keywords Continuous processing · Drug substance · Process safety · Pressure relief · Reaction safety

M. D. Johnson (✉) · J. Niemeier
Eli Lilly and Company, Indianapolis, IN, USA
e-mail: johnson_martin_d@lilly.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_13

Abbreviations

A/V	Surface area per unit volume
API	American Petroleum Institute
ASME	American Society of Mechanical Engineers
BPVC	Boiler and Pressure Vessel Code
CCPS	Center for Chemical Process Safety
CSTR	Continuous stirred-tank reactor
DIERS	Design Institute for Emergency Relief Systems
GC	Gas chromatography
LOC	Limiting oxygen concentration
MAWP	Maximum allowable working pressure
MOC	Materials of construction
MSMPR	Mixed suspension mixed product removal
P&ID	Process and instrumentation diagram
PFR	Plug flow reactor
PI	Pressure indicator
PPE	Personal protective equipment
PT	Pressure transmitter
SAChE	Safety and Chemical Engineering Education program
SURF	Scale-Up Review Form

13.1 Introduction

Before we look at the unique safety considerations for continuous processes versus traditional batch processes, let us look at some similarities. The same risks can exist for continuous and batch operations, including the potential for fires, explosions, runaway reactions, and toxic chemical exposures and environmental releases. All technical aspects of process safety assessment apply to both continuous and batch systems, including the need to assess reactivity, kinetics, thermodynamics, materials of construction, heat generation and removal rates, mass transfer rates, off-gassing, industrial hygiene, engineering controls, and by-product quantities and hazards. Chemical releases can occur in the same ways: equipment ruptures and failures, releases from pressure relief devices, and spills and inadvertent releases of gases. The principles used to design inherently safer processes still apply: minimize, substitute, moderate, simplify, limit effects (Kletz 1978; Crowl and Louvar 2012; CCPS 2009, 2012). Despite these similarities, continuous processes can offer a unique set of safety benefits and challenges, as described in this chapter.

13.2 Safety Advantages of Continuous Processes

Continuous processing offers safety advantages compared to batch processing (Baxendale 2013; Hessel et al. 2013; Webb and Jamison 2010; Yoshida et al. 2011; Gutmann et al. 2015), including the following:

- Decreased risk of significant processing events due to smaller equipment volumes.
- Decreased risk of forming a flammable vapor–air mixture inside the reactor in cases when a plug flow reactor (PFR) can be run 100% liquid full. Elimination of headspace in a PFR also eliminates the potential for volatile explosives, like hydrazoic acid, to partition into a headspace.
- Decreased potential for reactant or by-product accumulation when the reaction and quench are run continuously in series.
- Decreased amount of hazardous reagent gas (e.g., H_2 or CO) in a PFR that runs mostly liquid filled. A PFR with H_2 or CO reagent can operate outside the building because it is always sealed.
- Improved heat removal due to higher heat transfer surface area per unit volume (A/V).
- Decreased chance of worker exposure. Some commercial scale reactors are small enough that they can be operated in fume hoods (Johnson et al. 2020).
- Increased potential to telescope reactions and process steps and thereby reduce inventories of hazardous materials.

In many cases, the most atom-efficient synthetic routes would require high-energy chemical transformation or extreme driving conditions, such as high temperatures and pressures (Gutmann et al. 2015). Throughout the history of pharmaceutical production, not many chemistry types have been practical for a scale-up to manufacturing in standard multiuse batch equipment. However, with the advances of continuous processing technologies and their associated safety advantages compared to batch processing technologies, the previously impractical reactions, reaction types, or reaction conditions are now becoming available for pharmaceutical manufacturing. A comprehensive review of hazardous chemistries made safer in continuous mode versus batch mode was provided by Gutmann, Cantillo, and Kappe (2015). As described in that review paper, examples for oxidation include aerobic oxidation with molecular oxygen in packed catalyst bed PFRs (Obermayer et al. 2013) or with dissolved catalyst (Ye et al. 2010) and oxidation with ozone in microreactors (Wada et al. 2006) or CSTRs (Allian et al. 2010). Exothermic nitration reactions have been accomplished industrially on a large scale in parallel microreactors (Braune et al. 2009). Diazomethane was continuously produced and continuously consumed at a 60-ton-per-year scale, with maximum instantaneous inventory only at 80 g, at Phoenix Chemicals (Proctor and Warr 2002). An in situ continuous diazotization and consumption of the diazo compound, in series with continuous in situ formation and consumption of hydrazine in a three-step telescoped synthesis of n-aryl pyrazoles, was published by Pfizer researchers (Li et al. 2012). Here, continuous

reactions in series minimized the accumulation of diazonium salt and hydrazine intermediates. Sodium azide has been used to make organic azides in PFR tube reactors and microreactors with no headspace, therefore keeping hydrazoic acid in solution until it reacted or is destroyed in situ (Kopach et al. 2009; Zhang et al. 2014; Sahoo et al. 2007). While it may be safer to run many chemical reactions continuously instead of batchwise, safety hazards still exist. More importantly, continuous processing introduces certain new safety hazards compared to batch processing. Temperatures and pressures are often higher, and equipment is more subject to line plugging (e.g., in plug flow reactors) and backflow to other parts of the process train or utilities. More hazardous process materials and chemistries are often selected since small processing scale provides a measure of safety. The risk of an accident is higher during process start-up and shutdown, as opposed to a steady state.

13.3 Safety Checklist: Useful for Prestartup Safety Reviews

The following list gives safety guidelines for continuous processing. This list presumes the existence of basic safety provisions (e.g., safety data sheets (SDSs), PPE, emergency procedures, and facility-related safety items such as eyewashes, safety showers, alarms, sprinkler systems, fire extinguishers, fume hoods with sufficient linear velocity at the sash, and fume hood alarms).

Pressure Testing and Relief Devices

- Include pressure relief devices and/or auto-shutoff pressures on the discharge of pumps (especially positive displacement pumps). Pump reliefs should be set lower than downstream vessel reliefs.
- Provide pressure relief devices immediately downstream from back pressure regulators (lower pressure set point to protect downstream equipment).
- Consider installing pressure relief devices at the inlet and outlet of PFRs in case they plug in the middle, if this is a scale-up of a long PFR with a small i.d.
- Make sure to include pressure relief devices on all pressure vessels that are not always open to an atmospheric pressure vent.
- All pressure reliefs must relieve to a safe location, and all must be set to lower than the maximum allowable working pressure of the equipment.
- Make sure that there are no valves in the pressure relief vent line.
- If any equipment will operate under positive pressure, pressure test at 120% of the desired operating conditions at the desired operating temperature. The system should lose less than 5% pressure in 1 h.
- If you are running a reactor at extreme temperatures, for example, 300 °C or −78 °C, then pressure test again after heat/cool cycles to make sure that the thermal expansions and contractions did not cause new leaks in the compression fittings.
- Stay below the pressure and temperature limits of the equipment. Remember that pressure rating decreases with temperature.

- If you have a jacketed vessel like a flask or condenser and if the heat transfer fluid can be blocked in by valves or quick-connect fittings, then the jacket needs a pressure relief.

Venting

- Ensure sufficient venting so vessels do not pressure up when filling (with vent paths free of block valves). Vent lines should be larger in diameter than inlet lines.
- Use separate vent headers when multiple vessels have incompatible headspaces.
- Provide vent knockout tanks that are appropriately sized to catch process materials in case vessels overfill and also catch bubbler liquid in case of suck back.
- Make sure that all organic vapors are condensed and collected in product vessels prior to venting.
- Ensure that vent knockout vessels are monitored. It is best to monitor them automatically with alarmed level detectors or weigh scales, but they can be monitored manually if the frequency is sufficient to a point that it is not possible for the knockout vessel to completely fill and overflow.
- Provide a vent bubbler that keeps air from diffusing back into the vent header and make sure that the liquid in the bubbler is compatible with the process.
- Minimize venting of hazardous substances and use vent scrubbers where needed.
- Dilute vent gases to below flammability limits (e.g., venting $H_2 < 4\%$).
- Be especially careful with separation of liquids and vapors when a reactor is depressurizing and the gas/liquid mixture is flowing into a vessel and gas is exiting to a vent at the same time at large scales (e.g., use of hydrocyclones to prevent misting out the vent for >1 L/min vent gases).
- Vent down pressurized reactors slowly, especially if using a gas like H_2 or CO.

Leaks and Spills

- When reviewing an equipment setup, particularly for large-scale applications, do a risk assessment for the potential of spills (e.g., manual valves left open/pump failures/automation errors). Ensure that appropriate precautions are put in place to prevent spills.
- New operations using automation recipes should be evaluated with personnel oversight prior to execution/use outside of working hours.
- Avoid manual sampling valves in process streams or waste lines that could accidentally be left open. Choose an automated sampler instead. However, if you are using manual valves, then use an overflow tee with two block valves in series and gravity drain, such that both valves are not opened at the same time.
- Use automated bottom valves that can be interlocked to shut off the supply from feed vessels. For example, the interlock can be set to close the block valve if the weight of the feed vessel on floor scale is decreasing faster than the maximum allowable value.
- If you are running an experiment in a lab hood that does not have a floor drain and if you are using a pressurized feed vessel with flammable solvents, then it is best if the liquid volume in the feed vessel is less than the secondary containment capacity in the lab hood.

Plugging/Fouling

- Use a large enough tubing size to prevent plugging if you suspect that you could have solids. For example, do not use smaller than a ¼″ tubing in research-scale continuous crystallization and use high-velocity intermittent flow. What are the smallest flow restrictions for process materials, for example, needle valves? These will plug first.
- Provide monitoring for plugging and fouling visually and with pressure transmitters. Measurement of pressure versus time at pump outlets provides indication of gradual fouling.
- Provide mitigations for plugging and fouling (e.g., double block and bleed valves) so that you will not need to use a wrench on a clogged line under pressure.
- Be aware that the process tubing side of a pressure transmitter can plug with solids and make the reading inaccurate.

Static Electricity

- Make sure that process vessels, pumps, and all electrically powered equipment are properly grounded. Extra caution is needed when humidity is low.
- Inerting is the primary line of defense against fire. Make sure that all vessels are inerted before flammable solvents flow into the vessel, especially if they flow into the vessel through nonconducting tubing.
- If you are using an automated sampling and dilution cart for online high-performance liquid chromatography (HPLC) or GC, then make sure that the waste is pushing into an inerted vessel and not straight to the waste solvent drop.
- Use grounding and bonding when flowing from one vessel to another to prevent static charge buildup and sparking. If flowing into a nonconductive vessel, ground via a metal dip tube.
- If you intend to use nonconductive heat transfer fluid, for example, Galden, Syltherm, Paratherm, then use conductive piping for your heat transfer fluid pump-around, for example, stainless steel tubing or conductive impregnated perfluoroalkoxy (PFA). Otherwise, have a specific safety review for this.
- If you intend to use nonconductive piping in the heat transfer fluid pump-around loops, for example, Teflon or a PFA tubing, then use water or other conductive heat transfer fluids. Otherwise, have a specific safety review for this.

Heat Removal and Chemical Reaction Safety

- Provide sufficient heat exchange for process liquid flowing into and out of reactors or separators and for heat of reaction and other sources of energy. Make sure that cooling fluid cannot overpressurize or boil out from the jacket.
- Complete chemical reaction safety analysis. This includes accelerated rate calorimetry (ARC) and differential scanning calorimetry (DSC) and calculation of heat removal.
- Provide appropriate material compatibility, flammability, reactivity, reaction calorimetry, and thermal stability test data.
- Know the reactor surface area and heat transfer coefficient.
- Stay below the temperature limits for the reaction solvent.
- Provide calculations of worst case scenario heat release for exothermic reactions

- If running in CSTRs, make sure that the reaction is dilute enough so that all the solvent cannot boil off in an exothermic event, for example, a Grignard formation.
- Complete the Scale Up Review Form (SURF) if this is a scale-up.
- Test the stability and solubility of starting reagent solutions over time before running flow.

Safety Review Meetings

- Hold prestartup safety reviews to ensure that equipment and controls were installed as designed.
- Have prestartup safety meetings to communicate critical safety issues to the operation staff, if the people running the experiment were not all present when going through the safety checklist.

Facility

- A safe location should be used for the experiment or production run considering the number of people affected by a worst-case scenario.
- Have a plan for waste disposal before the start of the experiment or production run. Test waste streams for pH, especially if the waste stream flows directly to waste, and wet the pH paper first if the waste is mostly organic solvent.
- Provide chemical sensors to detect leaks and hazardous concentrations (e.g., hydrogen detectors in the tops of hoods where H_2 is used).
- Plan for what will happen in the event of an evacuation if you are kept out of the area for a time.
- If you are running flow chemistry in a lab, what is the maximum amount of flammable solvent allowed in that lab at one time, and is your process under the limit?
- Ensure that a fire extinguisher is present that is suitable for the process you are running (e.g., metal-rated fire extinguisher for lithiations or Grignard formations)

Equipment

- Make sure that you have a drawing or process and instrumentation diagram of the system, and make sure that the drawing matches the actual equipment, instrumentation, and piping for the entire process.
- Make sure that all manual valves are labeled and that the labels on the equipment match the labels on the drawings and that the procedure identifies what to do with each of the manual valves.
- If they are connected to a flexible tubing, mount valves solidly on brackets so that the tubing is not strained when turning.
- Make sure that valves are not mounted directly on glassware nozzles, especially if they are automated valves. When the actuator turns the valve, it puts stress on the glassware and it will eventually break the glass.
- Make sure that all manual block valves have been adjusted to the correct position for startup.
- Make sure you know and understand all your process lines for all flows in and out.
- Provide secondary containment and use netting or coating on vacuum glassware.
- Consider the special safety hazards during startup and shutdown transitions.

- Consider physical protection (shielding) of process lines or equipment from people kicking, bumping, or leaning.
- Provide redundancy for measurement of key process parameters; measurement devices can fail for any number of reasons; examples include plugging of a pressure transmitter with solids and coating of a thermocouple with residue. Use redundant measures for temperature (e.g., thermocouples in process and jacket or circulator), pressure (e.g., pressure transmitter (PT), pressure indicator (PI), supply gas cylinder regulator), mass flow (e.g., calibrated pump, mass flow meter, catch and weight, change in balance weight over time).
- All electrical equipment must be grounded and safety tested by electricians. Keep electrical cords off surfaces.

Materials of Construction

- Make sure that vessels and tubing are compatible with the chemistry at operating temperature and pressure to prevent corrosion or ruptures.
- Make sure that pumps are compatible with corrosive or abrasive feeds if a leak can cause a safety incident. For example, if BuLi contains solid salts that cause pump seals to leak, then the BuLi leak could cause a fire.
- Consider all tubing, fittings, transmitters, valves, O-rings, and gaskets for compatibility with the solvents at process temperatures and pH of solutions.

Automation and Emergency Stops

- Include alarms, interlocks, and/or auto-shutoffs based on temperatures, pressures, and fill levels where needed.
- Provide measurement of real-time mass and energy balances when necessary (e.g., scale-up of continuous Grignard formation reaction in a continuous stirred tank reactor (CSTR)).
- Provide emergency stops for power to the lab hood in case of building evacuation and one-button stops for the automation that would be used at the same time.

Heating

- Provide safe heating if using circulators. Do not heat higher than flash point. Inert the headspace of constant temperature baths if heat transfer fluid is flammable. Elevate the circulator above the process vessel if the vessel heating/cooling jacket volume is large. Otherwise, if the circulator is placed at a lower elevation than the vessel heating/cooling jacket, then gravity could cause the heat transfer fluid to leak out the circulator reservoir when it is not running, if it is not completely sealed.
- Make sure that quick-connect fittings in heat transfer fluid tubes are connected properly. Sometimes if these are not completely connected, it can greatly restrict the circulator flow.
- Understand the hottest possible surface temperature if using electrical heaters and make sure that this will not be a problem.
- Protect workers from accidental burns from touching a hot surface.
- Provide mitigations for thermal expansion, including thermal expansion in closed ball valves. Allow for thermal expansion of a metal-coiled tubing reactor.

Solvent-Only Run Prior to the Real Chemistry Run

- Use the exact same flows, temperatures, and pressures that will be in the real run.
- Test written startup, shutdown, and operating procedures. The procedure can be the description written in the electronic lab notebook.
- Test interlocks and/or auto-shutoffs and make sure that everyone involved in this run knows all of the interlocks in the automation.
- Close mass balances; make sure that there were no leaks. Verify correct temperature and pressure control and correct mass flow rates. You should achieve at least 95% mass balance on solvent run before real chemistry run. Perform catch-and-weigh testing to prove that your mass flows are correct.[1]
- Ensure that you will have safe and representative sampling of flowing streams, preferably from lower pressure parts of the system.

Gas Cylinders

- Mount compressed gas cylinders properly, regulate lower than downstream limits, use reliefs, close valve when not in use, and cap unconnected cylinders.
- Restrict flow if using hazardous gas for continuous reaction. Consider adding a metering valve immediately downstream from cylinder.
- Use the smallest cylinder available and feasible for the experiment, if using hazardous reagent gas like hydrogen. For example, use the D size (2.2 L water volume) hydrogen cylinder in a lab hood.
- Use automated block valve and programmed interlocks to shut off supply.
- Filter the gas from the cylinder to prevent fouling of metering valves or ruining automated block valves.

Prevent These Things

- Prevent reactant accumulation or total mass accumulation. Make sure that vessels are not gradually filling if they are supposed to be at a constant level.
- Prevent overfilling vessels. This especially applies to product collection vessels.
- Prevent undesired phase changes in process lines like boiling or freezing.
- Prevent plugging of vent lines.
- Prevent process materials from flowing out of vents.
- Prevent backflow into feed lines or feed vessels. Check valves can only be a backup line of defense, not the primary defense for backflow prevention, because check valves fail. Prefill the pumps and prefill the feed lines all the way to the tee or mixer. Make sure that peristaltic pumps are pumping in the correct direction.
- Prevent backflow of liquid from vent lines, such as backflow of caustic from scrubbers.
- Prevent backflow into utility lines (e.g., nitrogen). If you are running a high-pressure process, then only use cylinder nitrogen, not house supply nitrogen.
- Prevent closed block valves in vent lines.
- Prevent process liquid getting into vacuum lines. Use vacuum knockout pot.

[1] "Catch-and-weigh" means to collect the pumping liquid for a measured time period in a tared container or graduated cylinder. If collecting in a tared container, then measure the mass of liquid pumped. Measured mass per time is the actual pumping rate.

- Prevent sparks due to static electricity discharges by using grounding and bonding.
- Prevent flow of nonconducting solvents or fluids in Teflon or PFA tubing at linear velocities higher than 1 m/s.

Manufacturing facilities typically have thorough and detailed safety guidelines and process hazard reviews. Laboratories sometimes have less rigorous and less formalized procedures because of the smaller material amounts and smaller chemical and physical potentials. This guideline document can help fill the gap for the lab.

13.4 Laboratory Fume Hoods

The decision to not electrically classify the interior of the laboratory fume hoods is not inconsistent with NFPA 45, standard on fire protection for laboratories using chemicals, or NFPA 70, national electric code. Most organization labs make the decision to not classify their lab fume hoods. However, this does not mean that fire hazards are nonexistent in laboratory fume hoods. Even when design face velocities are met (85–115 ft/min), there exists the possibility of a flammable solvent/air mixture in the hood. Therefore, design specifications, policies, and procedures should be in place to help minimize the possibility of a flammable mixture coming into contact with an ignition source. In addition to the safety checklist just described, the following guidelines are helpful for minimizing the possibility of fire or the impact of a fire inside the lab hood.

- Electrical receptacles, switches, and controls should be located so as not to be subject to liquid spills.
- A master emergency stop switch or button should be mounted on the outside of the hood to turn off all power to the receptacle strip in the hood. Receptacle strips should be mounted at least 6.5 inches from the top of the bench or 42 inches from the floor in a walk-in hood.
- Minimize ignition sources in the hood at all times as much as possible.
- Keep the electrical cords off the surfaces.
- Elevate electrical equipment on lab jacks providing better flow around the equipment and reducing the potential for high vapor concentrations.
- Place solvent cooling baths and secondary containment in inert headspaces of solvent cooling baths as much as possible.
- Unplug electrical cords or only when equipment is off, or the power to the receptacle strip has been turned off, to eliminate the possibility of a spark when the plug is pulled.
- Inspect the condition of electrical cords and equipment in the hood.
- Keep the sash closed as much as possible.
- Do not store solvents in the hood.
- Include and maintain an audible or a visible alarm indicating loss of airflow in the hood for low face velocity.

- Respond to low airflow alarms as soon as possible. If this is not remediated immediately, then lab personnel should push the master electrical switch in the event of a loss of airflow alarm when conducting experiments using solvent that is open to the air, for example, during open filtration.
- Use a fire suppression system in the fume hood if particularly hazardous or flammable chemistries are done, for example, lithium aluminum hydride (LAH) or sodium hydride (NaH) reductions, lithiations, or deprotonations with n-butyllithium (n-BuLi) or lithium diisopropyl amide (LDA), or dry catalyst handling.
- Use alarmed gas detectors in the upper corners of the fume hood if using flammable gasses line H_2 or CO.

13.5 Process Documentation and Communication

Process documentation provides an invaluable tool for conducting hazard reviews and sharing knowledge with current and future operating personnel. The documentation needs will depend on the scale, duration, and hazards of experiments and may include the following:

- Diagrams
- Process descriptions
- Operating instructions
- Safe operating limits
- List of safety-critical equipment
- Emergency procedures
- Material safety data sheets (MSDSs)
- Corrosivity data
- Thermal stability data
- Reactivity data
- Emergency relief system design basis and design calculations
- List of interlocks

Process changes necessitate updates to process documentation. Operators should be knowledgeable of process safety documentation, especially emergency procedures. Critical valves should be clearly marked in the hood and on diagrams.

13.6 Pressure Relief

This section contains a brief overview of pressure relief for continuous laboratory process equipment. A properly designed, installed, and maintained pressure relief system will protect the equipment from damage and rupture and protect personnel from injury. Such systems are the last line of defense for overpressure scenarios.

A pressure relief device and/or an auto-shutoff switch is needed on pumps that are capable of exceeding the pressure rating of downstream equipment (positive displacement pumps are capable of generating pressures). Pump relief set pressures should be lower than downstream vessel relief set pressures. This relieves pressure near its source and avoids pressure drop that would occur if downstream relief devices are utilized. If material is released at the pump, then it would normally be at ambient temperature, but if material is released from the reactor, then it could be at elevated temperature.

Pressure relief devices are also needed on vessels, including reactors and feed and product tanks. Exceptions to this rule can be made. For example, pressure relief may not be necessary on a microtube reactor with a volume less than 1 mL if the volume of the interconnected system is small, the pump has pressure relief, the equipment is in a lab hood, and secondary containment is provided. The default procedure should be to use pressure relief devices on continuous reactors, and if they are not utilized, then justification should be given and approved. Pressure vessels that fall under the American Society of Mechanical Engineers (ASME) Boiler and Pressure Vessel Code (BPVC) are required to have an ASME-certified relief device. Many jurisdictions have adopted the ASME BPVC as law, and in some cases they modify the definition of a regulated pressure vessel.

Nonpressure vessels must have an open vent line at all times, typically inerted with nitrogen flowing out through a vent knockout pot and a vent bubbler. It is important to consider safe venting of product collection vessels as they are often overlooked.

A pressure relief device should be installed immediately after a back pressure regulator to protect downstream equipment. Such regulators are typically used to keep a continuous reactor at constant high pressure, while the downstream collection tank(s) or downstream unit operations run at lower pressure. The process lines on the upstream side are typically rated for high pressure, as opposed to downstream lines and equipment. Pressure relief is needed in case the process line plugs with solids downstream from the back pressure regulator, or there is a manual block valve mistakenly closed in the low-pressure section, because the tubing pressure downstream from a back pressure regulator has the potential to increase to the same pressure upstream from the regulator.

Two key design considerations for pressure relief systems are as follows:

- *The relief device set pressure*: it should be as low as practical[2] and no higher than the design pressure or maximum allowable working pressure (MAWP)[3] of the lowest rated equipment component being protected.

[2] A low set pressure is especially advantageous for runaway reaction scenarios, as it provides a better chance of turning around the pressure rise before the maximum allowable pressure is reached.

[3] Piping components are assigned a design pressure, while ASME pressure vessels are assigned an MAWP; ASME pressure vessels also have a design pressure, but this may be different from the final assigned MAWP.

- *The relief system capacity*: capacity must be sufficient to prevent the pressure from significantly exceeding the design pressure or MAWP of the equipment. The amount by which piping systems and ASME pressure vessels can exceed these pressures is defined by ASME B31.3 and the ASME Boiler and Pressure Vessel Code (BPVC), respectively.[4]

When designing a system, appropriate safety margins are needed between operating pressures and relief device set pressures. Operating too close to the set (burst) pressure of a rupture disk can eventually cause it to fail below its set pressure. Operating too close to the set pressure of a relief valve may cause it to begin to leak slowly.

Equipment pressure ratings and relief device opening pressures are temperature dependent. Equipment must operate within acceptable temperature ranges during both operating and upset conditions.

The relief system must have sufficient capacity for all scenarios deemed credible. Scenarios that may be credible in laboratory equipment include exposure to external fire, pump overpressure due to closed downstream valves, overpressure from a utility service (e.g., nitrogen cylinder), hydraulic expansion due to heating of trapped liquid, abnormal heat input, automatic control failure, and runaway reaction. API 521 provides guidance on scenarios to consider.[5] API 520, Part 1, provides guidance on relief device sizing methods.[6] Some sizing calculations are relatively straightforward, while others require significant expertise (e.g., runaway reaction scenarios).

Rupture disks and relief valves are the most commonly used types of pressure relief devices. Table 13.1 provides a comparison of the devices. Pressure relief valves are commonly used for continuous reactors in laboratory fume hoods because the volume of release would typically be small and the valve reseats.

The effect of normal superimposed back pressure must be considered since the actual opening pressure of most relief devices depends on the differential pressure between the device inlet and outlet.

Another type of pressure relief device is thermal plugs. These melt at high temperatures, providing pressure relief for external fire scenarios, and are most often used on cylinders.

Pressure relief effluent steams must be handled in a manner consistent with toxicity and flammability hazards. The need for effluent control equipment (e.g., catch

[4] ASME (American Society of Mechanical Engineers) B31.3 Code for Pressure Piping provides guidance on the design of piping systems. Section VIII of the ASME Boiler and Pressure Vessel Code is commonly used to design and test pressure vessels. For equipment designed to the ASME BPVC, the allowable accumulated pressure (pressure above the MAWP) is generally 10% for non-fire scenarios, and 21% for fire scenarios.

[5] API 521 is American Petroleum Institute Standard 521 on Pressure-Relieving and Depressuring Systems.

[6] API 520 Part 1 is American Petroleum Institute Standard 520 Part 1 on Sizing, Selection, and Installation of Pressure-Relieving Devices in Refineries. Its guidance is widely applied in the process industries.

Table 13.1 Comparison of rupture disks and relief valves

	Rupture disks	Relief valves
Reclosing characteristics	They do not reclose, allowing the release of material to continue.	They reclose when pressure drops below the blowdown pressure.
Maintenance	They require periodic inspection and/or replacement.[a]	They require periodic inspection, bench testing, and/or replacement.[a]
Ability to handle streams prone to plugging or formation of deposits	They are better suited to handle such streams.	They may stick shut and not provide overpressure protection.
Ability to handle highly corrosive streams	They can be supplied in a variety of exotic materials.	The use of exotic materials may be prohibitively expensive.
Ability to handle relief line pressure drops	No special considerations—the system is designed considering overall pressure drop, including inlet and outlet relief lines.	Excessive inlet or outlet line pressure drop will reduce the capacity of the valve and cause it to chatter, possibly leading to valve failure and loss of pressure relief.[b] Consult with the manufacturer for pressure drop limits.[c]

[a]See API 576 (American Petroleum Institute Standard for Inspection of Pressure Relieving Devices)

[b]Pilot-operated relief valves can be used if inlet line pressure drop is an issue; balanced bellows relief valves can be used if outlet line pressure drop is an issue

[c]The general guideline for conventional relief valves is to limit nonrecoverable inlet pressure drop to less than 3% of set pressure, and outlet line built-up back pressure to less than 10% of set pressure. See ASME BPVC Appendix M

tanks, cyclones, scrubbers), as well as the selection of the ultimate venting location, should be considered.

Pressure relief lines should not have block valves. Special circumstances could be considered if a block valve is necessary to allow the replacement of a relief device during process operation. The ASME BPVC contains extensive guidance on measures necessary to ensure that such valves do not increase the risk of equipment overpressurization.

The nominal pipe size of all piping, valves, fittings, and vessel components between a pressure vessel and its pressure relief valves shall be at least as large as the nominal size of the relief valve inlet. Likewise, the nominal pipe size of all piping, valves, and fittings on the discharge of relief valves shall be at least as large as the nominal size of the relief valve outlet. Pressure relief outlet lines should be as short and straight as possible. Bends and piping length reduce relief system capacity.

The overall system should be reviewed to determine which relief device(s) will relieve first for the various credible scenarios. Relief device set pressures should be selected so that lower hazard materials are vented preferentially.

The experimental apparatus should be pressure and leak tested at a pressure of 1.2× to 1.5× the maximum planned operating pressure after it is completely installed and before running process chemistry. Testing requirements will vary depending on

the design of the equipment. For example, the ASME Boiler and Pressure Vessel Code has specific requirements for vessels designed to that code. The system can be pressurized to its operating pressure using an inert gas such as nitrogen, then pressure versus time is observed. Observed pressure loss must not exceed a certain rate, for example, 1% per hour, after temperatures stabilize at controlled value. Increasing pressure suddenly causes an increase in temperature, and when the temperature drops back to ambient, the pressure will drop. This alone can cause a significant pressure drop, which is why the pressure versus time leak rate check should not start until temperatures stabilize. Leak detectors with audible alarms should be considered when handling hazardous gases such as H_2 and CO. Leak detectors must take into account the density of the gas (heavier or lighter than air). For example, mount H_2 leak detectors near the ceiling because it is lighter than air.

An incident that causes activation of a relief device will likely necessitate follow-up actions. If the pressure exceeds the yield strength of the equipment, it will be necessary to replace the equipment. Relief valves may need to be removed, cleaned, and bench tested (or simply replaced), and discharge lines may need to be flushed to remove any residual material.

Table 13.2 lists organizations that provide information relevant to pressure relief.

13.7 Venting

Vent lines from all process vessels and product receivers should be less restricting (i.e., have a larger inside diameter) than the inlet lines. This minimizes the possibility of vessel pressurization as fluids enter continuously or intermittently.

Vent bubblers are used to maintain an inert gas environment. Check valves are sometimes used for maintaining an inert headspace in process vessels by preventing air suck back from the outside, but vent bubblers are preferred. Check valves can plug, clog, or otherwise fail to allow adequate flow. Even when operating properly, they restrict flow more than open tubing. The goal is to have no flow restrictions in the vent line. If a vent bubbler is used, then a vent knockout vessel should be installed between the process and a vent bubbler so that the liquid from the vent bubbler cannot be sucked back into the process vessels. The liquid in the vent bubbler must be compatible with the chemistry. The vent tube piped between the knockout pot and the bubbler should be routed up as high as possible in the hood and then back down into the bubbler, and both the vent knockout pot and bubbler should be located at as low an elevation as possible in the hood. This arrangement helps to prevent bubbler suck back into the vent knockout pot in the unintended event that the system temporarily goes under slight negative pressure. For instance, the system could temporarily have slight negative pressure when a material is pumped or sucked out of a process vessel. In order for liquid to be sucked back from the bubbler, it must overcome gravity, thus keeping the system inert even if process vessels are under slight negative pressure. For example, if the tube between the vent knockout pot and the bubbler goes up 6 ft elevation in the hood and then back down 6 ft

Table 13.2 Some pressure relief information resources

Organization, Program, or Act	Comments
API (American Petroleum Institute)	It provides standards for design, installation, and maintenance of pressure relief systems; pertinent standards include API 520, 521, and 576.
ASME (American Society of Mechanical Engineers) B31.3	It provides guidance for the design of process piping; some jurisdictions have adopted this as law.
ASME (American Society of Mechanical Engineers) Boiler and Pressure Vessel Code	Section VIII of the code contains requirements for ASME pressure vessels and pressure relief for such vessels; it is often adopted as law in many jurisdictions (in whole or in part); applicability of the code depends on equipment size and operating pressure.
CCPS (Center for Chemical Process Safety)	The book Guidelines for Pressure Relief and Effluent Handling Systems (American Institute of Chemical Engineers 1998) provides design guidance and interpretation of codes and standards.
DIERS (Design Institute for Emergency Relief Systems)	The DIERS Project Manual (Fisher et al. 1993) provides in-depth guidance on pressure relief issues, including two-phase flow; the DIERS User's Group provides a forum for continual improvement of pressure safety understanding and technology.
SAChE (Safety and Chemical Engineering Education) program	It provides an overview of pressure relief issues.
US OSHA (Occupational Safety and Health Act) 1910.106, flammable liquids code	It provides requirements for the storage and handling of flammable liquids, including pressure relief requirements; this regulation is similar to NFPA 30 (National Fire Protection Association 30).

and if the density of the bubbler fluid is about 1 g/mL (silicon oil), then system pressure can drop 0.18 bar without liquid getting sucked back from bubbler and air getting into the vent system because it must overcome gravity to get up and over the tubing arch. Therefore, this technique provides some amount of air suck-back prevention without introducing any valves or restricting orifices in the vent system. Furthermore, the connecting vent tube between the vent knockout pot and the bubbler should extend into the bottom of both vessels. If the tube extends to the bottom of both the bubbler and the knockout pot and if the bubbler fluid does happen to get sucked back to the knockout pot unintentionally, then it will push back in the forward direction into the bubbler again once system pressure is reestablished and vent header purge gas begins to flow out of the system again.

All process vents should go through a knockout vessel before exiting the system. If possible, the volume of the knockout pot should exceed that of the process materials to prevent liquids from flowing out of the final vent line in the unintended event that a vessel is overfilled. The capacity of the vent knockout should be sufficient to hold all feed materials for the period of time it would take to recognize an overflow situation and shut down the process. Measurement devices should be included to

detect vessel overflow. For example, the vent knockout vessel could be on a weigh scale, or it could be equipped with a level transmitter tied to alarms and/or interlocks to stop pumps. The bubbler overflow line can be routed to a waste solvent drop, if one exists in the lab hood. Therefore, if so much liquid is pumped into the system that it overflows the process vessel into the knockout pot, overflows the knockout pot into the bubbler, and overflows the bubbler, then it will safely flow into the solvent waste drop.

When operating multiple low-pressure glass CSTRs in series for any operations (e.g., reaction, extraction, crystallization), their vent lines should only share the same vent header if it is acceptable for their contents to interact. For example, a vessel that must remain water free must not share the same vent header as a vessel containing an aqueous phase or water-containing vapor space because water vapors will diffuse from the "wet" vessel to the "dry" vessel. In addition, the potential to overflow from one vessel to another through the vent line must be considered.

An off-gas scrubber should be utilized when needed, which can simply be putting scrubbing solution into the vent bubbler instead of silicon oil. The knockout vessel must be of larger volume than the scrubber flask. If the process has acidic vapors, sodium hydroxide solution could be used in the vent bubbler. If a process could have toxic vapors, the selected scrubbing liquid should be capable of absorbing and, if possible, chemically reacting the toxic gas. The temperature of the scrubbing fluid should be modified (cooling may be necessary); positive agitation could be used to improve gas absorption.

13.8 Valves

Before startup, valves should be properly positioned (open, closed, or open the appropriate amount) for all lines, including process, vent, and utility lines. Making a list of proper valve positions for all modes of operation (startup, steady state, sampling, shutdown) is recommended. All valves should be shown on the equipment drawing.

Ensure that all valves have a solid mounting. If turning a valve handle turns the entire valve body, stress can be created on the fittings and tubing, possibly causing a leak or rupture. Such can be the case if unsecured metal valves are used with PFA tubing, for example. The valves should be solidly mounted first and then the tubing attached to the valve.

Standard ball valves can trap material that may thermally expand or decompose. Use valves that are designed to eliminate this risk if processing unstable compounds.

13.9 Spills and Releases

Secondary containment is recommended, especially when toxic or hot fluids are being processed. Glass process vessels should receive special attention; a catch pan or reservoir should be used in case of a sudden leak or break. The volume of the pan should exceed that of the process vessel.

Consider whether there are any points in the system where just one manual block valve separates the process from the outside. If so, and the process is running at elevated pressure, then cap the ends of the tubing so that opening the valves will not cause the release of process fluid. Consider using two valves in series, for example, in a sample port, and make sure that the procedure dictates opening and closing the valves in series and not simultaneously. If you do not expect to need to open the valves for a normal run, then cap the line on the discharge side of the final valve.

Consider whether physical protection is needed to prevent people from kicking, bumping, leaning, or falling into process equipment because if a process line is broken, then it would result in the potential exposure of people to chemical releases. Also, use shielding or insulation to protect people from equipment contact.

13.10 Inerting

All vessels and tubing should be inerted before flowing organic solvents. If it is possible to form flammable gas mixtures or if handling oxygen-sensitive reagents or products, then all process vessels must be connected to an inert gas (e.g., nitrogen) supply and vent system. In such cases, inert all process vessels and lines before running any operations (chemistry or auto-sampling). Flow a sufficient volume of inert gas through the entire system to ensure that the feed, processing system, and sample system are air free. If the reagents are air sensitive or moisture sensitive and if you are using large-volume syringe pumps, then consider keeping nitrogen pressure on the pumps when they are not in operation.

If you are inerting the system with an inert gas sweep, then the gas flow rate can be controlled using a metering valve. Keep the vent lines from all process vessels open to a vent bubbler. The bubbler will keep air from backing up into the process. Make sure that the liquid in the bubbler is compatible with the process because the solvent will diffuse back into the vapor space. Make sure to use a knockout vessel between the process vents and the vent bubbler so that bubbler liquid cannot suck back into the process if a negative pressure develops in the process.

If you are using vacuum on process vessels, for example, to enable the intermittent transfer of slurry from continuous crystallization mixed suspension mixed product removal (MSMPR) vessels, then make sure that the vacuum is not strong enough to suck back liquid from the bubbler or draw in sufficient air to cause a hazardous concentration of oxygen. Ensure that the vent knockout is sufficiently large to hold all the liquid in the bubbler if it is sucked back. Make sure that the

system headspace volume is sufficiently large so the brief vacuum pull will not draw significantly from the vent. Supply extra nitrogen flow to the vent header each time the vacuum is pulled on process vessels. Do not rely on check valves to prevent backflow from scrubbers or bubblers.

Flammable solvents have a limiting oxygen concentration (LOC), below which ignition is not possible. LOC is a function of temperature and pressure. Values for common solvents under standard conditions are available in open literature[7] (Osterberg et al. 2014); in some cases, testing may be necessary to obtain a LOC value. For processes that use oxygen as a reactant (aerobic oxidations) and processes that generate oxygen as a by-product (e.g., due to peroxide decomposition), it can be especially important to know the LOC since it may not be practical to maintain an oxygen concentration near zero.

13.11 Use of High-Pressure Gases

Mount pressurized gas cylinders properly, according to local regulations.[8] Regulate discharge pressure to lower than downstream equipment pressure limits. Consider installing pressure relief devices immediately downstream from regulators to protect against regulator failure. Vent relief valves to a fume hood, ventilated gas cabinet, or scrubber.

There are several possible ways to minimize the potential amount of gas that can leak from the system. A metering valve can be installed downstream of the regulator. This minimizes the hazardous gas release rate in the event of a downstream leak or line break. Automated interlocks can be used to shut an automated block valve at the outlet of a gas cylinder if the downstream pressure drops suddenly, if there is a system alarm, or if a gas mass flow meter reads that flow exiting the gas cylinder exceeds a set maximum limit. The valve should fail closed (i.e., “air-to-open”) so that the gas supply is closed if the laboratory automation system is not working. Alternatively, excess flow valves can be used to automatically shut off the flow if a sizeable leak occurs downstream. These are self-contained valves that do not rely on an automation system.

A smaller pressure pot can be used between the high pressure gas cylinder and the process (e.g., 1 L pressure cylinder), pressurizing it with only the amount of gas required to run the process for a period of time. This allows the main gas cylinder to be closed, thus limiting a leak to the quantity in the feed pot. If the process is running continuously, then the gas feed pot will need to be repressurized intermittently, which can be automated.

[7] NFPA 69 Standard on Explosion Prevention Systems.

[8] See for example, OSHA 1910.101 on Compressed Gases.

Leak test the system when using gas cylinders. After the cylinder is connected to the system, it and all system components should be leak tested using something like a soapy-water solution and a limited system pressure.[9]

Consider placing gas cylinders in remote locations, either outside or in a well-ventilated area, to minimize the consequences of a large leak when changing cylinders. Ideally, any tubing from the cylinder to the process would contain no or very few fittings. Furthermore, if gas is piped from outside the lab, then the gas supply can be accomplished with a tube inside a tube and leak detection in the space between tubes.

Filter the gas from cylinders to prevent fouling downstream metering valves or ruining downstream automated block valves with tiny solid particles. Compressed gas cylinders can contain debris from the filling process, as well as from chemical reactions that form solids or corrode cylinder and valve surfaces. To reduce the potential for system plugging and chemical contamination due to this cylinder debris, an inline, pressure-rated filter should be installed as near to the gas cylinder as possible.

For particularly toxic or flammable gases, install detectors with audible alarms in the same laboratory hood as the supply gas cylinder. Close gas cylinder valves when the system is not in use. Move cylinders only when capped and secured in a properly designed cylinder cart.

Use laboratory ventilated enclosures specifically designed for hazardous gas cylinders if using small volumes of pressurized hydrogen or other highly flammable or toxic gas. Use hard piping between ventilated enclosure with the cylinder and the fume hood where the reaction is running. Compressed hydrogen represents a commonly used cylinder gas with the potential to generate an explosive gas–air mixture in the event of a rapid leak. Consider the potential places where hydrogen could vent and whether a hazard would be created. Hydrogen–air mixtures are extremely easy to ignite, and the best strategy is to minimize the potential volume of any such mixture. A flow hydrogenation process is governed by all of the typical safety considerations and designs needed for a batch hydrogenation. Scale-ups are best performed in an unoccupied explosion-proof bunker. Because the reactor is typically a long plug flow tube or multiple vertical pipes in series, there may be an increased chance of plugging. In such cases, pressure relief valves are recommended for both the inlet and outlet ends of the reactor. There should be pressure reliefs for all positive displacement pumps, automated shutoffs for liquid pumps, hydrogen feed system interlocks to high- or low-pressure readings, and alarmed hydrogen gas monitors within the bunker. High bunker air turnover can help dilute hydrogen leaks; a typical turnover rate is ten times per hour. Automated interlocks are typically in place based on hydrogen leak detectors and process temperatures and pressures. Pressure checks of the entire system should be performed with inert gases (e.g., nitrogen and helium) prior to using hydrogen, CO, or other hazardous gases. Pressure test should be done at the actual target operating temperatures and 1.2× the operating pressure. Besides

[9] See the pressure relief section for more on leak testing.

leak checks, other prestartup checks may include volume measurements, flow rate, and sampling checks, as well as cleaning and equipment shakedown. Nitrogen can be used for initial checks, but helium is closer in physical size to hydrogen and able to reveal leaks that may occur when hydrogen use begins. The hydrogen feed line, which contains 100% H_2 gas, must be purged at the start. While batch operation quickly supplies a large quantity of hydrogen to the reactor from the beginning of reaction, continuous operation has the added safety benefit of greatly restricting the maximum hydrogen gas flow rate from the source cylinder. As a consequence, it may take longer time to replace nitrogen with hydrogen throughout the system.

If the system uses high-pressure hydrogen from a supply cylinder and if the system also uses nitrogen gas, then the nitrogen gas should also be supplied by a high-pressure N_2 cylinder rather than low-pressure house nitrogen system. Otherwise, it may be possible to backflow hydrogen into the house nitrogen supply. Check valves are used for the nitrogen supply, but check valves occasionally fail; therefore, the use of a nitrogen cylinder is a backup line of defense.

13.12 Chemistry Related

Before starting any experiments, all personnel (operator, engineer, process chemist, analytical chemist) should review the SDS of all chemicals. Safety hazards of the chemistry and the equipment should also be discussed by all team members before testing starts.

Specify personal protective equipment (PPE) (e.g., lab coat, safety glasses, gloves, other chemistry-specific articles) in the procedure. The chemist should evaluate chemical compatibility and provide safety guidelines for a given process. If possible, the chemist should complete batch-mode screenings for solvent, reagent, temperature, stoichiometry, and catalyst before running the reaction in continuous mode. The chemist should determine whether feed and reaction product solutions are soluble at room temperature. If products are not soluble and achieving solubility of all reagents, intermediates, and products is impractical, then additional measures will be necessary to enable continuous processing. Continuously adding another solvent (preferably preheated to reaction temperature) at the reactor exit (through a plumbing tee) while the product solution is still hot may keep products in solution after cooling. An alternative may be operation in a slurry-capable CSTR (similar to running continuous crystallizations) with a means of pumping out slurry.

During the batch runs, visual observations over time are critical in assessing whether the reaction mixture is a liquid or slurry at different conversions. Observed changes in physical state and/or the presence of solid or gas phases should be noted during batch reactions. Solids sticking to the walls of the batch reactor suggest potential fouling problems in a flow system, even if the flow reactor is designed for slurries.

Test the stability and solubility of starting reagent solutions over time in batch mode before running flow. For quality and safety reasons, feed solutions should be

prepared at least 1 day before continuous tests are to be started to verify stability (no solid formation or material degradation) before loading them into stock cans or pumps for continuous testing. Solids inside the pumps could cause plugging, ruin seals, and require pump disassembly inside a hood. Additionally, product solution or slurry stability should be similarly verified to ensure that collected and evaluated product is representative of that immediately produced by the process. Otherwise, proper product receiver storage (e.g., inerted vessel, cooling) must be provided.

The need for calorimetry testing should be assessed by a process safety expert before beginning continuous processing. Reaction calorimetry and thermal stability calorimetry can be used to measure the thermodynamics and kinetics of desired and decomposition reactions. Depending on the instrument design setup, it may also be possible to obtain gas evolution data.

Calorimetry data, along with equipment heat removal characteristics, can be used to select practical and safe operating temperatures, as well as critical safety limits. It is important to understand that there is not an inherent "decomposition temperature" above which a material will experience runaway reaction. Decomposition reactions follow the same basic kinetic laws governing all reactions: reaction rate will increase as temperature is increased. In order to set a maximum safe process temperature, the chemical thermokinetics must be considered in conjunction with the equipment heat removal characteristics. Development of mathematical process models based on calorimetry and equipment data can provide insight into hazards and operability (Rincón et al. 2011).

Continuous flow chemistry expands the allowable temperature and pressure operating windows compared to most standard batch equipment. Nevertheless, the chemical reaction safety measurements should be obtained and used in heat generation rate versus heat removal rate calculations. For slow reactions, simple heat removal calculations can be performed using the overall equipment surface area. For fast reactions, more complicated models are necessary to determine localized heating. Potential upset conditions should also be evaluated—for example, mixing reactants at different proportions or higher rates.

The equipment contents must never enter the flammable or explosive regime. For example, in an aerobic oxidation process the oxygen concentration must be maintained below the limiting oxygen concentration (LOC). Since LOC values vary with temperature and pressure, it may be necessary to obtain experimental LOC data.[7] In most cases, the LOC is greater than 5 vol%, but there are exceptions. Reliance on the elimination of ignition sources as the sole basis of safety is generally not an acceptable approach.

The experimentalist is responsible for safe feed solution makeup. This is according to batch guidelines for safe temperature-controlled agitated vessels and charging of solid reagents into mixed vessels. If the solution will be heated to help dissolve solids, ensure safe heating methods and appropriately rated vessels.

13.13 Heating and Cooling

Special care must be taken to avoid fires when heating to high temperatures. Circulators with hot oils, electrical heating mantles under glass flasks, heat tapes, and leaking fittings inside or above forced convection ovens can all cause fires. Do not use a circulator with the oil above its flash point. Avoid the use of electrical heating tapes, but if heat tapes are used, be careful to not overlap them. Do not use heating mantles at temperatures above manufacturer recommended limits. If the contents of the flask on top of the heating mantle breaks and the contents spill down onto the electric mantle, this can also cause a fire. Even in nonflammable solvents, some products or starting materials can present a fire hazard. If electrical heaters are used, what is the hottest possible surface temperature of the heater? What damage can be caused by process fluids or equipment contacting this hot surface at its maximum attainable temperature? Can tubes, containers, or process instruments be overheated or melted?

Rather than electrical heaters, it is better to use steam for temperatures below 200 °C. Steam heating has less potential to cause a fire than electrical heating. The maximum surface temperature you can get with saturated steam is limited by the steam pressure and thermodynamics. Above 200 °C, it is better to use inerted ovens or intrinsically safe heating elements because of other hazards associated with higher pressure steam.

Oil circulators require many safety precautions. Elevate the circulator above the process vessel if the reactor jacket volume is large so the oil cannot gravity drain back from the jacket and overflow the circulator when it is not running (for lab circulators that are not completely sealed). Do not heat higher than flash point. Consider inerting the interior headspace of lab circulators with nitrogen as a secondary line of defense against fires. If you know that your operating temperature is between 20 and 90 °C, you can use water instead of oil in the circulator, but you must make sure to replenish the liquid level if some of the water is lost to evaporation. Water is nontoxic, it eliminates the fire hazard with hot oil, and it minimizes the impact if the circulator fluid leaks out into the secondary containment in the hood.

If PFRs are heated to high temperatures inside forced convection ovens like GC ovens, then avoid fittings in the oven. Fitting leaks during an experiment can be spotted if the fitting is outside of the oven. If fittings are leaking inside the oven, then the leaks will not be promptly detected. Accordingly, coiled reactor tube lengths must allow the entrance and exit ends to extend outside the oven walls so fittings can be located outside the oven. If you need to use a tube reactor in an oven with fittings inside the oven, you can do it by exception. To prove that leaks do not develop over time due to temperature cycling, do several heat-cool cycles and redo the pressure test. This will show that thermal expansion and contraction have not loosened the fittings in any way.

If a coiled tube reactor is heated inside a GC oven, then fittings located above the oven electrical or electronic components should have secondary containment

between the fitting and the oven to prevent any process fluid from dripping into the oven circuitry. Large-scale ovens for heating pilot-scale tube reactors should be inerted with nitrogen and have interlocks so that they will not heat unless the nitrogen is on. Install flammable atmosphere detectors inside the oven at larger scales.

Cooling heat exchangers should be used downstream from high temperature reactions if running a large enough scale that ambient cooling is insufficient. For example, if volumetric flow rate exiting from a high-temperature continuous reactor is greater than 1 mL/min, then a cooling heat exchanger should be used. At research scale (e.g., 0.1–1 mL/min flow out of a reactor), ambient cooling of the 1/16″ tubing is probably sufficient. If you are using a cooling heat exchanger, then make sure that the cooling fluid on the outside of the tube cannot be heated by the process above its flash point or boiling point. Also, ensure that the cooling fluid on the shell side cannot be pressurized because of heat from the process, even under upset conditions. Use pressure relief on the coolant side of the heat exchanger, or design it so that the cooling fluid flows unrestricted from the heat exchanger to an inerted, vented surge vessel. The freezing of process liquids inside of cooling heat exchangers must be avoided. A cooling heat exchanger is typically used at the outlet of a high-temperature continuous reactor. The temperature of the cooling liquid should not be below the freezing point of the process fluid or the precipitation point of dissolved compounds. If this is not possible, then set up the heat exchanger to run cocurrent, to be automatically controlled, and to restrict the flow of the cooling fluid.

Energetics of reaction, reaction rates, overall heat transfer surface area, and overall heat transfer coefficient should all be known so that you can estimate whether or not heat removal is sufficient. For tubular reactors, the ratio of heat transfer surface area to unit reactor volume is typically about one to two orders of magnitude higher than a batch reactor that would give the same kg/day throughput. See the following table of published examples.

Example	T (°C)	P (bar)	τ (min)	Tube volume (L)	Tube od (mm)	Tube id (mm)	Tube length (m)	L/d ratio	A/V (m^2/m^3)	Batch A/V for the same L/day (m^2/m^3)	Ratio
1. Hydroformylation (Braden et al. 2009)	55	68	1440	32.5	19.1	16.5	151.8	9,194	242	12.0	20
2. Ortho-Claisen rearrangement (Rincón et al. 2011)	230	15	240	0.220	3.18	1.95	73.5	37,635	2049	42.0	49
3.Newman-Kwart rearrangement (Tilstam et al. 2009)	300	68	8	0.220	3.18	2.16	60	27,784	1853	13.5	137
4.Organic azide formation (Kopach et al. 2009)	90	14	20	0.020	1.59	0.635	63.1	99,360	6299	40.9	154
5.Aerobic oxidation (Ye et al. 2010)	100	34	270	7.14	9.53	7.75	151	19,500	516	13.7	38
6.Asymmetric hydrogenation (Johnson et al. 2012)	70	70	720	73	19.1	16.5	340	20,606	242	8.8	28
7.Thermal cyclization	140	68	90	7.14	9.53	7.75	151	19,500	516	9.5	54
8.Thermal deprotection (May et al. 2012)	270	68	9	0.221	3.18	1.96	73.6	37,600	2045	14.1	145

The last two columns in the table compare the continuous reactors to batch reactors. The column titled "Batch *A/V* for the same L/day, m^2/m^3" is an estimate of the heat transfer surface area per unit volume (*A/V*) for a batch cylindrical vessel large enough to achieve the same daily production rate as the flow reactor, as well as the assumed time for filling, reacting, and emptying. The column titled "Ratio" is the calculated ratio of *A/V* for the actual continuous reactor divided by the hypothetical batch reactor.

The rate of reaction is also important. If exothermic reaction is relatively fast compared to mean residence time (τ), then the heat of reaction could be released over a short fraction of the tube length near the inlet. This is the so-called hot spot in the tube reactor. This can also initiate a decomposition reaction. Know the decomposition thermochemistry for your process and ensure that the temperature cannot approach the critical safety limit and that any decomposition is minimal and incapable of causing a hazard due to overheating or overpressurization. See the Chemical Reaction Safety section.

13.14 Materials of Construction

The potential for corrosion and the associated safety issues must be understood. Reactor materials of construction (MOC) must be compatible with the chemistry. For example, if azides are present, then metals like copper can be very hazardous due to the potential for formation of insoluble explosive metal-azide salts. Also, corrosion must be avoided. If database information for estimating compatibility and corrosion rates is unavailable or uncertain, tests can be conducted using MOC coupons in batch reaction experiments. A coupon is a machined piece of metal that is placed inside the process and exposed to process fluids for corrosion tests. Select construction materials for reactors, pumps, tubing, fittings, transmitters, valves, and other wetted process equipment that are compatible with the chemistry at the intended maximum operating temperature. Watch out for stainless steel valves and fittings inadvertently installed in Hastelloy® reaction trains. This has the potential to occur commonly because the stainless fittings and ferrules can look identical to Hastelloy fittings and ferrules. Check rupture disks and disk holders for compatibility. Metering valves and check valves must have compatible O-rings or gaskets. This may require the use of nonstandard parts. Also, ensure that system materials are appropriate for reaction pressure and temperature combinations. For example, avoid the use of polytetrafluoroethylene (PTFE) or PFA tubing at elevated temperature and pressure because pressure rating decreases significantly at elevated temperature. Make sure that the feed solutions are compatible with the pump materials of construction and the pump seals. If not, then consider using a feed coil for research-scale experiments. Select a compatible coiled tube designed for low axial dispersion at low flow rates and use solvent in the pump to push the feed through the coil and into the reactor.

Pressure ratings decrease with temperature. Pressure limits at ambient, processing, and upset temperatures should be evaluated. Never allow a reaction to run at any temperature/pressure combination exceeding the ratings of the equipment. Keep process below the temperature limit for the reaction solvent, considering the potential for thermal decomposition. Be especially wary of exposing self-reactive polar solvents (e.g., dimethylsulfoxide, dimethylacetamide, nitromethane, and dimethylimidazolidinone) to elevated temperatures. As a rule of thumb, decomposition potential should be investigated when handling these solvents above 100 °C, but decomposition may be a hazard at even lower temperature, depending on the other mixture constituents (e.g., strong bases or hydrides) (Wang et al. 2012). See the Chemical Reaction Safety section.

13.15 Avoid Accumulation and Overfilling

Avoid having materials accumulate in the system by ensuring that incoming flows are not faster than outgoing flows for sustained periods of time. Otherwise, accumulation could either pressurize the flow system or cause process materials to exit through unintended pathways, for example, vents. Ensure that interlocks or other layers of protection are in place to prevent accumulation and to maintain target fill levels. Quantification of real-time mass balances is important for a continuous system. If the average mass flow out of the system does not equal the average mass flow into the system, then accumulation could be happening. Placing feed and product vessels on data logging balances is recommended. Overflowing product collection vessels can easily be avoided; nevertheless, it happens more often than it should. Use alarms with automated interlocks and shutdowns. Connect product collection vessels to a vent system with a knockout vessel. Planning should be in place in the event of an evacuation keeping the operators out of the lab for an extended period of time. Make sure that flows are turned off or that it is otherwise not possible for vessels to overfill if pumps continue to run for long time periods. As a final line of defense, if waste solvent sinks or waste solvent drops are available in the lab hoods, then make sure that the end of the vent tube downstream from product collection vessels is situated over the top of the waste solvent drop.

13.16 Static Electricity

Static charge is generated by the contact and subsequent separation of materials. If the charge dissipation rate is not sufficient, charge will accumulate on the process material and/or equipment. This can create the potential for electrostatic discharge, and if this occurs in the presence of flammable gases or vapors, there may be ignition.

All process vessels and pumps should be properly grounded to avoid static charge accumulation. All conductive continuous processing equipment should be grounded with proper grounding straps, including conductive reactors, portable pumps, auto-samplers, and any other type of conductive equipment through which process fluids flow.

Nonconductive equipment and tubing cannot be effectively grounded. In addition, it is difficult to ground nonconducting fluids (e.g., toluene and heptanes) with long charge relaxation (dissipation) times. A mitigation for these issues is to limit flow velocities. For homogeneous (one-phase) nonconductive fluids, limiting the flow to less than 1 m/s is usually sufficient (Expert Commission for Safety in the Swiss Chemical Industry (ESCIS), SUVA 1988). However, there are fluids for which even lower flow velocities may be necessary. These include nonhomogeneous (two-phase) liquids and slurries. In addition, filters have a tendency to create high charging rates (Britton 1999).

When handling flammable liquids, strong consideration should be given to inerting all parts of the system. Processes should not be run with open containers that hold flammable fluids. Process flows should only flow into inerted receiving vessels. Electrical grounding and bonding is only a backup line of defense against fire. Inerting is the primary line of defense.

Extra caution is needed when humidity is low in the laboratories. Low humidity increases the risk of a spark when handling flammable materials. It decreases charge dissipation and increases the potential to retain high voltages, which can later discharge as a spark. If it is necessary, handle flammable liquids or gases in low humidity areas, then inert when possible, minimize flow velocities to reduce charge generation, maintain good ventilation, and ground all conductors. Open filtrations involving nonconductive liquids create one of the highest risks and should be avoided under low humidity conditions. When in doubt, consult an expert.

13.17 Solvent-Only Run Prior to the Real Chemistry Run

Do a solvent-only run to ensure that the flow system is working as intended prior to actual chemistry experiments and that all safety features such as interlocks or auto-shutoffs are working. The solvent run is also a test to ensure that operating procedures are accurate, especially for startup and shutdown. Solvent-only run should be performed at the flow rates, temperatures, and pressures that are subsequently planned with the actual chemistry, with a representative solvent and, if applicable, inert gas instead of a reactive one. The duration of solvent run must be long enough to test all pumps and flow control systems, including switching back and forth several times if dual syringe pumps with automated switching valves are involved.

During the solvent run, system leaks are identified, automated safety interlock operation is confirmed, and proper temperature and pressure measurements are verified. It is also during the solvent run(s) that catch-and-weigh testing is

conducted to ensure that flow rates are accurately reported. This is an opportunity to make sure that mass balances are within the range 95–105%.

The solvent run is the best time to discover potential problems with mechanical aspects of the continuous unit operation, including mass flow rate control, mass flow ratio control, temperature control, pressure control, liquid volume, and residence time control. While all potential problems like fouling and plugging cannot be discovered during the solvent-only runs, the intent is that issues that can be discovered should be discovered at this time.

Verify that you have redundant temperature and pressure measurements. Make sure that the process has built-in redundant temperature and pressure sensors since this is the in-process verification that temperature and pressure readings are correct. The two pressure readings could come from multiple pressure transmitters or a manual pressure gauge on the reactor or a manual pressure gauge on a gas cylinder regulator or pump pressure readings, for example. The two temperature readings could be multiple in-process thermocouples or the temperature readings on a heating or cooling device such as an oven, a cooling or heating bath, or a circulator. Measure the difference between redundant measurements and do not let the difference exceed a specified value.

13.18 Backflow Prevention

When starting a flow chemistry experiment, first make sure to prefill the pumps and feed lines all the way to the tee or mixer. If the reaction will run under pressure, pressurize each feed line from the pump before opening the valve at the tee and starting the flows. Any unintended gas pockets in the feed lines are compressed before the start of the flow chemistry. In this manner, none of the reagent streams can back up into the feed system or any of the other reagent streams. Check valves are used to prevent backflow; they are helpful, but they will not work 100% of the time. They are more reliable in clean nitrogen supply lines, for example, and less reliable in process lines with reagent and product solutions flowing through them because the process materials can eventually cause fouling in the check valves. Check-valve failures should be expected; thus, it is safest to assume that every check valve will fail, and preemptive measures must be implemented to prevent flow reversal when differential pressure is positive in the reverse direction. Those measures include placing a block valve downstream of every check valve and closing that block valve whenever the downstream process pressure is expected to exceed that upstream of the check valve. In other words, do not count on the check valve to work as a block valve in the reverse direction when flows are stopped. Check valves are not a valid basis of safety and only serve as a backup line of defense. Prefilling and pre-pressurizing each reagent feed line and pump and using block valves are the primary lines of defense. Make sure that reagent gases like hydrogen will not backflow to liquid feed lines and process contents will not backflow into utility lines.

13.19 Beware Phase Changes and Thermal Expansion

Physical properties should be known. For example, unintentionally heating solvents beyond their boiling points could create inadvertent pressurization, just as unintentionally freezing system liquids can create blockage of process lines. Consequently, vapor pressure curves and the proximity of intended operating temperatures to freezing and precipitation/crystallization points should be known for all process liquids. Physical property models can be used to estimate pressures at various temperatures. When running superheated reactions, the selected operating pressure should be well above (e.g., 3 bar higher) the solvent vapor pressure to permit pressure fluctuations without inadvertently boiling solvent within the reactor. At a given temperature, the system pressure will not increase if some of the solvent vaporizes, assuming the liquid composition remains relatively fixed. However, the residence time of a PFR will be strongly affected, and solvent boiling can result in the precipitation of nonvolatiles, which may ultimately block the flow path.

If you are running a liquid-phase reaction at high temperatures, you need to understand the effects of thermal expansion. This can be tested experimentally by heating a filled PFR after filling and then stopping flows. If your reactor is filled with solvent and you heat it up from room temperature, then liquid will be pushed out the back end. Make sure that block valves are not closed when you do this or liquid will come out of the pressure relief valve. Also, make sure that your reactor has a pressure relief valve. If you close block valves and heat a liquid-filled tube reactor that does not have a pressure relief valve, the reactor could rupture. Is it possible to completely liquid fill a process line or vessel, valve it in, and then heat that section? If so, use caution and make sure that a valve is open to allow for thermal expansion of the liquid when you start to heat. Ball valves themselves can be the cause of thermal expansion issues; if they are closed at low temperature and then heated, thermal expansion can damage the valve.

If you are using a metal tube reactor that is tightly coiled and if there is a metal bracket tightly holding the tube coils together, then it is possible for the coiled tube reactor to expand more than the brackets. The brackets can actually squeeze and dent part of the tubing, and this can compromise its pressure rating. Therefore, make sure that the physical construction of the PFR support structure allows for a thermal expansion of the metal.

13.20 Automated Interlocks

An automation system should stop feed pump if a process pressure, temperature, level, or flow rate deviates from acceptable limits. If the interlock engages, it is important to know what will happen to all parts of the system when handling the process materials being considered. Likewise, it is important to consider the hazards

created by the failure of an interlock to function correctly. In some cases, such as for continuous Grignard formation reactions, the automation system can be designed to calculate a real-time energy balance and stop the feed if the heat generation drops below the low-low alarm limit (Johnson et al. 2019). There is usually a low-level alarm at which the operators are alerted so that they can make an adjustment and then a low-low-level alarm at which the automation system shuts down the process. Similarly, if a reactive gas like H_2 or CO is used, then automated shutoff valves are placed in the gas feed line after the gas regulator. These are programmed such that if a reactor or system pressure decreases below a certain value, they close to stop gas flowing out of a system leak or rupture. When the product solution is flowing into a collection tank on a data-logging balance, interlocks can be programmed based on a real-time mass balance and/or if the liquid level in the product tank gets too high. Automated interlocks are especially needed if a process will run unattended. The automated interlocks should be demonstrated and verified to work properly during or prior to the solvent-only run by artificially inducing shutdown criteria in a safe manner.

13.21 Sampling

All sampling methods, ports, and devices must be engineered to ensure that representative samples are collected safely, especially from pressurized reactors. It is best to sample from a low-pressure product collection vessel or low-pressure tubing located downstream from the reactor, for example, downstream from a back pressure regulator, rather than directly from a pressurized reactor. If sampling directly from a pressurized reactor, multiple block valves between the operator and reactor must allow isolation and depressurization of the sample within a small volume before it is manually drained to the sample container. Consider second and third samples and how hold-up volume might affect them. If sampling particularly hazardous materials, such as strong lachrymators or highly toxic materials, then PPE and engineering controls such as hoods, ventilated enclosures, and elephant trunks should be the backup lines of defense. For highly hazardous material sampling from a continuously flowing system, manual sampling ports should be designed to be fully contained and completely sealed. This means that a sample bottle includes a fill valve and a vent valve, that the process is connected to the sample bottle by fittings that seal (e.g., compression fittings), and that flow between the process and the sample bottle is restricted such that the sample bottle cannot pressure up. Consider the need for the sample bottle to be rated for the system pressure in case the sampling procedure is not followed correctly or upstream valves leak through. In many cases, automating the sampling process can make sampling more consistent and representative while also minimizing the potential for worker exposure.

13.22 Utilities

Plant or house nitrogen should not be used if its pressure is less than that of the process. Instead, a pressurized gas cylinder should be regulated down and used. When nitrogen is used, analyze all process scenarios and ensure that flow from the system cannot be pushed back into the house nitrogen system or cylinder. Check valves are useful, but only as a backup line of defense. If the process vessel is not pressure rated and is always vented, then the nitrogen supply should include both pressure relief (e.g., with a set pressure of 0.3 bar) and a flow-restricting orifice so that the flow rate into the system cannot exceed the flow capacity of the process vessel vents.

If using house vacuum, install a vacuum knockout vessel large enough to hold all process liquids so that they cannot enter the house vacuum system. Also, use interlocks that will prevent process liquid from overfilling the system and flowing into the vacuum line. Liquid sensors in the vacuum knockout pot should be used to automatically stop the process and close automated valves in vacuum supply lines if process fluids start to flow toward the vacuum supply. Check the vacuum rating of all equipment to be used in the process and ensure that it is sufficient. Use netting or a coating on all glass process vessels that operate under vacuum since any defect (e.g., a star crack) in the glass may render a flask susceptible to implosion. Inspect glass equipment for defects prior to use.

If using house glycol or other types of house cooling or heating fluid, ensure that the system setup prevents the backflow of process materials into the fluid return. Avoid fittings and tubing connections that can slip off or slip apart—compression fittings are recommended as the best practice. If the house glycol line becomes disconnected and spills into the lab hood, it may create a significant spill if the system volume is large.

13.23 Plugging and Fouling

Brainstorm where plugging will be most likely. What are the smallest flow restrictions in your process? These will probably plug first. Needle valves and control valves are prone to plugging with solids. Also, any valves or fittings with constrictions and expansions are likely locations of fouling and clogging. Plan ahead so you will not need to use a wrench to open any fittings on pressurized lines. Preemptively install tees and block valves strategically for pressure bleed points. Walk through all the process lines before startup. Go through all process lines in the process drawing. Think about all potential places where solid plugging could occur. For each scenario, have a plan and install the proper valves and vents so that pressure can be safely released on both sides of the plug. The recovery plan must be assessed for unintended safety issues before implementing. Use of a lock-out/tag-out program should be considered to prevent inadvertent restart of the system prior to the completion of maintenance; this is especially important if multiple people/shifts will be working on the system.

Monitor pressure versus time at pump outlets to see if fouling is gradually occurring. Double-block and bleed valves allow isolation and safe depressurization. In the event of a plug, slowly depressurize through vent lines and manual valves that are already hooked up and leak tested. Verify that a process line is depressurized before working on it. If possible, flush out process fluids with solvent and then blow with nitrogen before disconnecting any fittings.

Avoid plugging at the tube reactor outlet when the product solution cools. Continuously adding dilution solvent at the reactor exit where the product solution is still hot may keep cooled products in solution. Prevent plugging in vent lines caused by the misting of process solutions with dissolved solids: vent slowly, design for adequate vapor/liquid separation prior to the vent, use large vent lines, and avoid vent restrictions. If separating vapor from liquid downstream from a high pressure reaction or unit operation, assume that vent lines can plug and foul with solids. When the pressure increases to the pressure relief point, then the pressure relief will become the vent. Because of this, sensors and interlocks should be used so that the automation system detects when fluids are exiting relief devices. Otherwise, it can go unnoticed and process fluids can continue to flow out of pressure relief vents for extended time periods. Also, a pressure transmitter should be used with interlocks for automated shutdown if the pressure approaches the relief set point. If plugging is occurring in the vent, then it is possible that solids are also being deposited in the pressure relief line and device. It is important to regularly inspect pressure relief systems, especially in such cases.

If a process line is clogged, then you can use techniques to try to free the clog before resorting to a wrench. Alternating between high and low pressure repeatedly on the same side of the plug is often successful. For example, you may be able to alternate between pressuring up and down with nitrogen on the downstream side of the plug, which may gradually work out the solids and free the clog like a plunger. Sometimes it may be necessary to repeat the push–pull process many times from each side of the plug before it clears.

What if the process tubing side of a pressure transmitter (PT) plugs with solids? The transmitter will give inaccurate readings. The actual process pressure could be lower or higher than what the transmitter says. Therefore, be careful where you install pressure transmitters; try to put them in headspaces where the tube between the process and the PT will not fill with solids or liquids that contain dissolved compounds that could precipitate at lower temperatures. Avoid dead legs in the PT line. If the PT provides a safety-critical pressure measurement and is used for an interlock, consider redundancy in a different location that is less likely to foul with solids. Have the process control system perform a cross-compare and trigger an alarm if the difference between the PT values is too large.

Consider the potential for solids to coat an in-process thermocouple causing it to give inaccurate temperature readings. If the thermocouple provides a safety critical temperature measurement and is used for an interlock, consider redundancy in a different location that is less likely to foul with solids. Have the process control system measure a temperature differential between the two redundant measurements and trigger an alarm if the value is too large.

13.24 Mass Balance, Accurate Mass Flow Rates, and Material Tracking Inside the System

In every experiment, compound-specific and overall mass balances should be calculated. The overall mass balance includes all materials entering and exiting the process. It depends not on analytical results but on balance weights and/or volumes and densities. If closure of the overall material balances is not within the 95–105% range, then improvements should be made to the equipment and/or the measurements until total material balances again fall within the ±5% range.

Accurate and precise determination of mass flow rates is important for both quality and safety. If possible, do a catch-and-weigh testing at the beginning, middle, and end of each experiment so that you would know if mass flows are correct. When this is not possible, the catch-and-weigh procedure should at least be completed with solvent-only run before and after the experiment and at the same flow rates and pressures as are used in the experiment. Include redundancy of mass flow measurements if possible. For example, if a mass flow meter is used downstream from a pump, then the feed solution could also be placed on a data logging balance so that the two mass flow measurements can be compared.

Obtaining process parameters for nontransparent flow equipment can be a challenge. Develop methods to permit tracking of what is happening inside of high-pressure continuous reactors and separators, the interiors of which are not visible because the reactor is made with metal tubing or metal autoclaves. Methods must be devised to at least estimate how much liquid, vapor, and/or solid material is inside of a flow tube, stirred tank, packed column, vapor–liquid separator. This can be done when operating at steady state, transitioning to steady state, and during startup/shutdown transitions. One way to know the liquid to vapor ratio in a reactor at steady state is to measure the real τ and measure thermal expansion of the liquid. One way to measure the real τ is to measure the F-curve for transition from solvent to product solution during startup. If it is a symmetric F-curve, which indicates that variance of the C-curve is normally distributed about the mean, then the 50th percentile of the F-curve equals τ. This can also be quantified after a step change in the process or using a nonreactive tracer.

13.25 Shutdown Safety and Waste Disposal

Verify shutdown procedures with the chemist and the engineer. Obtain storage/stability information on vessels and know the proper procedures (including those for utility supplies such as vacuum, nitrogen, glycol, etc.) to safely shut down the system. The typical shutdown procedure continues flowing solvent through the continuous reactor after the feed material is stopped. Product solution or slurry leaving the reactor will get diluted if flow is not diverted to a different collection vessel during the shutdown transition. For example, at least two turnovers of a PFR system should be completed with only solvent before cooling down and depressurizing the

continuous reactor. This flushes out reagents, intermediates, and products. Shutting down the system with process materials inside could produce fouling and plugging or cross-contamination afterward. Flush the entire continuous reaction system, especially the pumps, with solvent immediately after the experiment. This is important for both the safety and the quality of subsequent experiments.

Plan for waste disposal before you start an experiment or production run. Always follow facility-specific waste disposal guidelines, including adjusting the pH of waste (between 6.5 and 10 is a typical range). Depending on the disposal guidelines, the waste may end up in a waste container in the lab or in a larger community waste tank. In either case, consider the potential for reactivity of your waste with other materials that it could be in contact with.

Acknowledgments The authors thank Edward Mark Davis for his consulting on this manuscript. We thank Bret Huff for leading and promoting the continuous process design and development at Eli Lilly.

References

Allian AD, Richter SM, Kallemeyn JM, Robbins TA, Kishore V. The development of continuous process for alkene ozonolysis based on combined in situ FTIR, calorimetry, and computational chemistry. Org Process Res Dev. 2010;15(1):91–7.

American Institute of Chemical Engineers. Guidelines for pressure relief and effluent handling systems. New York: American Institute of Chemical Engineers; 1998.

Baxendale IR. The integration of flow reactors into synthetic organic chemistry. J Chem Technol Biotechnol. 2013;88(4):519–52.

Braden TM, Gonzalez MA, Jines AR, Johnson MD, Sun W-M, Eli Lilly and Company. Reactors and methods for processing reactants therein. United States patent application WO 2009023515 A2. 19 Feb 2009.

Braune S, Poechlauer P, Reintjens R, Steinhofer S, Winter M, Lobet O, Guidat R, Woehl P, Guermeur C. Selective nitration in a microreactor for pharmaceutical production under cGMP conditions. Chim Oggi. 2009;27(1):26–9.

Britton LG. Avoiding static ignition hazards in chemical operations. New York: AIChE; 1999, ISBN 978-0-8169-0800-4.

CCPS. Inherently safer chemical processes: a life cycle approach. Hoboken: John Wiley & Sons; 2009, ISBN 978-0-471-77892-9.

CCPS. Guidelines for engineering design for process safety. Hoboken: Wiley; 2012.

Crowl DA, Louvar JF. Chemical process safety: fundamentals with applications. 3rd ed. Upper Saddle River: Prentice-Hall; 2012.

Expert Commission for Safety in the Swiss Chemical Industry (ESCIS), SUVA. Static electricity: rules for plant safety. Plant/Oper Saf. 1988;7(1):1–22.

Fisher HG, Forrest HS, Grossel SS, Huff JE, Muller AR, Noronha JA, Shaw DA, Tilley BJ. Emergency Relief System Design Using DIERS Technology: the Design Institute for Emergency Relief Systems (DIERS) project manual. New York: American Institute of Chemical Engineers; 1993, ISBN: 978-0-8169-0568-3.

Gutmann B, Cantillo D, Kappe CO. Continuous-flow technology—a tool for the safe manufacturing of active pharmaceutical ingredients. Angew Chem Int Ed. 2015;54(23):6688–728.

Hessel V, Kralisch D, Kockmann N, Noel T, Wang Q. Novel process windows for enabling, accelerating, and uplifting flow chemistry. ChemSusChem. 2013;6(5):746–89.

Johnson MD, et al. Development and scale-up of a continuous, high-pressure, asymmetric hydrogenation reaction, workup, and isolation. Org Process Res Dev. 2012;16(5):1017–38. https://doi.org/10.1021/op200362h.

Johnson MD, et al. Continuous reactors for pharmaceutical manufacturing. In: Nagy ZK, El Hagrasy A, Litster J, editors. Continuous pharmaceutical processing. Cham: Springer International Publishing; 2020. (this volume).

Johnson MD, May SA, Kopach ME, Groh JM, Cole KP, Braden TM, Shankarraman V, Merritt JM. Design and Selection of Continuous Reactors for Pharmaceutical Manufacturing. Chapter 16 in Chemical Engineering in the Pharmaceutical Industry, Second Edition – Drug Substance/API. 2019. Edited by David J am Ende and Mary Tanya am Ende.

Kletz TA. What you don't have, can't leak. In: Chemistry and industry. Richmond: The National Archives; 1978. p. 287–92.

Kopach ME, Murray MM, Braden TM, Kobierski ME, Williams OL. Improved synthesis of 1-(azidomethyl)-3,5-bis-(trifluoromethyl) benzene: development of batch and microflow azide processes. Org Process Res Dev. 2009;13(2):152–60.

Li B, Widlicka D, Boucher S, Hayward C, Lucas J, Murray JC, O'Neil BT, Pfisterer D, Samp L, VanAlsten J, Xiang Y. Telescoped flow process for the syntheses of N-aryl pyrazoles. Org Process Res Dev. 2012;16(12):2031–5.

May SA, Johnson MD, Braden TM, Calvin JR, Haeberle BD, Jines AR, Miller RD, Plocharczyk EF, Rener GA, Richey RN, Schmid CR. Rapid development and scale-up of a 1H-4-substituted imidazole intermediate enabled by chemistry in continuous plug flow reactors. Org Process Res Dev. 2012;16(5):982–1002.

Obermayer D, Balu AM, Romero AA, Goessler W, Luque R, Kappe CO. Nanocatalysis in continuous flow: supported iron oxide nanoparticles for the heterogeneous aerobic oxidation of benzyl alcohol. Green Chem. 2013;15(6):1530–7.

Osterberg PM, Niemeier JK, Welch CJ, Hawkins JM, Martinelli JR, Johnson TE, Root TW, Stahl SS. Experimental limiting oxygen concentrations for nine organic solvents at temperatures and pressures relevant to aerobic oxidations in the pharmaceutical industry. Org Process Res Dev. 2014;19(11):1537–43.

Proctor LD, Warr AJ. Development of a continuous process for the industrial generation of diazomethane. Org Process Res Dev. 2002;6(6):884–92.

Rincón JA, Barberis M, González-Esguevillas M, Johnson MD, Niemeier JK, Sun WM. Safe, convenient ortho-Claisen thermal rearrangement using a flow reactor. Org Process Res Dev. 2011;15(6):1428–32.

Sahoo HR, Kralj JG, Jensen KF. Multistep continuous-flow microchemical synthesis involving multiple reactions and separations. Angew Chem. 2007;119(30):5806–10.

Tilstam U, Defrance T, Giard T, Johnson MD. The Newman−Kwart rearrangement revisited: continuous process under supercritical conditions†. Org Process Res Dev. 2009;13(2):321–3.

Wada Y, Schmidt MA, Jensen KF. Flow distribution and ozonolysis in gas-liquid multichannel microreactors. Ind Eng Chem Res. 2006;45(24):8036–42.

Wang Z, Richter SM, Gates BD, Grieme TA. Safety concerns in a pharmaceutical manufacturing process using dimethyl sulfoxide (DMSO) as a solvent. Org Process Res Dev. 2012;16(12):1994–2000.

Webb D, Jamison TF. Continuous flow multi-step organic synthesis. Chem Sci. 2010;1(6):675–80.

Ye X, Johnson MD, Diao T, Yates MH, Stahl SS. Development of safe and scalable continuous-flow methods for palladium-catalyzed aerobic oxidation reactions. Green Chem. 2010;12(7):1180–6.

Yoshida JI, Kim H, Nagaki A. Green and sustainable chemical synthesis using flow microreactors. ChemSusChem. 2011;4(3):331–40.

Zhang P, Russell MG, Jamison TF. Continuous flow total synthesis of rufinamide. Org Process Res Dev. 2014;18(11):1567–70.

Chapter 14
Evaluating the Business Case for Continuous Manufacturing of Pharmaceuticals: A Supply Network Perspective

Jagjit Singh Srai, Ettore Settanni, and Parminder Kaur Aulakh

Abstract This chapter addresses the challenges of evaluating the business case for continuous manufacturing of pharmaceuticals, looking beyond traditional technical assessments made at the unit operations or individual production facility level. It provides an overview of key concepts, approaches, and tools for the early assessment of supply network configuration opportunities enabled by continuous production processing interventions. Multiple levels of analysis are considered with the aid of examples based on major UK research programs on continuous production process technologies. Particular emphasis is placed on the potential for achieving enhanced product flexibility (in terms of volume and variety) and, depending on scale, the optimum number and location of manufacturing operations to support speed to market and system-level cost benefits. In the case of multiple manufacturing operations using continuous production process technologies, where production facility replication through digital twins is becoming a key enabler, the chapter sets out a supply network design and analysis approach that evaluates the commercial and operational viability of alternative manufacturing supply network scenarios.

Keywords Pharmaceuticals · Supply chain · Supply network design · Continuous manufacturing · Business case · Volume-variety matrix

J. S. Srai (✉) · E. Settanni · P. K. Aulakh
University of Cambridge, Department of Engineering, Institute for Manufacturing, Centre for International Manufacturing, Cambridge, UK
e-mail: jss46@cam.ac.uk

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_14

14.1 Introduction

Conventional pharmaceutical manufacturing of so-called small molecules is generally accomplished using batch processing, an approach that has been successful in providing quality "blockbuster" pharmaceuticals to the public for decades. Growing pressure on cutting costs in healthcare systems, often achieved by favoring nonproprietary products in the procurement process, has constrained the opportunities for brand manufacturers to profit from the introduction of new products while exposing the shortcomings of traditional batch manufacturing in terms of responsiveness and excessive inventory (Shah 2004). An example of rigidities embedded in the current manufacturing footprint of international medicine supply networks is that some manufacturers' response to a possible "no-deal" Brexit was to increase finished packs stockpiling across Europe by 20% (BBC News 2018).

Some shortcomings of traditional medicine manufacturing have been addressed in terms of manufacturing practices through initiatives aimed to promote operational excellence (Friedli et al. 2013). However, most opportunities to radically improve pharmaceutical development, manufacturing, and quality assurance are driven by targeted technology interventions in product and process development, in process analysis and control, as well as in the broader healthcare ecosystem. Over the past few years, sector outlooks have particularly emphasized the role of small-scale, modular continuous processing technologies to deliver highly engineered drug products and more personalized healthcare enabled by digital technologies—see, for example, Stegemann (2015), Srai et al. (2015a), and Rantanen and Khinast (2015). This tendency reflects a more general call to effectively deploy the latest pharmaceutical science and engineering principles and knowledge to improve the efficiency of manufacturing and regulatory processes, respond to the challenges of new discoveries (e.g., novel drugs), and provide novel business opportunities, such as individualized therapy (FDA 2014).

Evidence suggests that industry is responding favorably to the adoption of continuous processing. Exemplar success stories are outlined in Table 14.1. While recognizing potential advantages in terms of agility of supply, speed of development in drug products, reduced factory footprint, and improved quality, most companies regard their continuous manufacturing strategy as exploratory and themselves as early adopters (Dell'Orco and Tix 2018). Especially in the case of pharmaceutical substances (or primary manufacturing), difficulties in producing compelling business cases are perceived by both research-driven and contract manufacturing organizations as a barrier for the uptake of continuous manufacturing in the pharmaceutical industry, despite the merits of individual technologies in enabling reaction chemistry (McWilliams et al. 2018).

To address the challenges of making the business case for emerging manufacturing technologies such as continuous processing, this chapter outlines supply network design rules for the early-on assessment of reconfiguration opportunities in medicine supply enabled by the perspective adoption of such technologies. To achieve this aim, the remainder of this chapter summarizes key insights from major UK research programs delivering demand-led intensified, continuous

Table 14.1 Summary of companies with significant investments in continuous processing

Company	Application	Location	Operations	Throughput	Scale	Investment
Eli Lilly	R&D	US	Reaction Workup Isolation	1–10 (g/h) 5–15 kg/day	Lab Pilot	
GSK	R&D	UK			Pilot	
	Comm.	Singapore				USD 50m
Lonza	Comm. (43 products)	Switzerland	Reaction (micro/mini)			EUR 1m+
Novartis/ MIT	Comm.	Switzerland	Primary and secondary		Pilot	
Pfizer	R&D	US				
		Ireland		Kilo	Lab	USD 11m
Sanofi-Genzyme		UK	Primary			
Sigma-Aldrich	Comm. (70+ products)	Switzerland	Multipurpose		Medium	

Based on Srai et al. (2015b)

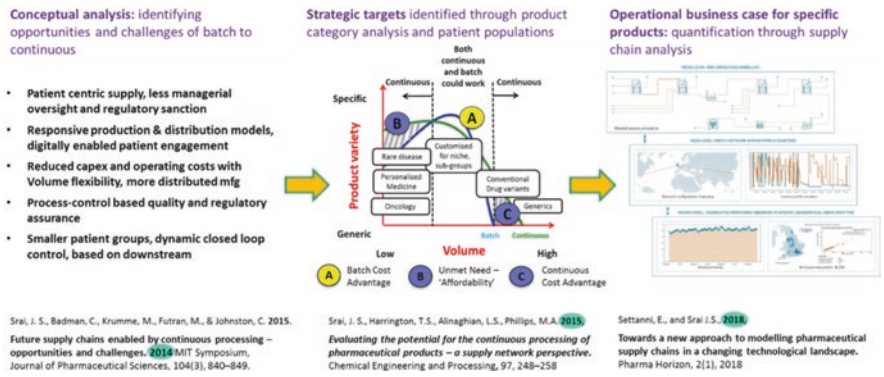


Fig. 14.1 Suggested approach to the business case for transformation problems in continuous manufacturing

technologies—from conceptual analysis through strategic target formulation to operational business case analysis, as shown in Fig. 14.1.

14.2 Overview of the Current State and Emerging Opportunities for Reconfiguring Medicine E2E Supply

As in most highly regulated activities, the structure of global pharmaceutical supply networks is fairly rigid to changes. In the 1980s, the industry shifted from owning most of its infrastructure through vertical integration to regarding activities outside

research and development (R&D) and marketing as noncore (Rees 2011); hence, starting material supply, drug substance manufacture and processing, and downstream supply chain operations are now typically outsourced through third-party commercial relationships.

This industrial setting has incentivized the contract manufacturing of pharmaceutical substances (pharmaceutical ingredients—active pharmaceutical ingredient (API)), as well as chemical intermediates in facilities compliant with good manufacturing practices (GMP): API manufacturing is estimated to account for two-thirds of the fine chemical industry's global production worth circa USD 85bn (Pollak and Vouillamoz 2012) with a progressive move of pharmaceutical manufacturing footprints toward Central/Eastern Europe and China through the 2000s (Boswell 2004, 2007). In this context, processing typically takes place in multipurpose plants consisting of batch equipment where APIs are manufactured in campaigns, the number and length of which largely depends on the need for validated cleaning and changeovers, with cycle times up to 300 days from starting materials to finished goods (Shah 2004).

The typical production-to-reactor-volume range of a commercial API in a fine chemical plant is approximately 15–30 t/m^3 per annum, with a production bay equipped with four to six vessels corresponding to 100 t/year – roughly the yearly production volume of one-third of the 500 top-selling drugs (Pollak and Vouillamoz 2012). In the presence of demanding changeovers, offering product variety can be insidious for a batch environment: for example, due to the number of languages required for blister packing, achieving an overall equipment effectiveness as low as 20% is not uncommon across European production lines (Rees 2011). However, the current industrial trends suggest an emerging need for smaller product volumes to meet more personalized, geographically dispersed demands through an increasingly fragmented product portfolio (Stegemann 2015; Srai et al. 2015a).

A market outlook with more niche volume products and fewer blockbusters will likely demand greater flexibility of unit operations, avoiding prohibitive scale-ups from laboratory to market volumes (Bieringer et al. 2013). An example of this trend is the concept of modular fine chemical plants, where continuous, intensified flow chemistry facilitates "scaling-out" the volume for a given module or "numbering up" the modules operating in parallel (Pollak and Vouillamoz 2012).

From a supply network perspective, the challenge goes beyond making a "binary" choice between batch and continuous production process technologies; rather, the ambition is to explore the potential of targeted technology interventions across the industry to transform pharmaceutical supply chains so that they are more efficient and adaptive, ultimately leading to benefits to the patients (Srai et al. 2015a).

Key opportunities and challenges arising from the pharmaceutical industry's shift toward a primarily "continuous processing"-based supply chain were articulated at the first International Symposium on Continuous Manufacturing of Pharmaceuticals, held in Cambridge, MA, on May 20–21, 2014, and are summarized in Table 14.2. In particular, system-level benefits of targeted technological

Table 14.2 Opportunities and challenges in primarily "continuous processing"-based supply chain

Challenges	Opportunities	Expected benefits
Globalized, volatile markets, increased demand for specialty products, smaller patient populations	Increasing responsiveness to specific market needs through greater product (e.g., dose) and volume flexibility Allowing for late customization in packaging through more distributed and geographically dispersed manufacturing Reducing constraints in terms of "minimum order quantities" within a supply chain design	From >200-day to <70-day inventory within primes and up to 50% reduction in the current 1–2-year inventory Halved cycle time from starting materials to packed product Achieve >5σ, right first time in manufacturing Reduce cost to market by 10%
Costly "buffers," product shelf-life shortening and country-specific regulatory nuances	Designing adaptable, "inventory-light" supply chain configurations through economical supply of smaller volumes Relying on near real-time demand signals as opposed to long-term forecast-driven campaign planning Shortening lead times in product replenishment	
Regulatory requirement for high-quality, validated, efficient, and safe production process	Improving control over process conditions and increasing consistency in chemistry "right-first-time" manufacturing Facilitating and expediting quality control and release through, e.g., in-line process analytical technologies (PAT) Increasing options in terms of synthesis routes involving short-lived unstable intermediates typically avoided in batch processing	
Reduced time windows to reach markets with branded products	Accelerating scale-up post clinical trial design through easier quality-by-design filings Supporting a more "just-in-time," exploratory approach to clinical research Reducing investigational product wastage	
Operational cost linked to increasing complexity	Cutting management burdens and overhead structural costs by minimizing human-to-human hand-offs through enhanced flow-through supply concepts Limiting inventory holding costs and working capital "held up" in multiechelon inventory Reducing footprint through reconfigurable assets with options for numbering up/scaling up	

Based on Srai et al. (2015a)

interventions, such as continuous medicine manufacturing, can be better appreciated when the analysis is carried out "end-to-end" (E2E in short), beyond the four walls of the individual factory or lab, and compared and contrasted with incumbent supply network configurations.

Along with challenges and opportunities, capturing value across the E2E supply chain through continuous manufacturing suggests new developments, some of which have been demonstrated through collaborative, industry-led projects; see, for example, Badman and Srai (2018):

- Digital-technology-enabled, reconfigurable, and adaptive supply networks, whereby response to demand fluctuations is increasingly driven by product demand requirements captured directly between the patient and drug provider through Patient Diagnostic and Management Systems
- Convergence between continuous-processing-based and other medical technologies such as "smart packaging" and additive manufacturing (3D printing) to support more integrated and patient-centric product-service solutions, for example, facilitating product assurance, as well as compliance and adherence to treatment
- More decentralized models of supply and derisked technical transfer earlier in the development timeline through continuous supply centers and reconfigurable assets

In particular, digitally enabled production technologies such as continuous processing with advanced process analytics are now regarded by industry and policy makers alike as viable alternatives to "monolithic" centralized production of medicines in pharmaceutical supply, both commercial (UNCTAD 2017) and clinical (Made Smarter Review 2017).

Demonstrating end-to-end benefits across the pharmaceutical supply chain, however, is not a trivial task. For example, while the use of continuous-flow chemistry, in particular microreactor technology, represents one of the few radical improvements to fine chemical plants in over 25 years, only 10–15% of chemical reaction that could be carried out in microreactors are deemed economically justified by experts (Pollak and Vouillamoz 2012). Also, existing business case for continuous processing seldom extends beyond the four walls of the individual manufacturing plant—see, e.g., Schaber et al. (2011). The following sections focus on how to address these gaps, providing a framework as well as analytical insights.

14.3 Identification of Strategic Targets Through Product Category Analysis

A credible conceptualization of the relative benefits potentially arising at the system level—not just at the manufacturing plant level—from the continuous processing of pharmaceutical products should follow a four-staged analytical framework (Srai et al. 2015b):

1. *Identification of opportunities and barriers to adopting potential alternative product-process technologies*: the first step is to explore which existing markets can be served more effectively and which unmet end-user needs can be targeted through continuous processing. *Volume* and *complexity* play a crucial role in identifying E2E opportunities. Two major questions that arise at this stage are as follows: "What is the volume-scale where the transition from batch to continuous becomes attractive?" and "Which arrangement of batch and continuous processing can achieve the desired volume flexibility?" The resulting

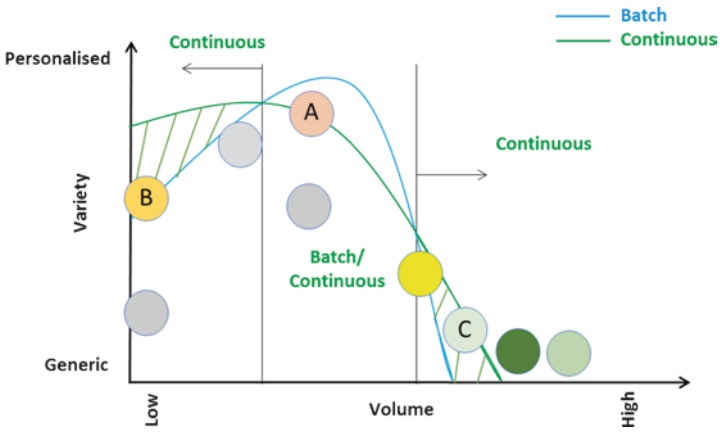


Fig. 14.2 Conceptual volume-variety matrix for pharmaceutical products. (Based on Srai et al. (2015b))

"volume-variety matrix" conceptualization is represented in the central portion of Fig. 14.2: three hypothetical "dots" denote where continuous processing may be beneficial (points "B" and "C" in Fig. 14.2) while considering areas where batch can still be a viable option (point "A"). This stage leads to the identification of *product-process archetypes* that are best candidates for network reconfigurations (i.e., transformative changes in unit operations and actors) enabled by targeted technology interventions, possibly resulting in the emergence of new products.

2. *Subsystem identification and mapping of "current state"*: the incumbent configuration of pharmaceutical supply networks reflects the fact that extensive regulatory requirements operate to a large extent by segregated manufacturing steps, namely, primary manufacturing (i.e., synthesis of active pharmaceutical ingredients/drug substances), secondary manufacturing (formulation of drug products as dose forms), packaging and distribution, and finally dispensing to the patient. Relevant network configuration design rules and trade-offs between them are identified at this stage for each semi-independent subsystem. Value network mapping techniques are particularly useful for defining the overall industrial landscape, as well as actors, technologies, and processes operating within the landscape and its subsystems—Srai (2017) provides a detailed methodology.
3. *Alternative, "future state" supply network scenario formulation and analysis*: alternative scenarios involving continuous technology processing options may be considered and their technology readiness levels assessed. Understanding emerging technologies is crucial at this stage, bringing into the picture scales of production, reordering policies and routes to market that were previously thought unfeasible in a pharmaceutical context (e.g., direct-to-patient e-commerce last-mile delivery, advanced patient diagnostics, close-to-market low-scale integrated manufacturing plant, etc.). Typically, alternative scenarios will

challenge the incumbent supply network configurations for the identified product-process archetypes and introduce new possibilities in terms of product, process, and service offering.

4. *End-to-end supply network system analysis*: the final step consists of E2E network performance analysis, defining potential benefits for the most promising scenarios previously identified. At this stage, existing subsystems are reevaluated from a business perspective as part of a more integrated network reconfiguration agenda considering targeted technological interventions. For example, potential benefits may accrue in terms of revenue/margins, inventory/service levels, etc. against the required capital investments and the technological feasibility of potentially disruptive technologies.

Figure 14.2 shows at a conceptual level the opportunity in the batch-continuous context for particular processing models. The conceptual framework described above has been implemented extensively through case study research for representative products from different regimes of Fig. 14.2, with data being gathered from expert informants during workshops and interviews, as well as secondary data (see, for example, Harrington et al. 2017; Srai et al. 2015b). The potential integration between the framework and technical workflows developed for the implementation of continuous crystallization technologies has also been demonstrated, leading to the identification of critical "attributes" for product-technology screening (Brown et al. 2018). For illustrative purposes, Table 14.3 summarizes insights from the application of the framework to prescreen candidate product segments within specific therapeutic areas for the future application of continuous processing.

14.4 A Multilayered Approach to Analytical Supply Network Modeling

Proceeding from left to right in Fig. 14.1, the last step of the proposed approach to making the business case for continuous processing is the incorporation of quantitative data and analytical tools to complement the conceptual framework discussed in the previous section. Previous work reviews the extant literature, suggesting that incumbent approaches to pharmaceutical supply chain modeling exhibit the following features (Settanni et al. 2017a):

- Quantitative data and analytical tools tend to be applied to specialist subsystems (e.g., plant-level manufacturing scheduling, especially for APIs; clinical trial multiechelon inventory positioning; distribution and logistics; etc.), suggesting difficulties in overcoming functional silos.
- Most pharmaceutical supply chain models are agnostic to specific technologies and market/patient dynamics. In particular, knowledge about specific manufacturing and information technologies is typically a secondary aspect in pharmaceutical supply chain models, rarely contributing to model formulation. Similarly, knowledge of medicine demand/prescription profiles is often exogenous to these models.

Table 14.3 Exemplar results from the application of the conceptual framework

Framework step	Case-specific insights by framework step (nonexhaustive)			
	Case 1: lipid lowering agent[a]	Case 2: antiretroviral[b]	Case 3: antimalarial[b]	Case 4: diabetes[b]
1: Barriers and opportunities	Recommended for high cardiovascular-risk subjects but debated relationship between mortality and LDL cholesterol levels, technical challenges arising from particles with needle-like morphology	Complexity and adherence management, affordability of new combination drugs in least developed countries	Highly subsidized product with target population expected to increase, posing security issues	Mostly "type 2" with greater concerns for treatment affordability/accessibility, major healthcare cost
2: Current-state mapping	Synthetic statins among top 200 most prescribed drugs, decline of natural statin lovastatin market: low cost (~4 USD/kg)[c], very high volume (~17,000 t/a)[c] manufactured via batch/semi-continuous biosynthetic route	High variety (50+ combination), low volumes (<100 t/a), and high cost (>1000 USD/kg); multistage batch characterized by low yield and high wastage	Chiefly artemisinin, multiple combinations, high volume (up to 250 t/a), medium cost (~400 USD/kg), high counterfeit risk, high yield variability due to plant-based active ingredient	Chiefly metformin, very high volumes (tens of thousands t/a), low cost (<20 USD/kg) with wide choice of nonproprietary equivalents, low variety (<10 combinations), simple chemistry
3: Future-state scenario	Potential repurposing to reduce the growth/metastatic spread of melanomas, prevention of Alzheimer and treatment of some types of muscular dystrophy, particle engineering improvement through continuous processing in secondary manufacturing	Service improvement opportunities through greater API control enabling novel, high-leading combination products using continuous formulation	Potential for continuous manufacturing in alternative synthetic routes with lower capex on newly installed capacity to meet increasing demand	Continuous processing enabled by alternative chemistry with improved kinetics, potential capacity gains with lower manufacturing footprint

(continued)

Table 14.3 continued

Framework step	Case-specific insights by framework step (nonexhaustive)			
	Case 1: lipid lowering agent[a]	Case 2: antiretroviral[b]	Case 3: antimalarial[b]	Case 4: diabetes[b]
4: E2E benefits	Better response to demand flexibility under a repurposing scenario through distributed manufacturing, reduction of manufacturing steps through direct compression enabled by continuous crystallization and spherical agglomeration	Better demand matching through continuous processing can reduce complexity, with a positive impact on inventory levels and lead time; limited scale-up opportunities due to complex synthesis; regulator's priority might not align	Lower dependency on harvesting, better quality control and more consistent purity, reduced risk of counterfeiting through "continuous printing" and "serialization," rapid and local response to outbreaks/ emergency situations leveraging available precursors[d]	Greater benefits under high-volume/flexibility scenarios (e.g., doubling global capacity and novel combinations), demanding asset reconfiguration, and ability to relocate closer to specific markets

[a]Based on secondary data/desk research by the authors
[b]Based on published works (Harrington et al. 2017; Srai et al. 2015b)
[c]Based on Clarivate Analytics' Newport estimates
[d]Based on Bio (2018)

The literature provides little guidance on how to carry out early-on quantitative assessment of future supply network configuration scenarios enabled by intensified, continuous processing technologies. Partly, this could be due to the conservativism with regard to manufacturing technologies discussed in the previous sections. However, exploring the future performance space in which an emerging technology may operate is inherently difficult because it is typically carried out before quantitative evidence is available for specific product-process-market contexts: at best, available data are limited to experimental proof of concepts or partial validation at lab scale (Cucurachi et al. 2018).

Some of these challenges can be addressed by moving away from the incumbent silo approach: as multiple layers of analysis are brought together, a richer picture of the pharmaceutical supply chain and the underpinning technologies can be gained (Settanni and Srai 2018b). These layers are represented schematically in Fig. 14.3 and summarized as follows:

- Analysis of country- or sector-level data on, e.g., markets, interorganizational networks and resource availability ("macro" layer)
- Global supply network configuration optimization and analysis, multiechelon inventory simulation ("meso" layer)

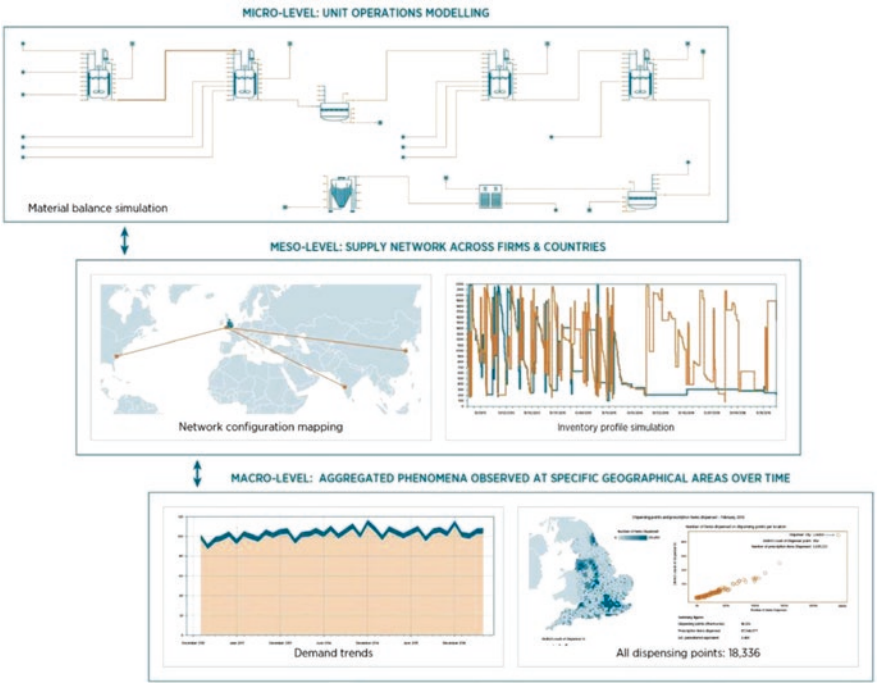


Fig. 14.3 A multilayered approach to pharmaceutical supply chain modeling (expanded detail from Fig. 14.1). (Source: Srai and Christodoulou (2017))

- Formulation of technology intervention scenarios in unit operations/business processes ("micro" layer)

For example, continuous manufacturing technologies are likely evaluated by process engineers at the microlayer (upper slice in Fig. 14.3), emphasizing whether alternative equipment or synthesis routes reduce the number of processing steps, improve yield, and make reactions more consistent. Flowsheet diagrams are the tool of choice at this stage of analysis. However, a purely technical focus could easily lead to a locally optimized system, whereby the benefits of "polishing" a specific factory might be offset by poor (or oversimplified) understanding of the wider supply network behavior downstream in terms of responsiveness to the market dynamics and/or upstream in terms of interactions with the supply base.

Carrying out a technology assessment beyond the four walls of the production plant requires moving towards the mesolayer (intermediate slice in Fig. 14.3) while ensuring that relevant insights from the analysis of unit operations are appropriately "plugged into" network-level configuration and inventory analysis tools. In so doing, difficulties may arise due to the streamlined and "discretized" representation of unit operation technologies that supply chain analysts typically work with due to assumptions that are implicit, such as the following:

- Supply network design is prevalently used in the assembly industry.
- The scope of the analysis is more likely to include site locations than unit operations.

Although without specific application to pharmaceuticals, these assumptions are challenged to an extent by the concept of enterprise-wide process optimization (Quaglia et al. 2012). In practice, the main implications for the specific assessment of continuous manufacturing technologies is that widely used off-the-shelf network design tools are unlikely to provide an adequate technical representation of the corresponding unit operations, thus calling for bespoke analysis.

The third layer of analysis, shown in Fig. 14.3 (bottom slice), represents the analytical exploration of key dynamics in resource availability and demand patterns through, e.g., visual and data-driven analytics. The analysis carried out at the macrolayer can be descriptive and predictive in nature. The latter is perhaps most valued in its own right because big data and digitalization narratives typically suggest that telling associative relationships between unstructured variables can be quickly and effortlessly identified (Waller and Fawcett 2013).

A key link between macro- and mesolayer is the analytical representation of demand signals—actual or forecasted—as well as resource constraints across multiple locations. As in the case of operation technology, the macrolayer is often a black box to the supply chain analyst. For example, in clinical trial supply chains, comprehensive statistical analysis of patient enrollment and dispensing events is common practice; however, insights from field data analysis are rarely integrated in clinical supply network design models, and so are scenarios including alternative technological interventions—with a few exceptions (Settanni and Srai 2018a).

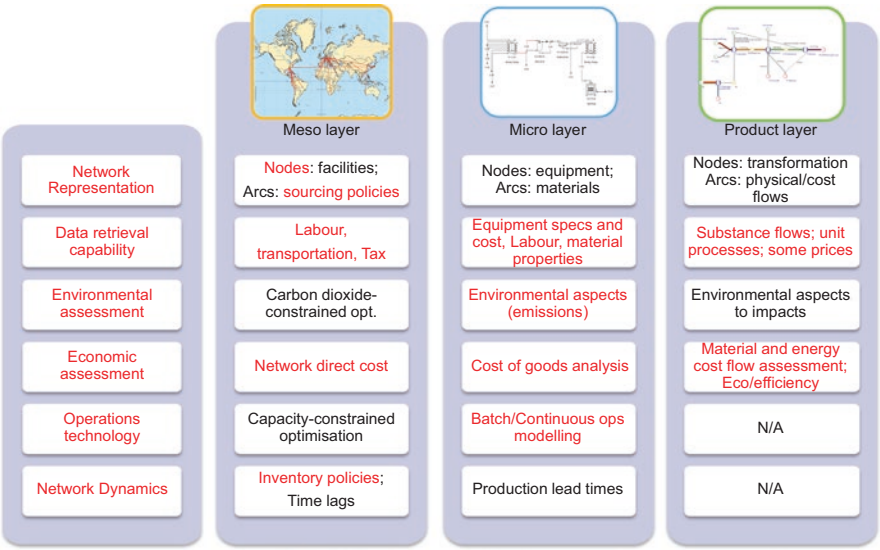
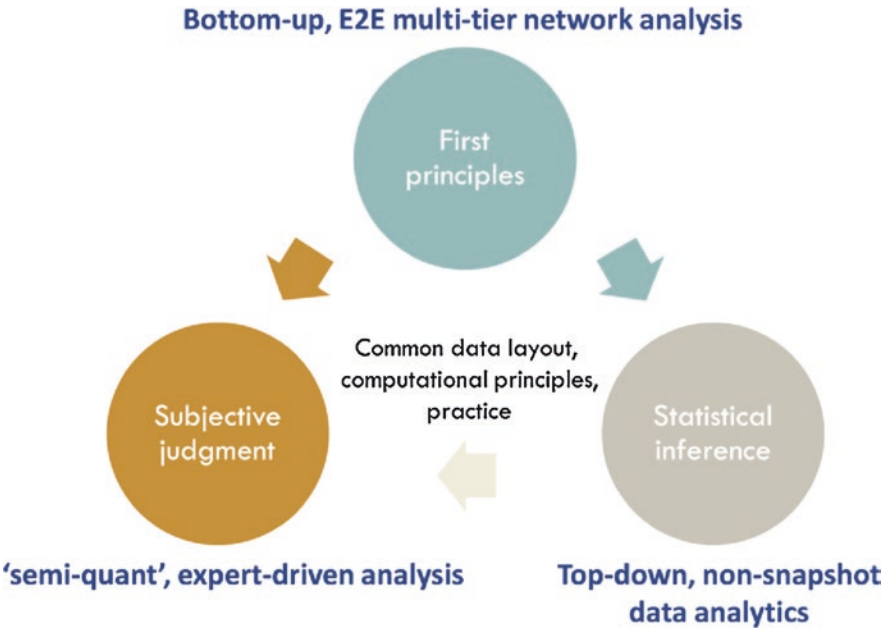


Fig. 14.4 Exemplar evaluation of selected digital network design tools operating at different layers of analysis. (Source: Settanni and Srai (2018b))

From a practitioners' perspective, each layer is readily operationalized through specific digital tools that are commonly employed by most multinationals for technology assessment and supply network design. This operationalization can greatly facilitate bridging different layers of analysis within an organization—and, potentially, the corresponding functional silos. However, in doing so, it also exposes inconsistencies that may arise in the way specialists operating at each layer approach and solve the same problem situation. For example, Fig. 14.4 shows an evaluation of industry-grade, commercial-off-the-shelf tools commonly deployed at the micro- and mesolayer. Despite apparently similar capabilities claimed by specific software vendors in terms of generating optimal supply chain network configurations that jointly reduce costs and environmental burdens, the comparison shown in Fig. 14.4 helps identify whether and to what extent individual tools align in terms of underpinning conceptual models, computational structures, required data, visualization facilities, and data retrieval capability.

Finally, at each layer of analysis described so far, different approaches and tools may be deployed depending on the nature of the data available at specific stages of technology development. Figure 14.5 summarizes which "families" of analytical approaches are typically used with different data—quantitative "snapshot" data, longitudinal/time series data, or hybrid/"semi-quant" (e.g., categorical); it also exemplifies which analytical insights can be generated through each approach to support early-on exploration of a given technology's performance space.



Analytical purpose	First principles	Statistical inference	Subjective judgment
● Explore likely scenarios of future performance for a technology at scale. ● Compare the performance of emerged technology with incumbent technology.	For example, flowsheet simulation informing multi-tier supply network modelling	For example, unsupervised segmentation to quantify 'similarities' between existing technologies	For example, structural modelling, to provide analytical structure to expert opinions and secondary data to reduce subjective bias

Fig. 14.5 Taxonomy of approaches for data-driven, ex ante assessment of emerging technologies. (Based on Settanni et al. (2017b), with additions and modifications)

14.5 Illustrative Applications of the Multilayered Approach

Key concepts introduced in the previous section are further exemplified here using streamlined applications to medicine manufacturing, highlighting specific implications in terms of business case for continuous processing technologies. In the interest of space, the illustration is kept concise and descriptive without disclosing analytical details.

14.5.1 Macrolayer: Data-Driven Industrial System Mapping

Analysis carried out at the macrolayer is concerned with systematically capturing fundamental aspects in mapping supply network configurations, such as the following (Srai 2017): top-level structure of international manufacturing networks, aggregated material flows, indicative product structure and variety, and key actors and transactional relationships.

For example, Fig. 14.6 provides data-driven insights for paracetamol, one of the largest-selling nonproprietary drugs worldwide and the most widely used and prescribed first-line analgesic in the UK (also known as acetaminophen and, by its chemical name, *N*-(4-hydroxyphenyl)acetamide). The focus is on the UK, all data being obtained by leveraging resources in the public domain.

The top half (Fig. 14.6a) looks closer to the point of dispensing to the patient providing a foundation for estimating demand signal at specific geographies in the absence of business-specific data. Proceeding from left to right in Fig. 14.6a, location-specific "hotspots" in terms of prescription dispensing are identified regardless of specific products and trends for specific product formulations in terms of the number of script, expenditure, and units dispensed (500 mg tablets can be considered a typical formulation for this drug), both in absolute terms and relative to total prescriptions across all prescribed medicines in a month. These aggregate, sector-level figures can then be used to simulate more granular demand patterns such as those required at the mesolayer to evaluate alternative supply network designs.

The bottom half (Fig. 14.6b) leverages a completely different set of data (regulatory API registration certificates across the EU) to highlight that, at present, no manufacture of paracetamol as an active pharmaceutical ingredient (API) takes place in the UK or Europe. The likely supply base for crude paracetamol API is located in the USA, China, and India, shipping API varieties (color coded in Fig. 14.6b) toward UK activities that qualify as either importers (17) or distributors (5). Indicative capacity is summarized through a box plot to emphasize the variability in the scattered data disclosed by manufacturers. From a practitioner's perspective, the insights summarized in Fig. 14.6 can be particularly valuable, considering the limited coverage of APIs in published chemical market reports, such as from the early 2000s, and the increasing commercial interest in healthcare data, especially with regard to prescription records (Steinbrook 2006).

Besides buyer–supplier relationships defined in terms of material flows, analysis carried out at the macrolayer can shed some light on contractual relationships between key actors fostering collaboration on specific technologies. For example, Fig. 14.7 visualizes a network of collaborative agreements concerning two natural statins, lovastatin and simvastatin, considered in Table 14.3.

Within this specific network, hubs (nodes with the greatest number of inbound/outbound connections) can be determined by visual inspection to include Kos Pharmaceuticals (acquired by Abbott in 2006) and, rather expectedly, Merck (appearing as multiple nodes due to changes that occurred over time in the

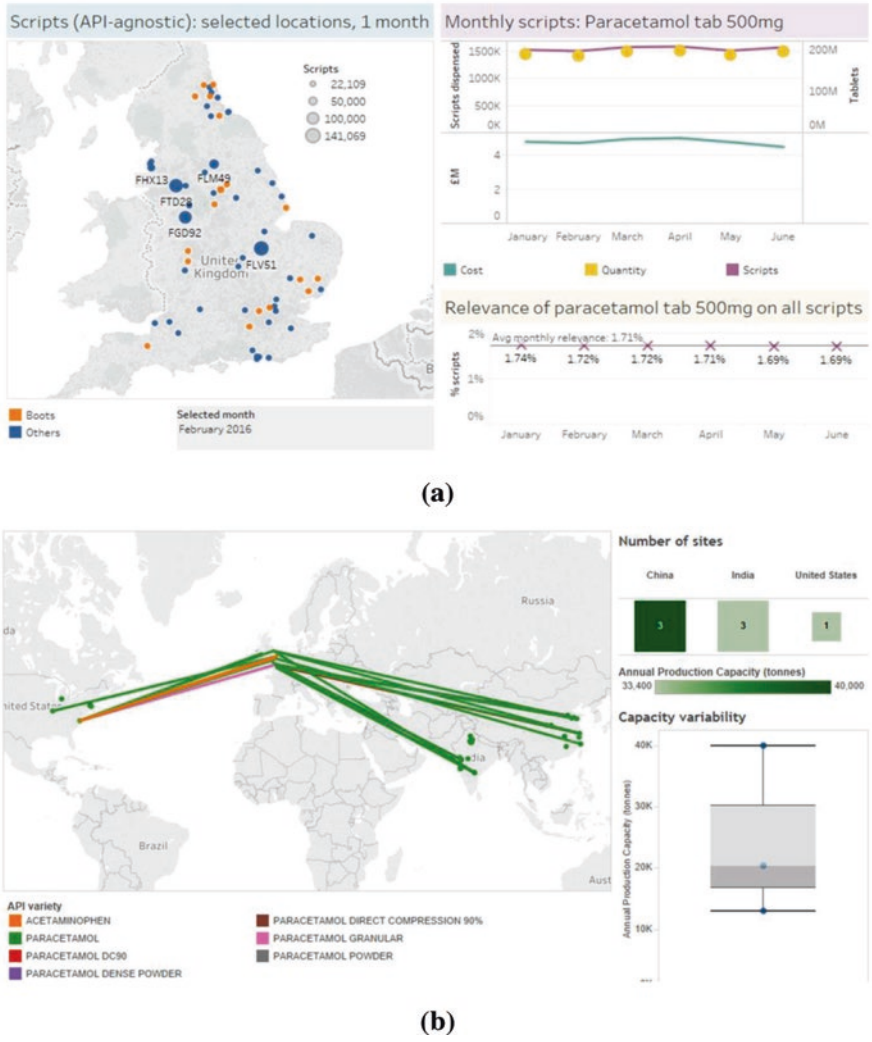


Fig. 14.6 (**a**) Product-specific dispensing estimate in England (raw data: NHS digital: https://digital.nhs.uk/), (**b**) authorized manufacturing footprint based on authorized API importers (raw data: EudraGMP: http://eudragmdp.ema.europa.eu)

company's structure), which pioneered natural statins in the late 1970s (Barrios-González and Miranda 2010). A closer inspection of the edges in the graph highlights the following aspects:

- Only 10% of the captured relationships concern manufacturing and supply agreement of API and/or formulation—over half of which concern Merck's outsourcing of its statin manufacturing activities.

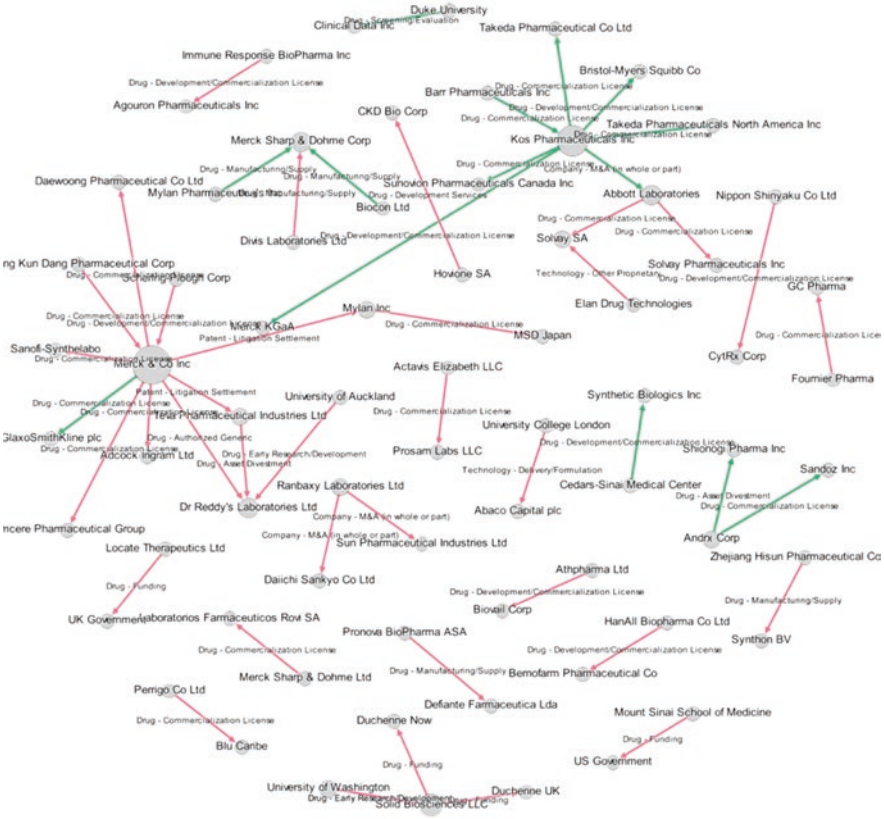


Fig. 14.7 Network representation of collaborative agreements concerning natural statins lovastatin (green arcs) and simvastatin (red arcs)—including combinations. (Authors' elaboration based on Clarivate Analytics' Cortellis data (as of February 2018))

- Twelve percent of arcs (labeled "early research development" or "funding") concern collaborations on simvastatin being repurposed for the treatment of certain types of muscular dystrophy or atherosclerotic plaques.

The above application shows an industrial landscape where potential users of continuous technologies are most likely contract manufacturing organizations seeking to provide greater flexibility through scaling up/numbering up within exclusive supply agreements (Pollak and Vouillamoz 2012). Continuous processing may also play a role in the possible reshoring of key intermediates to protect against disruptions of highly prescribed medicines due to, e.g., foreseeable changes in international trade agreements. With specific regard to simvastatin, opportunities for lab-scale continuous manufacturing may arise from repurposing as novel combinations are developed in adjacent therapeutic areas such as musculoskeletal and cardiovascular.

14.5.2 *Microlayer: Assessment of Current and Perspective Unit Operations*

Models developed at the "micro" layer (upper slice in Fig. 14.3) generate analytical insights into relevant manufacturing technologies and how these are, or could be, implemented under specific supply network configuration scenarios. Following the scheme presented in Fig. 14.5, it is possible to implement distinct modeling approaches at this layer, depending on which type of data is available as the technology of interest develops. In what follows, examples illustrate different principles that can be used to model unit operations depending on what is known about the underpinning technology. Numerical results from an overarching exemplar application are shown in Table 14.4, with a focus on the paracetamol current-state example introduced earlier and selected environmental performance indicators (Settanni et al. 2017b).

Case 1: Chemistry and Manufacturing Technologies Are Fairly Well Understood In the presence of sufficient evidence from a theoretical and experimental viewpoint, it is possible to deploy data-driven models of unit operations using computer-aided flowsheet simulation diagrams. Flowsheet simulation of fine chemical synthesis and pharmaceutical product formulation are supported by fairly well-established modeling platforms available off the shelf and typically yield "first principle" estimates of the following (Petrides et al. 2010):

- Cycle time and throughput estimate
- Bill of material generation from known kinetics/stoichiometry
- Equipment sizing and corresponding resource demand (labor, energy, utilities, etc.)
- Manufacturing cost analysis (typically through a direct costing approach)
- Campaign scheduling feasibility analysis and debottlenecking

In general, a "first principles" approach to unit operation modeling can be used to replicate the actual behavior of an installed base operating, e.g., at the commercial scale for incumbent batch technologies or at lab scale for continuous processing technologies; the same insights can also be helpful to explore the likely behavior of novel operation technologies under scale-up/number-up scenarios. This approach has been used, for example, to compare the economics of batch and continuous technologies for an existing commercial API all the way through tablet formulation (Schaber et al. 2011; Gerogiorgis and Jolliffe 2015). Flowsheet simulation often plays a major role in the environmental life-cycle assessment (LCA) of medicines, LCA being now regarded as a key engineering research area supporting the industry-wide adoption of "green chemistry" practices (Jiménez-González et al. 2011).

When developing a business case for the shift to continuous technologies, the first challenge is to create a flowsheet simulation for existing as well as new technology-based processes using process simulators, e.g. Aspen Plus, SuperPro

Table 14.4 Example results for the environmental assessment of paracetamol obtained from multiple approaches shown in Fig. 14.5

Modeling approach	Unit op data	Approach/tool	Unit of measure	Manufacturing supply chain stage			
				Raw materials	API manufacture	Formulation	Packing
"First principles"	Snapshot quantitative data on material flows based on secondary evidence (literature, patents)	Combined flowsheet diagram simulation and life-cycle assessment (LCA)—see right-hand side of Fig. 14.4 for details	$kgCO_2$-eq/pack	33.35	44.48	0.06	0.91
"Subjective judgment"	Environmental aspect scoring via categorical data based on secondary evidence (literature)	Multicriteria ranking of life-cycle stages by impact category using analytical hierarchy process (AHP)	Dimensionless rank	N/A	0.663	0.259	0.078
Statistical inference	Industry-driven data set on ~30 APIs linking carbon footprint to process characteristics (e.g., annual production volume, no. of synthesis steps, chirality, etc.)	Multivariate regression and supervised segmentation, visual analytics	$kgCO_2$-eq/kg API	N/A	>900	N/A	N/A

Based on Settanni et al. (2017b)

Designer, gPROMS Process Builder, or to make own process models using MATLAB, gPROMS Modelbuilder, etc. The process needs to incorporate all the steps involved in manufacturing a drug, i.e. API synthesis, API separation and purification, and tablet manufacturing so as to conduct an end-to-end case study. Each of these steps consists of an integration of individual unit operations. For the existing process, generally, these unit operations are performed in the form of batches, while for the new process, we consider a part/all of these processes being performed in continuous flow, depending upon which the new process can be either integrated continuous and batch (Schaber et al. 2011) or completely continuous (Jolliffe and Gerogiorgis 2015). The development of flowsheet simulation for the existing process requires data related to process diagram, process conditions, recycling, energy integration, batch sizes, and scheduling. In most situations, these data have to be collected based on extensive/exhaustive search. Different products, however, may present specific challenges. For example, "white biotechnology" products, such as lovastatin, are more widely discussed in terms of API manufacturing technologies, with comparative analyses focused on solid-state versus liquid-submerged fermentation, either batch or semi-continuous—see, e.g., Goswami et al. (2012). Thus, one has to use secondary data and his/her process engineering expertise to make up for the missing information. For the new process, the data requirement remains the same with some alterations: residence times and equipment sizes instead of batch sizes and details of new technology and additional unit operations added/removed from the new process, which can be detailed in a process flow diagram.

Figure 14.8 shows an example of a current-state/future-state unit operation map based on paracetamol as a demonstrator—a detail process description is provided elsewhere (Aulakh et al. 2018b). The existing/current-state flowsheet was developed based on secondary data collected from the literature search; however, due to limited information, scheduling was carried out using in-house simulation in SuperPro Designer. Also, where possible, the process was linked to specific supply geographies within the network configuration presented in Fig. 14.6. The flowsheet simulation for current-state manufacturing would, for example, take into account the fact that central or nearly all the paracetamol manufacturing routes currently exploited commercially involve a reduction-acetylation system where *p*-nitrophenol is reduced to *p*-aminophenol, and the result is acetylated to *N*-acetyl-*p*-aminophenol (Mitchell and Waring 2000), and that in such countries as India, where a large part of the UK paracetamol supply base is located, the *p*-nitro chlorobenzene (PNCB) route is estimated to account for 80% of production (NIIR 2004). With specific regard to environmental assessment metrics, flowsheet simulation on secondary data is sufficient to generate detailed results for each manufacturing stage, as shown in Table 14.4. For the new paracetamol process, a conceptual process flow diagram and experimental, lab-scale data were obtained as a part of research collaboration by CMAC (Continuous Manufacturing and Crystallization) Future Manufacturing Research Hub. Flowsheet simulation was deployed to connect the newly developed and individually optimized processes to form a complete process.

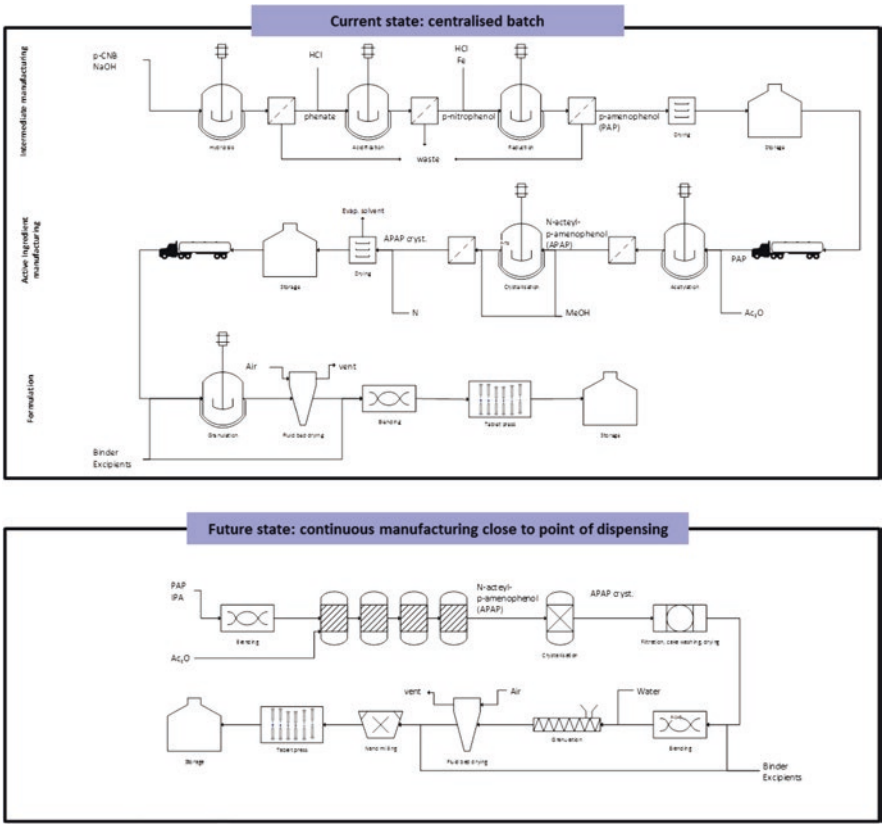


Fig. 14.8 Current- and future-state unit operations using process flowsheet diagrams (details omitted) for paracetamol API manufacturing. (Based on Aulakh et al. (2018b) and Brown et al. (2017))

From a continuous processing perspective, sometimes greater attention is paid to specific unit operations due to manufacturability implications, e.g., lovastatin crystals' characteristic high aspect ratio needle-like morphology or seeded cooling continuous crystallization of paracetamol (Brown et al. 2018). In such situations, the process engineer needs to use secondary data and his/her process engineering knowledge to connect the individual unit operations and respective material and energy flows to develop a continuous process that uses the experimental data as its foundation. For the continuous paracetamol process, similar exercise was performed for developing a process flowsheet that was simulated in SuperPro Designer. Further to note is that while comparing the existing and new processes, one might have to scale up the continuous processes from lab scale to the desired scale of comparison with batch unless data exist for both at similar scales of operation. This also depends upon the type of manufacturing set-up we are considering while developing the business case. In case of paracetamol, a process flowsheet for a

small/microfactory was used to compare it with batch process while setting it up in a distributed manufacturing scenario, as will be discussed in Sect. 14.5.3.

Case 2: Extensive Data (Longitudinal/Historical) Available for Fielded Technologies With the advancement of predictive analytics, the discovery of relevant measures of association from past observations across a variety of products is believed to greatly simplify the cumbersome task of understanding the "mechanics" of specific manufacturing processes in detail.

As in the case of first-principle flowsheet modeling, the cost and environmental footprint of medicine manufacturing are example performance metrics where the principle of statistical inference from an extensive data set of historical or longitudinal observations is extensively used. For example, de Soete et al. (2014) identify potential explanatory variables to predict environmental hotspots from 40 batch "recipes" for 5 APIs covering almost 3000 tasks performed by unit operations. Basu et al. (2008) use historical trend analysis to estimate the cost of goods sold across different types of pharmaceutical manufactures—research driven, generics, biotech—as well as the relationship with R&D effort.

With specific reference to the paracetamol example, results in Table 14.4 were obtained by analyzing excerpts from an industry-contributed data set covering circa 30 undisclosed APIs obtained in confidence. The excerpt covering only primary manufacturing results could be obtained only for one stage of the paracetamol manufacturing chain. No significant association could be found; however, a supervised segmentation approach was deployed to predict whether the product's carbon footprint would be greater than a certain threshold, rather than its magnitude.

The application of statistical inference principles to evaluate alternative manufacturing technologies has been less extensive so far. Adjacent application includes possible synergies between flexible synthesis workstations based on continuous manufacturing and automated selection of optima synthesis paths for a given molecule based on predictive analytics (Peplow 2014). Lapkin et al. (2017) reported the generation of multiple possible reaction routes to convert a biowaste feedstock, limonene, to paracetamol through data mining using Reaxys and a network analysis of the combined literature and in-house reaction set. The approach was based on combining data mining with heuristics and using network representation of chemical knowledge for automating the analysis to evaluate possible routes for paracetamol. The routes were analyzed based on the criteria of mass and energy efficiency, along with route reliability and the selected criteria of environmental importance. Finally, based on the literature and additional in-house data, a complete process flowsheet for manufacturing paracetamol from limonene was developed and modeled (Fig. 14.9). Such advances in automated synthesis planning and automated process design in future might also be extended to flexible and transformative augmentation of optimal supply chain configurations and of the corresponding underpinning technologies.

Case 3: Emerging Technologies Are Only Described Through Secondary Data and Expert Opinion A final case concerns situations where quantitative data

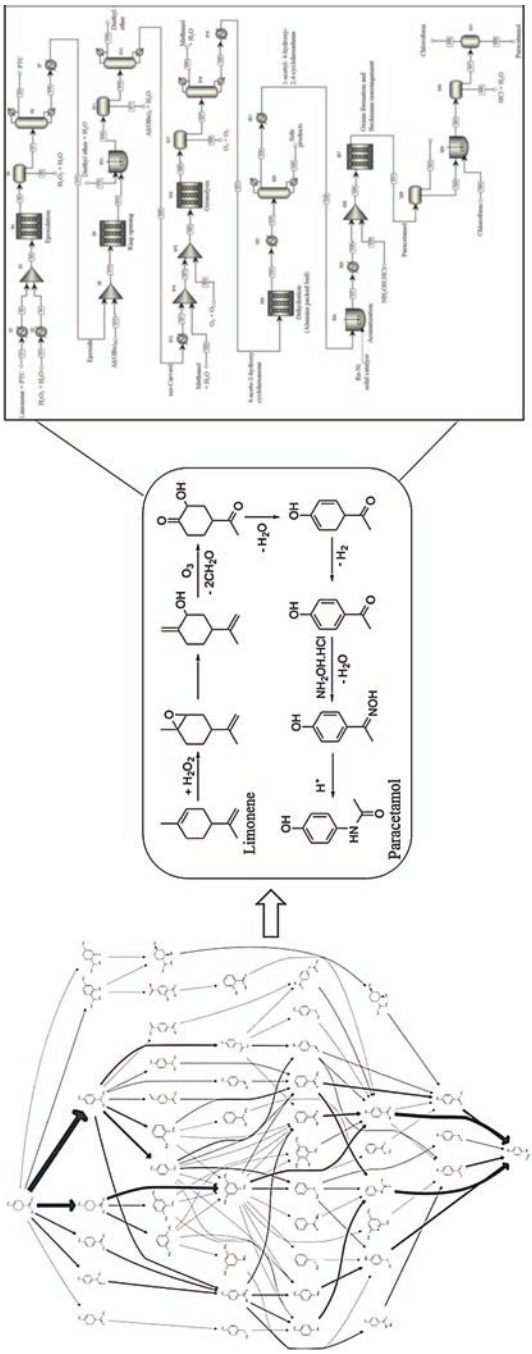


Fig. 14.9 A network of possible synthetic routes from limonene to paracetamol and process flowsheet of one selected route. (Based on Lapkin et al. (2017))

describing the performance of a given technology is not readily available. It is fairly common to elicit expert knowledge for ranking the relative importance of alternative technology options with respect to given criteria. Unlike the previously examined approaches, this type of analysis requires scoring the examined options and criteria with respect to each other, leveraging evidence available through secondary data or elicited from experts. For example, Choudhury et al. (2004) combine expert judgment and quantitative data to estimate the relative priority of different stock-keeping units in balancing the workload of pharmaceutical packaging lines. Another example closer to the chemical processing operations is provided by Manipura et al. (2013), who use expert knowledge to identify lowest risk/highest capability reaction schemes available at the early stage of process development, considering a range of criteria that include the application of continuous flow reactors for an improved process control.

In the paracetamol example, the performance metric shown in Table 14.4 under the subjective judgment approach was derived by gathering evidence from the literature to score resource intensities in a two-step process: first, by evaluating the relative importance of four criteria representing classes of environmental aspects (direct emissions and emissions embedded in resources, energy, and waste) and, second, by assessing three aggregated manufacturing supply chain stages (API, formulation, and packing) with respect to each criteria.

Other applications of the subjective judgment principle include the design of multicriteria decision-making tools to prioritize technological interventions from subject expert knowledge and/or systematic secondary data analysis. The following example is based on Aulakh et al. (2018a), who suggest a multicriteria decision-making approach to support the comparative evaluation of emerging technologies; considering multiple criteria, however, can be a challenge in the absence of detailed and specific quantitative data. In this study, the analytical hierarchical process (AHP) was used to support early-stage assessment of five representative API manufacturing reactor technologies: conventional batch, microreactor, microwave, supercritical fluid (SCF), and continuous stirred-tank reactor (CSTR) cascade. AHP allowed for a set of attributes, which have an impact on technology selection, to be compared with the importance of each attribute relative to its impact on the selection.

The typical steps involved in AHP are as follows: (1) identify multiple criteria/attributes, (2) identify multiple technology solutions, (3) determine the ranking/weights of attributes, and (4) evaluate the ranking of alternatives. In the specific case of rector technologies for the manufacture of pharmaceuticals, each technology's performance was evaluated qualitatively based on secondary data with regards to a range of 11 criteria through a rigorous scoring system.

To illustrate the scoring process using reports from the literature and industrial expert opinions, an example is shown in Table 14.5. The example refers to scoring microreactor and microwave reactor against batch reactor based on the attributes of solids handling capability; batch reactor was scored the highest since it could handle all solids, whereas as per the reports, only 38% of reactions could be handled in microreactor, and microwave was demonstrated to handle slurries above 30%. On the other hand, microreactor was scored higher than microwave reactor when the

Table 14.5 Example of the literature used for reactor technology selection based on attribute scoring (Aulakh et al. 2018a)

	Solid handling	Process cost	Continuous processing
Microreactor	Fifty percent of the reactions can benefit from microreactor technology. However, 62% of these reactions cannot be performed in microreactors due to the presence of solid (Lonza).	The investment costs for a continuous multipurpose microreactor plant were calculated to be as high as or even higher than that of a conventional batch plant (Lonza).	Demonstrated since two decades. It is a commercial product of AzurChem GmbH, 4-cyanophenylboronic acid, made by microprocess technology (1 production period = 10 kg of product).
Microwave	CamWave technology made continuous flow applicable to handling slurries and suspensions in microwave reactors. This opens up continuous flow to virtually all reactions of pharmaceutical interest, well beyond the estimated 30% that is or can be made homogeneous (Cambrex).	Selective heating can significantly increase efficiency and decrease operating cost (AstraZeneca).	Cambrex CaMWave KiloLAB reactor was operated continuously for 32 h, giving 22.3 kg of product.
Summary	All these reactions can be/are being carried out in conventional batch reactors.	Yield improvements compared to conventional technology will be necessary to justify higher investment costs.	Continuous processing is quite well established in microreactors.

continuous manufacturing capability was considered for which batch was scored the least. Through the use of AHP, a larger number of, perhaps individually less important, indicators do not get ignored in the final decision during a simplification that relies too heavily on a small number of key criteria.

The findings of the study highlighted that the ability of a specific emerging technology to replace conventional batch technology depends on the relative importance that experts assign to specific performance areas (Fig. 14.10). The importance of specific attributes is ought to change for specific product segments/therapeutic areas, and therefore the analysis not only needs to be performed individually on different products but also needs to be updated as technology improvements take place. The repeated deployment of the tool for a range of products can define the feasible zones for a given technology within the volume-variety matrix. A possible mapping of technology on volume-variety matrix is shown in Fig. 14.10 (for illustrative purpose only). However, such a multicriteria decision-making tool allows an early-stage assessment of processing technology options for given product-process-market contexts prior to detailed modeling and full supply network configuration design.

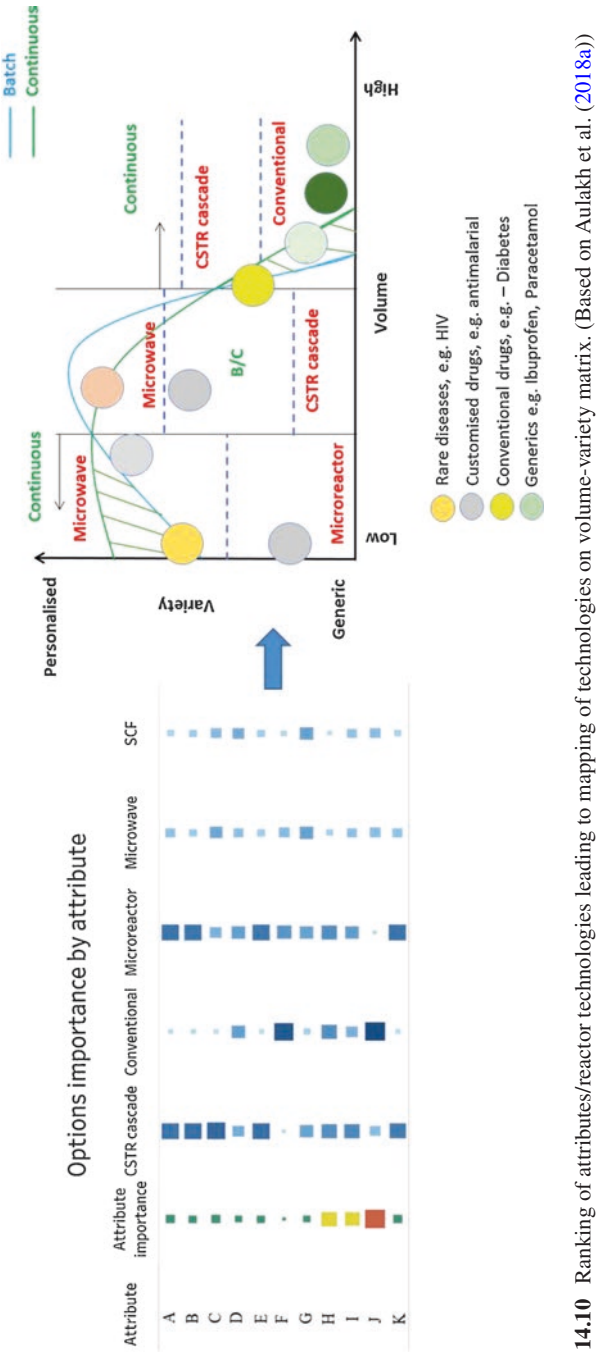


Fig. 14.10 Ranking of attributes/reactor technologies leading to mapping of technologies on volume-variety matrix. (Based on Aulakh et al. (2018a))

14.5.3 *Mesolayer: Network Optimization and Inventory Modeling*

The intermediate layer in Fig. 14.3 combines insights from the other layers (macro and micro) to enable supply network configuration design. In general, the pillars of data-driven supply network design typically require detailed representation of site locations, product architecture, and demand signals (Watson et al. 2013). Considered jointly, these elements define a "superstructure" whose behavior is determined by current and envisaged flows of information and materials through the network, governance relationships between key organizations, and product-process technological intervention scenarios (Srai and Gregory 2008).

Figure 14.11a, b shows a typical structure for a generalized, commercial pharmaceutical supply network configuration; it also maps against such the figure shows example techonlogy interventions at specific supply chain stages, based on evidence from a recently concluded, industry-led UK research program (Badman

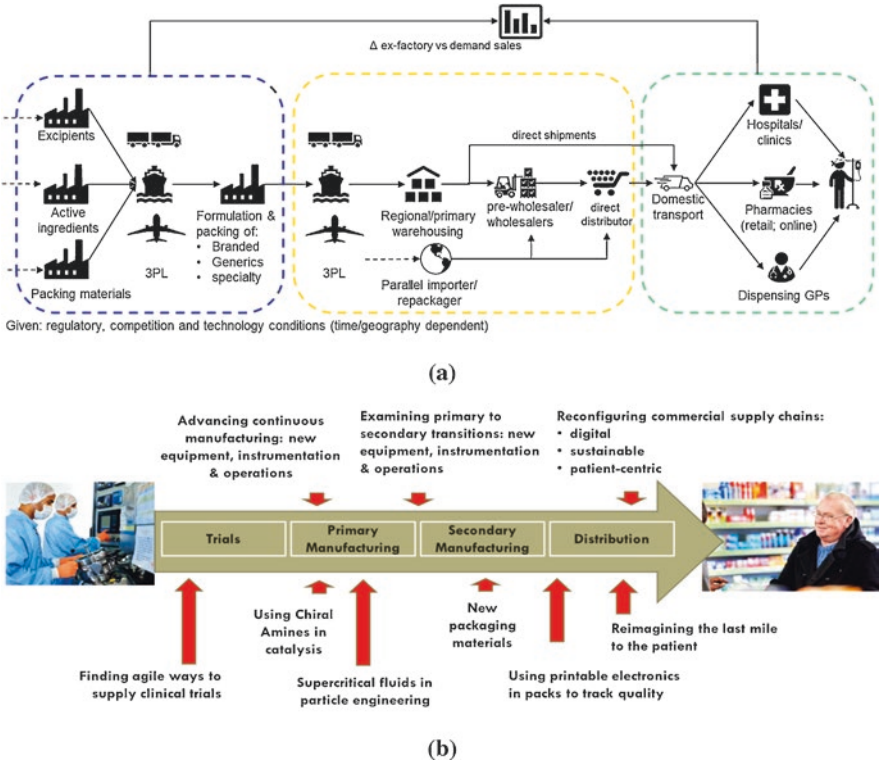


Fig. 14.11 (**a**) "Linear view" of a streamlined pharmaceutical supply chain, (**b**) exemplar interventions at each stage based on outcomes from the industry-led, UK-based ReMediES research program (www.remiediesproject.com)

and Srai 2018). The supply chain structure displayed in Fig. 14.11 is linear as it "follows the pill," rather than mirroring a geographically dispersed manufacturing footprint. Within this typical supply chain structure, one distinguishes the following actors: manufacturing organizations delivering APIs (primary), intermediates and formulated pharmaceutical products (secondary), packaging and leaflets, and nonmanufacturing organizations, including (at minimum) third-party logistics providers (3PL), medicine distributors (prewholesale, wholesale, and direct retail), and dispensing points—e.g., pharmacies, hospitals, dispensing physicians, and even fast-moving consumer goods (FMCG) retailers for some medicines sold over the counter. Moving beyond the physical flows, additional actors emerge. For example, country-specific context such as the US may have a wider range of intermediaries, such as payers and controllers (Rees 2011).

In principle, demand signals for in-market product are generated at the point of dispensing; in practice, misalignment between these signals and product ex-factory forecasts is typically due to intermediaries' behaviors, such as speculation, pipeline fill-and-bleed policies, etc. (Cook 2015). The distinction between generics, branded, and speciality pharmaceutical, as well as the possibility of parallel imports and repackaging in Fig. 14.11, adds to the behavioral complexity of the pharmaceutical supply chains.

Specialized approaches and tools can be deployed at the mesolayer to evaluate the operational performance of novel manufacturing network configurations building on insights generated by the early-on characterizations of relevant technologies carried out at the microlayer. For illustrative purposes, the interplay between knowledge of the operating conditions of an emerging technology and supply-network-level repercussions of deploying continuous processing can be illustrated through hypothetical scenarios of particular relevance for continuous processing, such as those arising from the following network design requirements:

- Capacity flexibility through the use of small-scale modular production concepts with the possibility of "numbering up" or decoupling container-scale devices as required capacity (see, e.g., Bieringer et al. 2013)
- Responsiveness to local demand fluctuation through a "pharmacy on demand" concept enabled by miniaturized and mobile "reaction toolboxes," achieving fully integrated multistep synthesis, purification, and formulation (see, e.g., Lewin et al. 2016)

Continuous technology alone cannot formulate a business case; it is imperative to consider which manufacturing scenario might be credible for a product and technology combination. Figure 14.12 illustrates possible scenarios to evaluate potential repercussions of supply chain configuration redesigns that leverage continuous manufacturing in terms of the necessary scale to respond to specific volume/variety market requirements (high/low) and location (distributed/centralized). When considering the products that are manufactured in medium-to-high volumes, the current-state-of-art manufacturing, followed by the pharmaceutical industry is batch. The possible replacement manufacturing strategy at this scale could be pro-

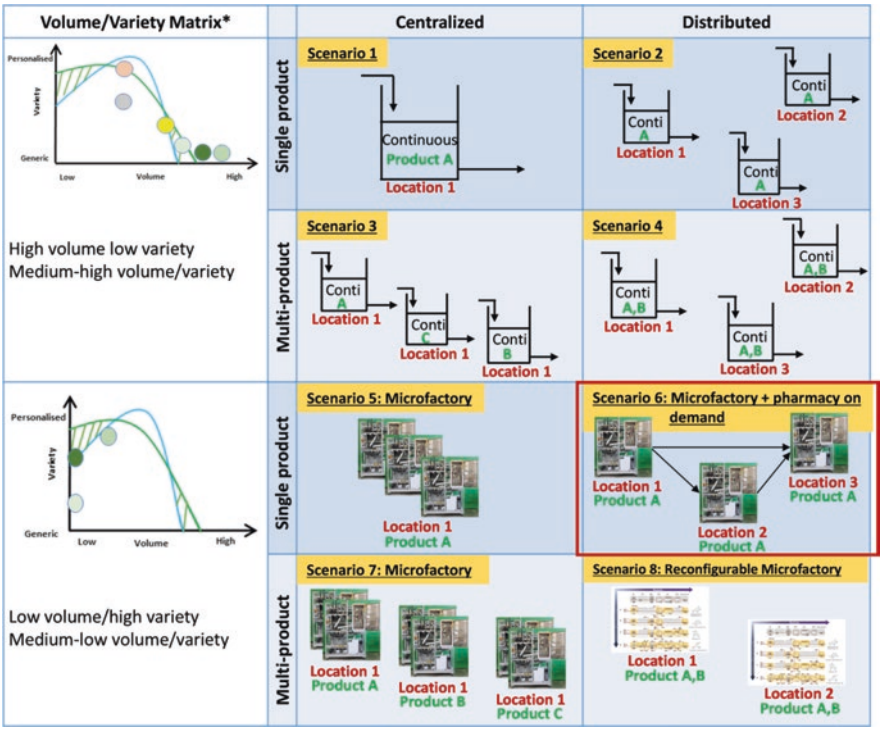


Fig. 14.12 Manufacturing network scenarios enabled by continuous processing technologies. Volume/variety matrix concept as per Fig. 14.2

ducing the products at the same scale in a continuous fashion in single location or in multiple locations, as depicted by scenarios 1 and 2 in Fig. 14.12. The same manufacturing regime can be followed for multiple products (scenarios 3 and 4). For small-scale products, the current batch manufacturing can be substituted by modular units or small/microfactories centralized at a single location (scenario 5) or distributed geographically over multiple locations for on-demand production (scenario 6). A similar strategy can be applied for multiple products (scenario 7) wherein, in scenario 8, there exists an additional future opportunity of exploiting reconfigurable modular units (Lewin et al. 2016). These are highly adaptable small pieces of equipment, which implement real-time monitoring that demonstrates the concept of continuous, small-scale, on-demand production of pharmaceuticals (Adamo et al. 2016).

Scenario 6 (combining the concepts of microfactory and pharmacy on demand) will be further considered for the sake of example. In this scenario, a "microfactory" technology, including an integrated one, such as the one described in Sect. 14.5.2, is deployed in a distributed manufacturing network configuration scenario whereby production activities of a pharmaceutical product are located closer to, or even colocated with, key dispensing points and hence closer to the patient. In the hypothetical

paracetamol demonstrator case used so far, a microfactory unit can plausibly achieve integrated continuous operations from acetylation of intermediate material *p*-aminophenol, through crystallization and filtration, to tablet formulation.

Figure 14.14 shows results obtained when, for a simplified numerical example, considering only six hypothetical dispensing points located in England, underpinned by actual data on high-prescribing points of dispensing, shown in Fig. 14.6, consistently with which a 30-day random demand signal is generated for each location. The same demand signal would drive both current- and future-state scenarios, shown in Fig. 14.13. For each scenario, lead times (in the case of batch manufacturing) or residence times (in the case of continuous processing), bill of materials, and manufacturing cost were obtained by flowsheet simulation (as explained in Sect. 14.5.2). The current-state supply base was mapped from public-domain data, as explained in Sect. 14.5.1.

Insights from unit operation modeling were integrated into a broader network design/optimization model using off-the-shelf tool Supply Chain Guru (Llamasoft, Ann Arbor, MI). To define the network structure, policies were declared for each product in the estimated bill of material and for each manufacturing or distribution node in a specific network configuration with regard to sourcing (multiple/single), transportation (e.g., less than truckload), and inventory replenishment (e.g., RQ—fixed reorder point/order quantity).

With regard to the final product (paracetamol tablets), the right-hand side of Fig. 14.13 shows production and final inventories over time. For the "current-state" scenario, an inventory of the final product is recorded at three echelons (distribution centers and secondary manufacturing facility), whereas for the distributed manufacturing case, an inventory is held at each dispensing point.

The hypothetical future-state scenario in Fig. 14.13 is underpinned by two "penalizing" assumptions about the continuous technology: (1) the microfactories hold no initial inventory at the beginning of the time window, and (2) the minimum run length is 24 h during the time window considered. While different microfactories have been allowed to operate below full capacity while meeting peak demand locally, once production started, it was not allowed to stop and restart. To compensate for the first assumption, in the network optimization step, the microfactories were allowed to meet each other's excess demand if capacity was insufficient. The results in Fig. 14.13 show that, under the abovementioned penalizing assumptions, both scenarios have comparably similar inventory positions at the end of the time window.

To demonstrate the sensitivity of network-wide performance to the level of understanding of the operational conditions for an emerging technology, a second iteration was carried out, allowing a less stringent run size of 5 days, strictly consecutive, with the possibility to operate at any point between minimum and maximum capacity during such time window. Also, microfactories were only allowed to respond to local demand while introducing one-day equivalent initial inventories at each location to prevent infeasibility. As shown in Fig. 14.14, under these revised

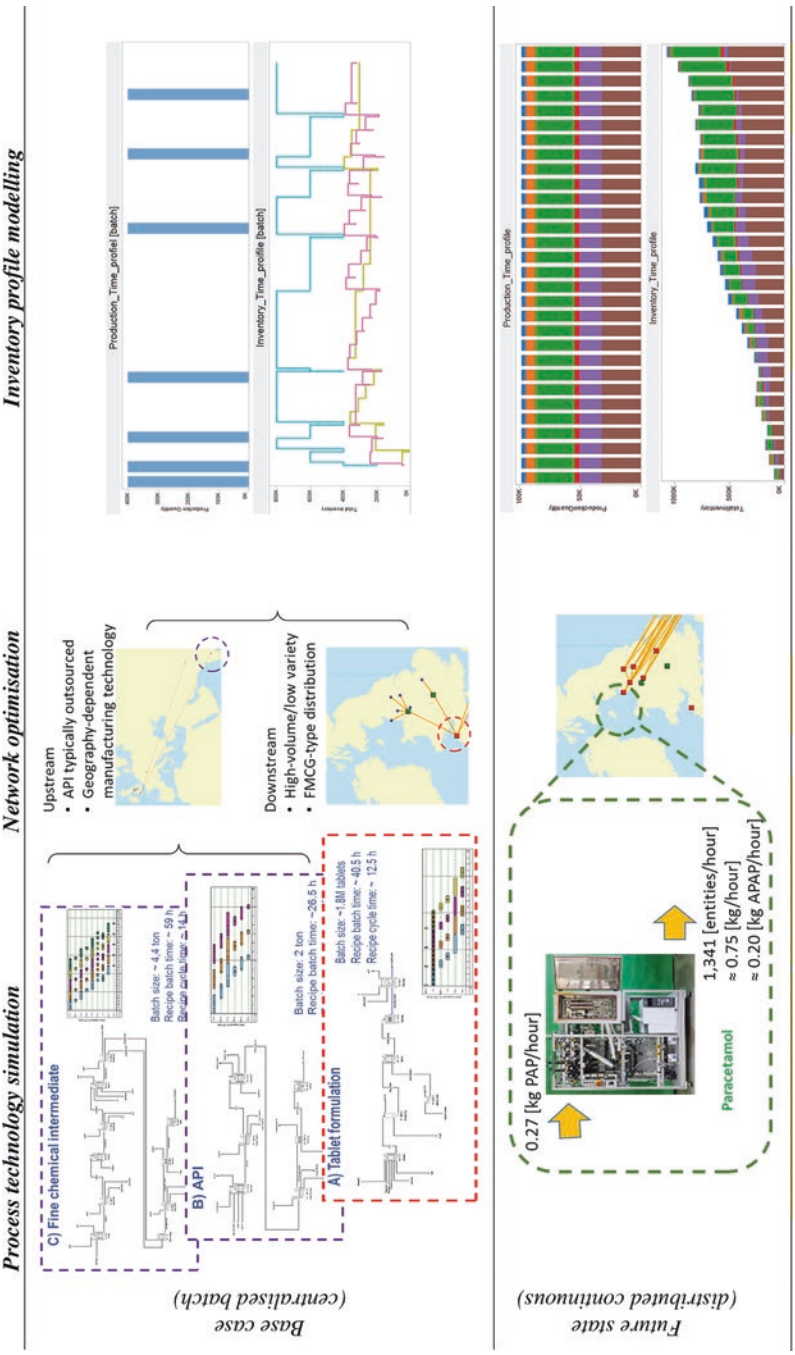


Fig. 14.13 Pharmaceutical supply chain network configuration and inventory modeling for a centralized batch current state (above) versus future state (below) involving distributed continuous microfactories

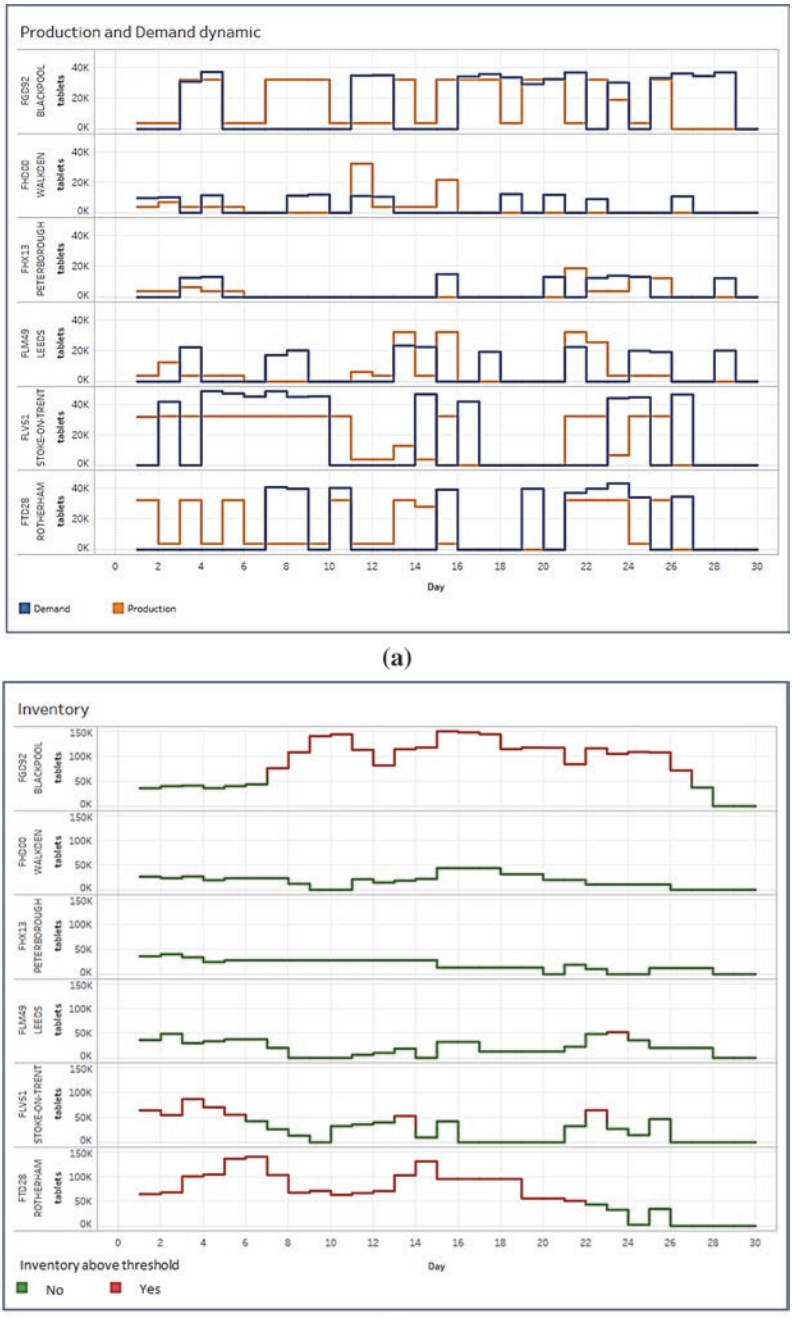


Fig. 14.14 Revised network-level production (**a**) and inventory (**b**) profile simulation for the distributed microfactory scenario (run size of 5 days, one-day-worth initial inventory)

assumptions, the distributed microfactory scenario operates very close to the concept of "on demand pharmacy" with zero inventory at the end of the time period at each location.

14.6 Concluding Remarks

This chapter has provided an overview of the key concepts, approaches, and tools to assess opportunities enabled by targeted technological interventions such as continuous medicine manufacturing from an end-to-end pharmaceutical supply network configuration perspective. It recommends that a "multi-layered" analysis is adopted when evaluating a business case for continuous manufacturing. With the aid of numerical examples based on outcomes from leading UK research programs, we demonstrate how to achieve greater integration between product/process technologies and supply chain dimensions for the evaluation of emerging technologies for specific product categories/therapy areas.

Modeling principles for the exploration of their operational space have been discussed, with an emphasis on the nature of the data that may be available as emerging technologies are developed.

Identification of "feasible zones" for a given technology for desired levels of volume and variety requires that the analysis is iterated for a range of products and under specific hypothesis with regard to the technology's operating conditions. Throughout this chapter, it is emphasized that early-stage production process design is a crucial input to a detailed evaluation of manufacturing and supply chain benefits for specific product-technology configurations through a multilayer system modeling platform.

Achieving a realistic understanding of the operating conditions for chosen continuous technologies enables the supply network designer to evaluate alternative configurations and select those that better support network-level performance improvements in terms of inventory reductions and manufacturing responsiveness. Particular emphasis is placed on the potential for achieving enhanced product flexibility (in terms of volume and variety) and, depending on scale, the optimum number and location of manufacturing operations to support speed to market and system-level cost benefits. In the case of multiple manufacturing operations using continuous production process technologies, where production facility replication through digital twins is becoming a key enabler, the chapter sets out a supply network design and analysis approach that evaluates the commercial and operational viability of alternative manufacturing supply network scenarios.

Funding Acknowledgment Support from the Advanced Manufacturing Supply Chain Initiative (Grant No. 35708-233529, ReMediES—Reconfiguring Medicines End-to-End Supply) and the EPSRC Future Continuous Manufacturing and Advanced Crystallization (CMAC) Research Hub (Grant No. EP/P006965/1) is gratefully acknowledged.

References

Adamo A, Beingessner RL, Behnam M, Chen J, Jamison TF, Jensen KF, Monbaliu J-CM, Myerson AS, Revalor EM, Snead DR, Stelzer T, Weeranoppanant N, Wong SY, Zhang P. On-demand continuous-flow production of pharmaceuticals in a compact, reconfigurable system. Science. 2016;352(6281):61–7.

Aulakh PK, Settanni E, Srai JS. Application of analytical hierarchy process (AHP) for comparative evaluation of API manufacturing technologies. In: 22nd Cambridge International Manufacturing Symposium, September 27–28, 2018, Cambridge. 2018a.

Aulakh PK, Settanni E, Srai JS. Unit process modelling of pharmaceuticals using public domain data: the case of paracetamol. Unpublished working paper. Institute for Manufacturing, University of Cambridge, Cambridge, UK. 2018b.

Badman C, Srai JS. ReMediES: collaborative research in action. White Paper. 2018. Available at: https://remediesproject.com/wp-content/uploads/2018/11/ReMediES-Collaborative-Research-in-Action.pdf.

Barrios-González J, Miranda RU. Biotechnological production and applications of statins. Appl Microbiol Biotechnol. 2010;85(4):869–83.

Basu P, Joglekar G, Rai S, Suresh P, Vernon J. Analysis of manufacturing costs in pharmaceutical companies. J Pharm Innov. 2008;3(1):30–40.

BBC News. Public may have to stockpile drugs in no-deal Brexit. 2018. Accessed 23 Oct 2018.

Bieringer T, Buchholz S, Kockmann N. Future production concepts in the chemical industry: modular – small-scale – continuous. Chem Eng Technol. 2013;36(6):900–10.

Bio MM. Flow chemistry as an enabling technology for essential medicine. Pharma Horiz. 2018;2(1):38–9.

Boswell C. Western manufacturers buy in China. Chem Mark Rep. 2004;265(23):FR4, FR6.

Boswell C. Drug deals go east. ICIS Chem Bus. 2007;272(16):23–5.

Brown CJ et al. MicroFactory prototype report V1.0 10-07-17. Unpublished technical report. CMAC, University of Strathclyde. 2017.

Brown CJ, McGlone T, Yerdelen S, Srirambhatla V, Mabbott F, Gurung R, Briuglia ML, Ahmed B, Polyzois H, McGinty J, Perciballi F, Fysikopoulos D, MacFhionnghaile P, Siddique H, Raval V, Harrington TS, Vassileiou AD, Robertson M, Prasad E, Johnston A, Johnston B, Nordon A, Srai JS, Halbert G, ter Horst JH, Price CJ, Rielly CD, Sefcik J, Florence AJ. Enabling precision manufacturing of active pharmaceutical ingredients: workflow for seeded cooling continuous crystallisations. Mol Syst Des Eng. 2018;3(3):518–49.

Choudhury AK, Tiwari MK, Mukhopadhyay SK. Application of an analytical network process to strategic planning problems of a supply chain cell. Case study of a pharmaceutical firm. Prod Plan Control. 2004;15(1):13–26.

Cook AG. Forecasting for the pharmaceutical industry: models for new product and in-market forecasting and how to use them. Aldershot/Burlington: Gower; 2015.

Cucurachi S, van der Giesen C, Guinée J. Ex-ante LCA of emerging technologies. In: 25th CIRP life cycle engineering (LCE) conference, 30 April – 2 May 2018, Copenhagen, Denmark, vol. 69. 2018. p. 463–468.

de Soete W, Debaveye S, de Meester S, van der Vorst G, Aelterman W, Heirman B, Cappuyns P, Dewulf J. Environmental sustainability assessments of pharmaceuticals: an emerging need for simplification in life cycle assessments. Environ Sci Technol. 2014;48(20):12247–55.

Dell'Orco P, Tix M. The state of continuous processing: thirteen companies share their experiences. Presented at the 3rd international symposium on continuous manufacturing of pharmaceuticals (ISCMP). 2018.

FDA – U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for industry: PAT – a framework for innovative pharmaceutical development. Manufacturing, and Quality Assurance. 2014.

Friedli T, Basu P, Bellm D, Werani J, editors. Leading pharmaceutical operational excellence: outstanding practices and cases. Heidelberg: Springer; 2013.

Gerogiorgis DI, Jolliffe HG. Continuous pharmaceutical process engineering and economics. Chim Oggi. 2015;33(6):29–32.
Goswami S, Vidyarthi AS, Bhunia B, Mandal T. A review on lovastatin and its production. J Biochem Technol. 2012;4(1):581–7.
Harrington TS, Phillips MA, Srai JS. Reconfiguring global pharmaceutical value networks through targeted technology interventions. Int J Prod Res. 2017;55(5):1471–87.
Jiménez-González C, Poechlauer P, Broxterman QB, Yang B-S, Am Ende D, Baird J, Bertsch C, Hannah RE, Dell'Orco P, Noorman H, Yee S, Reintjens R, Wells A, Massonneau V, Manley J. Key green engineering research areas for sustainable manufacturing: a perspective from pharmaceutical and fine chemicals manufacturers. Org Process Res Dev. 2011;15(4):900–11.
Jolliffe HG, Gerogiorgis DI. Process modelling and simulation for continuous pharmaceutical manufacturing of ibuprofen. Chem Eng Res Des. 2015;97:175–91.
Lapkin AA, Heer PK, Jacob P-M, Hutchby M, Cunningham W, Bull SD, Davidson MG. Automation of route identification and optimisation based on data-mining and chemical intuition. Faraday Discuss. 2017;202:483–96.
Lewin JJ, Choi EJ, Ling G. Pharmacy on demand: new technologies to enable miniaturized and mobile drug manufacturing. Am J Health Syst Pharm. 2016;73(2):45–54.
Made Smarter Review. 2017. Available at: https://www.gov.uk/government/publications/made-smarter-review.
Manipura A, Martin EB, Montague GA, Sharratt PN, Houson I. Risk-based decision making in early chemical process development of pharmaceutical and fine chemical industries. Comput Chem Eng. 2013;55(Supplement C):71–82.
McWilliams JC, Allian AD, Opalka SM, May SA, Journet M, Braden TM. The evolving state of continuous processing in pharmaceutical API manufacturing: a survey of pharmaceutical companies and contract manufacturing organizations. Org Process Res Dev. 2018;22(9):1143–66.
Mitchell SC, Waring R. Aminophenols. In: Ullmann's encyclopedia of industrial chemistry. Weinheim: Wiley-VCH; 2000.
NIIR – National Institute of Industrial Research. Drugs & pharmaceutical technology handbook. Delhi: Asia Pacific Business Press; 2004.
Peplow M. Organic synthesis: the robo-chemist. Nature. 2014;512(7512):20–2.
Petrides D, Koulouris A, Siletti C, Jiménez JO, Lagonikos PT. The role of simulation and scheduling tools in the development and manufacturing of active pharmaceutical ingredients. In: am Ende DJ, editor. Chemical engineering in the pharmaceutical industry. Hoboken: John Wiley & Sons, Inc; 2010. p. 521–41.
Pollak P, Vouillamoz R. Fine chemicals. In: Ullmann's encyclopedia of industrial chemistry. Weinheim: Wiley-VCH; 2012.
Quaglia A, Sarup B, Sin G, Gani R. Integrated business and engineering framework for synthesis and design of enterprise-wide processing networks. Comput Chem Eng. 2012;38:213–23.
Rantanen J, Khinast J. The future of pharmaceutical manufacturing sciences. J Pharm Sci. 2015;104(11):3612–38.
Rees H. Supply chain management in the drug industry: delivering patient value for pharmaceuticals and biologics. Hoboken: John Wiley & Sons, Inc; 2011.
Schaber SD, Gerogiorgis DI, Ramachandran R, Evans JMB, Barton PI, Trout BL. Economic analysis of integrated continuous and batch pharmaceutical manufacturing: a case study. Ind Eng Chem Res. 2011;50(17):10083–92.
Settanni E, Srai JS. Combining field data analysis and simulation to evaluate an alternative Just-In-Time clinical trial supply strategy. In: 22nd Cambridge International Manufacturing Symposium, September 27–28, 2018, Cambridge. 2018a.
Settanni E, Srai JS. Towards a new approach to modelling pharmaceutical supply chains in a changing technological landscape. Pharma Horiz. 2018b;2(1):26–8.
Settanni E, Harrington TS, Srai JS. Pharmaceutical supply chain models: a synthesis from a systems view of operations research. Oper Res Perspect. 2017a;4:74–95.

Settanni E, Srai JS, Yatskovskaya E, Harrington TS. Exploring generalisations for sustainability assessment in medicine manufacturing networks. In: 24th International Annual EurOMA Conference, July 3–5, 2017, Edinburgh. 2017b.

Shah N. Pharmaceutical supply chains: key issues and strategies for optimisation. Comput Chem Eng. 2004;28(6–7):929–41.

Srai JS. Mapping industrial systems – a supply network perspective on enabling technologies, processes and actors. Int J Manuf Technol Manag. 2017;31(1/2/3):82–99.

Srai JS, Christodoulou P, Settanni E. Next generation supply chains: making the right decisions about digitalisation. Institute for Manufacturing, University of Cambridge, 2017. Available online at https://www.ifm.eng.cam.ac.uk

Srai JS, Gregory M. A supply network configuration perspective on international supply chain development. Int J Oper Prod Manag. 2008;28(5):386–411.

Srai JS, Badman C, Krumme M, Futran M, Johnston C. Future supply chains enabled by continuous processing – opportunities and challenges. May 20–21, 2014 Continuous Manufacturing Symposium. J Pharm Sci. 2015a;104(3):840–9.

Srai JS, Harrington TS, Alinaghian L, Phillips M. Evaluating the potential for the continuous processing of pharmaceutical products – a supply network perspective. Chem Eng Process. 2015b;97:248–58.

Stegemann S. The future of pharmaceutical manufacturing in the context of the scientific, social, technological and economic evolution. Eur J Pharm Sci. 2015;90:8–13.

Steinbrook R. For sale: physicians' prescribing data. N Engl J Med. 2006;354(26):2745–7.

UNCTAD – United Nations Conference on Trade and Development. World investment report: investment in the digital economy. 2017.

Waller MA, Fawcett SE. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. J Bus Logist. 2013;34(2):77–84.

Watson M, Lewis S, Cacioppi P, Jayaraman J. Supply chain network design: applying optimization and analytics to the global supply chain. Upper Saddle River: FT Press; 2013.

Chapter 15
Regulatory Considerations for Continuous Manufacturing

Elaine Morefield

Abstract This chapter assesses the current regulatory environment, including the relevant regulations and guidelines, and their impact on continuous manufacturing. It outlines key regulatory aspects, including continuous manufacturing process description, control strategy, process validation, and key GMP requirements. This chapter summarizes information generated in the MIT Continuous Manufacturing Workshop 2014. The chapter covers key elements to consider when developing and implementing continuous manufacturing.

Keywords Regulatory · Guidelines · GMP · Submission · Control strategy

15.1 Introduction

This chapter assesses the current regulatory environment, relevant regulations and guidelines, and their impact on continuous manufacturing. It summarizes current regulatory experience and learnings from both review and inspection perspectives. It outlines key regulatory aspects, including continuous manufacturing process description, control strategy, process validation, and key GMP requirements. This chapter summarizes information generated in the MIT Continuous Manufacturing Workshop 2014 and published as a whitepaper (Allison et al. 2015).

Continuous processing provides opportunities for improvements in pharmaceutical manufacturing, including the following:

1. An integrated process with fewer separate unit operations (e.g., safer, faster response times; more efficient, shorter times; integrated control)
2. Smaller equipment footprint (e.g., potentially smaller active pharmaceutical ingredient (API) requirements, more flexibility, lower costs, environmental friendliness)

E. Morefield (✉)
VaxForm, LLC, Davenport, FL, USA
e-mail: elaine.morefield@vaxform.com

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2_15

3. An enhanced development approach (Quality by Design, incorporating process analytical technology (PAT))
4. Real-time product quality information
5. Potential for easier scale-up/down to accommodate changing supply needs

15.2 Current Regulatory Environment

The current regulatory environment supports advancing regulatory science and innovation, which may include abandoning some traditional manufacturing practices in favor of more flexible and efficient continuous manufacturing. Regulatory authorities in the three ICH regions and an increasing number of non-ICH regulators are encouraging the industry to adopt new technology, as supported by ICH Q8 (R2), Q9, Q10, and Q11, and to introduce Quality by Design (QbD) concepts, emphasizing science and risk-based approaches to assure product quality. The US FDA has issued a draft guidance, Quality Considerations for Continuous Manufacturing Guidance for Industry. This guidance discusses topics including defining batches, control strategy, process validation, scale-up, stability, and bridging existing batch to continuous manufacturing.

The regulatory expectations are the same for batch and continuous processing. The assurance of reliable and predictive processing that is technically sound, risk based, and relevant to product quality in a commercial setting is expected. The risks to quality in continuous manufacturing may differ in some aspects from batch processing, and thus the control strategy will need to vary to assure quality.

15.3 Guidelines

15.3.1 ICH Guidelines

The emergence of ICH Q8 (R2), Q9, Q10, and Q11 guidelines and the accompanying ICH Q-IWG Points to consider and the Q&A documents emphasized that a prospective science and risk-based, “enhanced” approach to development and life cycle management could increase the assurance of quality in the manufacture of pharmaceutical products away from traditional constrained manufacturing processes. Collectively, these guidelines reinforced the adoption of risk-based (Q9) systematic and science-based mechanistic approaches (Q8(R2) and Q11), within a robust pharmaceutical quality system (Q10), to establish an increased level of process understanding and product knowledge that could be translated into a defined design space and governed by an appropriate control strategy (Q8(R2) and Q11). Many of the tools described in these ICH guidelines were not, by themselves, new.

The implementation of these concepts within a more systematic and integrated framework based on sound science and quality risk management introduced a fundamental paradigm shift in product development and manufacturing. It offered a preemptive opportunity for manufacturers to mitigate risks, simplify the regulatory process, and potentially reduce issues and costs in development and manufacturing.[1]

15.3.2 US FDA Guidances

1. The *FDA Guidance for Industry PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance* specifically identifies that the introduction of continuous processing may be one of the outcomes from the adoption of a science-based approach to process design. Process understanding, control strategies, plus on-line, in-line, or at-line measurement of critical attributes that relate to product quality provides a scientific risk-based approach to justify how real- time quality assurance is at least equivalent to, or better than, laboratory-based testing on collected samples. Continuous manufacturing combined with Process Analytical Technology (PAT) can be used for in-process controls and monitoring of in-process materials and process parameters to predict end product performance. Process analytical technology can be the tools needed for successful continuous processing and this guidance indicates the Agencies acceptance of these technologies and approaches.
2. The *FDA Guidance to Industry: Process Validation: General Principles and Practices* aligns process validation activities with a product life cycle concept. The guidance encourages the use of modern pharmaceutical development concepts, quality risk management, and quality systems at all stages of the manufacturing process life cycle. The life cycle concept links product and process development, qualification of the commercial manufacturing process,[2] and maintenance of the process in a state of control during routine commercial production. This guidance supports process improvement and innovation, including continuous manufacturing.
3. The *FDA Draft Guidance to Industry: Quality Considerations for Continuous Manufacturing* discusses topics including defining batches, control strategy, process validation, scale-up, stability, and bridging existing batch to continuous manufacturing.

[1] PhRMA White Paper Implementation and Application of Quality by Design Feb 2013.

[2] The term *commercial manufacturing process* refers to the manufacturing process resulting in *commercial product* (i.e., drug that is marketed, distributed, and sold or intended to be sold). In this usage, the term *commercial manufacturing process* does not include clinical trial or treatment IND material.

15.3.3 ASTM International

The ASTM International Committee E55 on Manufacture of Pharmaceutical and Biopharmaceutical Products has key documents that could be useful for continuous manufacturing:

1. The *ASTM E2537 Standard Guide for Application of Continuous Quality Verification to Pharmaceutical and Biopharmaceutical Manufacturing* describes continuous quality verification (CQV) as an approach to process validation where a manufacturing process (or supporting utility system) performance is continuously monitored, evaluated, and adjusted as necessary. It is a science-based approach to verify that a process is capable and will consistently produce products meeting its predetermined critical quality attributes (CQAs). With real- time quality assurance (which CQV will provide), the desired quality attributes are ensured through continuous assessment during manufacture. Data from production batches can serve to validate the process and reflect the total system design concept, essentially supporting validation with each manufacturing batch.
2. The *ASTM E2968 Standard Guide for the Application of Continuous Processing in the Pharmaceutical Industry* describes best practices for implementing continuous pharmaceutical manufacturing.
3. The *ASTM E2474 Standard Practice for Pharmaceutical Process Design Utilizing Process Analytical Technology* describes the use of process analytical technology in designing and controlling a pharmaceutical manufacturing process.

15.3.4 European Union (EU) Guidelines

The ICH Guidelines referenced above apply in the European Union. Other EU guidelines that might be particularly relevant to continuous manufacturing include the *Guidelines for Process Validation*, where the concept of continuous process verification is introduced; the *Guideline on NIR*, as it is often used as a process analytical technology (PAT) tool for process monitoring and/or control; and the *Guideline on Real Time Release Testing*. Although not required, continuous manufacturing is commonly coupled with real-time release testing (RTRT). Additionally, the European Medicines Agency (EMA) set up a Process Analytical Technology Team in 2003 to support PAT and QbD activities in the EU. The team provides a forum for dialogue between the Quality Working Party, the Biologics Working Party, and the Good Manufacturing Practice/Good Distribution Practice Inspectors' Working Group.

In summary, global and regional regulations, guidelines, and standards are supportive of innovative pharmaceutical development and manufacturing

approaches. Current guidelines may need to be reevaluated with consideration of continuous manufacturing operations as experience is gained.

15.4 Regulatory Considerations

As the pharmaceutical industry and regulatory agencies gain more experience with continuous manufacturing, several regulatory aspects will need to be explored in order to link the principles and practice. While the current regulatory framework is adequate to allow for continuous manufacturing, traditional concepts may need to be further explored or challenged to advance the implementation of continuous processes.

The following aspects are applicable to both batch and continuous processing. In evaluating the differences and similarities between batch and continuous processing, it is important to note that different approaches may be needed for continuous processing:

1. As a matter of principle, the definition of a batch should be stated prior to manufacture. Although each continuous process has unique considerations, one may consider a batch definition based on the quantity manufactured or the duration of the process.
2. In-process controls (IPCs) and sampling considerations will be different. For example, continuous unit operations may have different operating principles; therefore, the sampling considerations (e.g., size, frequency, location) may differ. Setting up acceptance criteria considering representative tested sample size (e.g., large N) should be considered.
3. Acceptable procedures for handling deviations, including detection and removal of nonconforming material in continuous manufacturing processes, must be defined.
4. The rationale for the testing of a continuous batch should be based on considerations of the system dynamics, such as time or amount of material impacted by a deviation or the reaction time for material rejection. Knowledge of the process flow time, back mixing, and clearance rate will be useful in determining the control strategy.
5. The importance of raw material specifications and the lot-to-lot variability of raw materials to the process performance should be considered and studied as appropriate. This understanding should be used to develop the control strategy.
6. Sources of variability should be considered during development, controlled in the control strategy and the success of the control strategy evaluated during validation and/or continuous verification.
7. The evaluation of manufacturing changes and their impact on product quality should reflect relevant risks associated with continuous manufacturing, which may be different from batch processes.

Early and frequent communication between manufacturers and regulators is encouraged to ensure alignment and clarify continuous manufacturing requirements. Some regulatory agencies have the opportunity for site visits prior to the submission of a regulatory application. This can be useful to enhance communication about the manufacturing process for complicated continuous manufacturing and control systems, which can be difficult to communicate in an application.[3]

15.5 Development Considerations for Continuous Manufacturing

15.5.1 Process Development

Pharmaceutical companies can use a variety of manufacturing strategies in developing continuous processes for drug substance and drug product manufacture. Possible options include the following:

1. A fully continuous process where all drug substance and/or drug product unit operations are sequenced together to form a single production line
2. A fully continuous process as above but with two or more production lines in parallel
3. A "hybrid" of batch and continuous mode unit operations

A continuous manufacturing process emphasizing key design and control aspects would be described in sufficient detail in regulatory submissions similar to traditional/batch manufacturing processes.

The regulatory submission could include a general description of the overall manufacturing strategy. This general description could consist of a brief outline of each unit operation and its mode of operation (i.e., batch or continuous), including interfaces between unit operations, the material flow, proposed flow rate and total process operation time, critical process parameters (CPPs), and their ranges and IPC points.

The pharmaceutical development section of the regulatory submission can also include information specific to the development and modeling of the continuous process. These aspects may include residence time distributions, system dynamics, disturbance propagation, information on model setup, maintenance, and model improvement. This information would need to be generated during process development. In general, the regulatory submission for a continuous process will usually contain more information than a batch process because of the use of models, feed forward or feedback controls, and *PAT*.

[3] US FDA ORA Field Management Directive No. 135. http://www.fda.gov/ICECI/Inspections/FieldManagementDirectives/ucm096042.htm.

The definition of a batch or lot has significant regulatory implications, particularly with respect to current good manufacturing practices (cGMPs), product recalls, and other regulatory or enforcement actions. Although the definition of a batch or lot could differ for individual continuous manufacturing operations, the underlying regulatory expectation is that the batch or lot is of "uniform character and quality within specified limits." The manufacturing process should include a clear definition of a batch or lot.

Additional considerations for the continuous manufacturing process include the following:

1. Flow rate of material through the process.
2. Factors affecting the "scale" of the continuous manufacturing process. For example, "scale out" plans (i.e., multiple lines operated in parallel considered to be the same lot), flow rate ranges, and operation time ranges.
3. In-process control (IPC) points.
4. Start-up and shutdown procedures will need to be developed that include provisions on when the process is producing quality product so collection can begin and when the collection should be stopped. Acceptable hold times during shutdowns and "pauses," as well as restart procedures, will be needed. Unexpected occurrences such as power outage, fire alarms, equipment breakage should be considered.
5. Control systems integral to the control strategy: for example, feedback or feedforward controls utilized for maintaining a state of control in the system or automated valves used for rejecting materials deemed to be out of specification and diverting nonconformance materials will need documented procedures.

15.6 Control Strategy

The same regulatory requirements apply for continuous manufacturing as for batch manufacturing, specifically in that a control strategy should be developed that ensures that the manufacturing process produces product of intended quality in a reproducible way. Similar to any other mode of manufacturing, a control strategy is unique for different products and manufacturing processes. A control strategy developed for a batch process may not be appropriate when the same unit operation is operated in continuous mode. Therefore, the control strategy should be reexamined if a unit operation that was operated in a batch mode is replaced by a unit operation in a continuous mode.

Aspects unique to a continuous operation should be assessed in developing the overall control strategy of a continuous process. As material flows through the system and product is formed continuously over a long period of time, the process, product, or environmental conditions could potentially vary over time, resulting in a product of variable quality. A robust control strategy is essential to ensure the consistent quality of products formed over the total operation time.

15.6.1 Special Considerations for Control Strategy in Continuous Manufacturing

Some aspects to consider in establishing the control strategy for a continuous process are listed below.

State of Control
A continuous manufacturing process maintaining a state of control provides assurance that the desired product quality is consistently met. There may be situations such as sudden or uncontrolled changes in a process variable, such as at start-up and shutdown, where assurance is needed that the product is homogeneous and of acceptable quality. However, the process is expected to reach and maintain a state of control after some time. Start-up, shutdown, and transient states should be considered. The control strategy could establish criteria for determining that the process is under a state of control and procedures for handling process start-up, shutdown, or process variables change. The control strategy could include appropriate process attributes or ranges for monitoring or a multivariate process control approach. The ability to detect process upsets and institute corrective actions to bring the process back into conformance, such as feedback control, help ensure the consistency of a continuous manufacturing process over the production time.

Raw Materials and Intermediates
Continuous processing may require additional raw material control, if multiple lots of a raw material are used in a single continuous manufacturing batch. Control approaches should be based on product and process understanding and may include the use of PAT tools. The determination of the characteristics of an intermediate product that may or may not be isolated may be more difficult in a continuous process due to the limited sampling ports and high sampling frequencies. The quality attributes of raw materials and excipients should be linked to the product CQAs and the needs of the process.

Equipment
It is important to consider equipment control aspects for continuous processes. Equipment such as chemical reactors, loss-in-weight feeders, twin screw blenders, extruders, and tablet presses may need to run for long periods of time and may require special maintenance, calibration, and periodic review to ensure their performance. Cleaning needs should also be considered, especially for processes that might be susceptible to microbial growth, blinding or residual buildup.

Uniform Quality and Character of Product
The criteria for determining that the product manufactured is of uniform quality and character, the robustness of the process to produce product of desired quality in the presence of variability, and the ability of the system to detect nonconforming product should be established.

Product Collection or Rejection
Although the continuous process is expected to maintain a state of control, there may be temporary process upsets or disturbances over the total operation time. There may be situations where product made during the disturbance is removed while the remainder of the product is retained. Other situations may warrant rejection of the entire batch instead of a portion of a batch. Establishing a priori criteria, based on the level of risk, for product collection, product rejection, rejection of an entire batch, and indicating how or who makes those decisions prevent ad hoc decisions by manufacturing personnel and help to ensure the desired quality and consistency of the collected product. The disposition strategy of products obtained during start-up and shutdown should also be established.

Traceability
Traceability of incoming materials to the final product should be understood and documented. Traceability can be supported by data on system dynamics such as residence time distributions and back mixing. Planned disturbances such as feeder refills and how those disturbances propagate through the system should also be considered.

Process Monitoring and Sampling
The purpose of the monitoring system is to manage planned changes and respond to unplanned disturbances. Samples (whether analyzed in line or extracted) should be representative, and the frequency of measurement or sample acquisition time should consider material flow rate, system dynamics, and unit dose. Potential failure modes of any sampling devices employed should be understood. Flow rate, frequency of sampling, time constants, and residence time distributions – all of these have an impact on how to test for quality at any point (raw material attributes, IPCs, final quality) and also how to achieve feedback and feed forward control.

Consideration can be given to define a flexible monitoring frequency where more monitoring is expected in periods where there is a greater risk of variability (e.g., following addition of a new lot of input material, or following process parameter adjustments based on feed-forward/ feed-back loops).

Risk Assessment and Failure Modes
Process robustness is an important factor for consistent operation of a continuous process, which in turn helps to ensure that the product formed is of uniform quality and character. A thorough understanding of the risks and failure modes of the process and its associated measurement and control systems allows the development of effective risk mitigation strategies and helps support manufacturing changes and process improvements that may occur over the life cycle of a product. Knowledge of risks and failure modes is also useful to making risk-based decisions.

Scale-Up
Scale-up can be achieved in several ways, including running a longer time, increasing throughput, or parallel units (scale out). Increasing throughput at fixed size units can have effects on residence time distributions and time constants. These effects

should be considered during development. Representative sampling may be affected. Physical and physicochemical conditions may be affected by throughput, and criticality of parameters may change following a change in throughput.

Specifications
Specifications will be required as part of the control strategy. Continuous processes may include an RTRT approach for some quality attributes, but it is possible to foresee traditional end-product testing on off-line samples. RTRT approaches may require an enhanced sampling plan compared to traditional release testing and may involve a larger sample size.

15.7 Stability Considerations for Continuous Manufacturing

Regulatory requirements for having adequate stability data do not change between batch manufacturing and continuous processing. There are some differences that should be considered when developing the stability plan.

15.7.1 Representative Stability Batches

Since scale may not be a significant risk to stability when using continuous manufacturing, deciding how to determine a representative batch may be different than when using a batch process. Current ICH stability criteria for process scale-up provide a good starting point for consideration. A risk assessment can be used to understand the potential risks of the proposed lot sizes and determine a representative lot that has similar characteristics to the lots being manufactured.

Stability data should fulfill the initial filing requirements where data are generated on the critical to quality and stability-indicating attributes on representative batches. Representative batches for annual stability requirements will also need to be defined.

15.7.2 Stability Considerations at Scale-Up

The risks arising from batch scale-up are different for continuous manufacturing. These potential risks may include heat buildup over time, material buildup in the equipment, events occurring at the interface between unit operations, different surface area/volume rations, and other risks that may be product specific. These risks

are usually easily manageable for continuous manufacturing so change in scale might not need additional stability testing.

A change in scale for a continuous process could include volume, time, and/or multiple manufacturing trains that run in parallel. Each of these has its own potential risks to stability that should be considered when determining the type of stability testing needed to assess the impact, if any, of scale-up on product stability.

15.7.3 Stability Considerations for Site Change/Technology Transfer

When transferring a continuous process from one site to another, the impact on stability should be considered. Considerations can include equipment changes, scale changes, and potential location impact such as different raw material suppliers, and/or different environmental conditions.

15.7.4 Location of Information in Regulatory Submissions

The development of a continuous manufacturing process is likely to include information obtained from enhanced process development approaches. ICH Q11 and ICH Q8(R2) recommend that process development information be submitted in section 3.2.S.2.6 of the CTD for drug substance and 3.2.P.2 (Pharmaceutical Development) of the CTD for drug product. The three guidance documents also contain specific suggestions for the provision of information from development studies.

In general, the recommendations of ICH Q11 and ICH Q8(R2) could be adopted for the placement of information supporting a continuous manufacturing application. ICH Q11 and Q8(R2) recommend that the control strategy information be summarized in the specification sections, 3.2.S.4.5 and 3.2.P.5.6, for the drug substance and drug product, respectively. The ICH Q8 and ICH Q11 suggestions for the placement of information in a regulatory filing could also be used for continuous manufacturing applications. The US FDA draft guidance, Quality Considerations for Continuous Manufacturing, has a section discussing the location of information in an application, which suggests CTD locations for aspects of continuous manufacturing information such as residence time distribution, raw material concerns, and control strategy description. As in other regulatory submissions, the applicant could clearly indicate where the different information is located in the application. Similar to batch processes, certain aspects of the control

strategy are handled under the applicant's pharmaceutical quality system (see ICH Q10).

15.8 Quality/GMP Considerations

Current good manufacturing practices do not preclude the introduction of new manufacturing technologies and embody the flexibility needed to support continuous manufacturing. Points to consider when implementing continuous manufacturing in a cGMP environment are noted below.

15.8.1 Pharmaceutical Quality Systems (PQS)

To implement continuous manufacturing in an existing PQS, a site should evaluate its PQS and associated elements to determine if the design and content of the PQS should be modified. In addition to the areas described in Q10 (e.g., pharmaceutical development, manufacturing, quality, regulatory affairs, and medical), the manufacturing site should establish continuous manufacturing expertise in the quality organization. As with traditional batch processes, change management for continuous processes should incorporate an assessment of potential risks.

15.8.2 Batch Release

Current regulatory cGMP guidance considers a "batch" as a defined quantity of product processed in one process or a series of processes so that it is expected to be of uniform character and quality, within specified limits, and produced according to a single manufacturing order during the same cycle of manufacture. This principle applies equally to continuous processes where the amount of material subject to a quality disposition decision could be defined as:

1. Process time when all of the material is discharged from the process between two specific time points
2. Product quantity when a specific quantity of material is produced
3. Process event when all of the material is produced between two specific process events
4. Raw material quantity input when all of the material that is "intended" to contain a specific lot or quantity of a specified input material

If a continued state of control operation is demonstrated, it could be possible to designate large quantities of material as "homogenous" even though different lots of

raw materials and different processing conditions may have been used. The key is to have clearly defined criteria, which describe state of control operation, and to establish the product and process data, which demonstrate continued conformance with these criteria.

Continuous manufacturing within a controlled and reproducible operation may have periods of process perturbation. Therefore, during development, these perturbations should be considered, and criteria developed, to define the state of control. Procedures should be established to define a process of diversion of waste should adverse perturbations occur.

Material traceability is vital in the event of problems with product quality, such as contamination, raw materials, recalls, or other GMP failures. An understanding of material flow in the system is essential to track or divert the potentially affected material. Close monitoring of product and process data could support further decisions to be made, such as bringing the process back to target via process control measures, which could help minimize the impact of process failures.

15.8.3 Start-Up and Shutdown Procedures

During periods of start-up, shutdown, and processing of material, it is possible that not all unit operations within a continuous production line will be in a state of control at the same time. For example:

1. During shutdown, material may not be fed and discharged simultaneously. Material will continue to be processed and discharged after the feeding operation has stopped.
2. Where small amounts of material are produced, the first unit operation could already be shut down while the material is processed further.
3. The start from which onward material gets collected for later processing can be defined relative to the time point when the process is in a state of control, and significant process parameters, in-process material attributes, CPPs, and CQAs are all within their specified criteria, needs to be defined. The same determination can be made for shutdown periods and for transient adverse perturbations requiring material diversion.
4. The ability of the process to reach and detect the period of normal production is normally demonstrated during process verification.

The time available for a given process transformation can be determined by the residence time of the material in a specific process environment and can be used to accommodate the necessary reaction time for completion. As the material flows through the system, rate-limiting elements within the process can be considered to ensure that the end-point conditions are met within the time available (e.g., required process to complete a chemical reaction or drying operation).

An understanding and subsequent verification of the various time constants of the process is specifically important in determining the expected behavior of the process during start-up and shutdown and hence the impact on quality decisions regarding the disposition of material manufactured during this period.

15.8.4 State of Control: Product Collection and In-Process Sampling

A state of control provides assurance of continued process performance and product quality, as described in ICH Q10.

Acceptance criteria based on monitoring at an adequate frequency can be established to ensure that the entirety of the material subjected to the release decision is compliant with the applicable specifications. Decisions of diversion and/or rejection of material that does not meet acceptance criteria should be based on reliable data and an understanding of process dynamics.

Consideration should be given to confirm the ability of the system to produce consistent product over extended operation and to understand potential mechanisms of failure and degradation of performance, together with the methods of detection. Risk analysis techniques, including practical tests and/or modeling tools, can be employed to ensure that any impact on product quality is understood and appropriately managed over all operating states, including during normal operations.

In order to define the period of product collection, the process residence time and residence time distribution are typically quantified during start-up and normal operation conditions, as well as during shutdown conditions. In particular, an understanding and quantification of the residence time distribution can be used to determine which material may have been affected by a deviation in process conditions and hence the range of product within the scope of any investigation or disposition decision.

In some cases, the maximum length of time over which the process is run could be determined by monitoring specific product attributes or process parameters and equipment capability rather than by validating a single fixed length of run time.

Sampling, testing, quality control procedures and equipment mechanisms to detect and reject materials that are out of specification are typically necessary. In order to ensure that a process parameter or product attribute cannot move outside the predefined acceptable process window or acceptable range without being detected, it is important to ensure that the control and monitoring system is able to take measurements at a frequency that is appropriate to the dynamic response time of the parameter or attribute. Measurement frequency considerations include the intrinsic process risk (e.g. higher risk at lower API dose), the known process variability, and the required residence time in the equipment to complete transformation (e.g., higher risk at faster throughput).

15.8.5 Process Validation and Continuous Process Verification

Within process validation and continued process verification, process robustness and reproducibility are typically evaluated. The development of a continuous process can follow established principles that are applied generally within pharmaceutical process development. Existing guidance and standards can also be consulted for process validation and, when applicable, continuous process verification (CPV). In general, the requirements for including process validation and life cycle management information in the regulatory submission can be expected to be the same as that for batch processes.

In verifying the ability of the system to control and achieve the specified performance, the following should be considered for continuous processes:

1. Are the process conditions, which determine that the system is under a normal state of control, including verification of CPPs and CQAs, within target range?
2. Is the process control system able to reach and detect the start of acceptable product production? In order to demonstrate this ability as part of the process verification, a set of start-up and shutdown activities could be included. The number of start-ups and shutdowns included in the verification activities could be determined based on a risk analysis for a given process and the unique critical considerations for that process, including process robustness, and the number and interrelationship of CPPs/CQAs.
3. Is the system able to reach and maintain the intended process conditions over the entire process? The expected process run time and worst case (longest) process run time could be considered as a component of the process validation activities.
4. What is the ability to detect excursions from the target CPP or CQA values? These excursions can require the diversion of nonconforming material based on the understanding of process dynamics or shutdown of the process.
5. What is the impact of changes in the process production rate and/or equipment scale changes on the process dynamics?
6. Does the process remain in a state of control (the validated state) during commercial manufacture? This is one goal of the final validation stage.
7. Is the validation data statistically trended and reviewed by trained personnel? The information collected can help verify that the quality attributes are being appropriately controlled throughout the process.

   (a) Quantitative, statistical methods are recommended whenever appropriate and feasible.
   (b) Scrutiny of intrabatch as well as interbatch variation should be considered.

CPV as an alternative validation approach may be particularly well suited to the evaluation of continuous manufacturing processes. It can utilize in-line, on-line, or at-line monitoring or controls to evaluate process performance. These are based on product and process knowledge and understanding. Monitoring can also be combined with dynamic control systems in order to adjust the process to maintain output quality. This capability also provides the advantage of enhanced assurance of intra-

batch uniformity, fundamental to the objectives of process validation. Some process measurements and controls in support of RTRT can also play a role in CPV.

Using this approach, data from production batches can often serve to validate the process and demonstrate processing in accordance with the total system design concept, essentially supporting validation with each manufacturing batch replacing a conventional process validation approach (e.g., three-batch validation at set point) that was historically used.

As with traditional batch manufacturing, system qualification of commercial equipment and other supporting systems, including PAT and/or automation, is necessary. This may be especially critical if some systems are providing real-time monitoring and control of a continuous manufacturing process.

15.8.6 Material Traceability in Continuous Manufacturing

For any specific quantity of product produced from a continuous processing system and released to the market, it may be necessary to reliably link the relevant process information to the specific quantity of product in a timely manner and to identify the lots of raw materials from which it has been manufactured. This typically includes an understanding of residence time and residence time distribution at relevant flow rates and operating conditions. An appropriately reliable and timely link between relevant product quality information and any specifically identified product can be decisions related to later release, including considerations such as diversion of unacceptable product during the process.

Other traceability considerations include the overall flow of product in the system or subsections of the system and the ability to account for material that may be removed deliberately from the system for sampling or unintentionally lost from the system due to unforeseen events or diverting.

15.8.7 Handling of Raw Material and In-Process Material

Continuous processing may pose challenges due to behaviors of both the equipment and material, for example, starting materials in a hopper, or intermediates in process that occur gradually over a long period and that are not easily observed during batch processing or short tests runs. Determining the handling and flow properties of materials to be processed as early as possible within the development of the product can greatly aid in the appropriate design of the process equipment. Transport processes may cause some degree of transformation (e.g., segregation and/or attrition of powders), and therefore careful consideration should be given.

Risk analysis, practical tests, and modeling techniques should be considered in order to determine and evaluate potential challenges in maintaining stable process conditions during the operation of a continuous process over the full length of the production run. Consideration should be given to the potential for undesirable

buildup of material due to physical and chemical processes, stability of starting materials, or intermediates being held in buffer tanks.

15.8.8 Detection and Treatment for Nonconformity

A key consideration in any quality system is handling nonconformities and/or deviations. While process and product understanding are extremely important, unexpected discrepancies will undoubtedly occur during the product life cycle. These issues may raise questions regarding the existing process and product understanding and may require additional process development. A robust corrective action and preventive action (CAPA) system is integral to product and process improvement. The methodology used can aid in product and process improvements and enhanced product and process understanding.

There are some key elements for consideration in a continuous manufacturing process. The process control/monitoring system should be adequately developed to recognize a normal process and be able to identify when the data are divergent enough to represent a departure that could have a direct impact on quality. In these cases, the product should be diverted for further determination and potential rejection/waste. As a consequence, it is possible that not all of the materials that were originally fed into the process, as part of the original single manufacturing order, will be in the finished product intended for release to the market.

Continuous manufacturing could also have more complex in-process controls and monitoring that could include the potential for unintended failure modes. Such failure modes should be considered in setting a robust control system.

Handling of nonconformities for continuous manufacturing and batch manufacturing is generally similar.

Personnel Procedures and Training

In a robust pharmaceutical quality system, when new technology such as continuous manufacturing and PAT tools are implemented, it is important to evaluate the impact, if any, on existing quality, production, and engineering procedures. Procedures that define who is responsible for halting and resuming operations and how nonconformities are documented, investigating discrepancies, and taking remedial action may need to be modified based on the new technology. New procedures and/or modifications will typically require additional personnel training.

Material Carryover

It is important to ensure that any investigations are properly extended to other batches of the same drug product as appropriate. Thus, understanding how the facility defines a batch is critical to ensuring that the investigation is properly

extended to related batches. The amount of allowable carryover volume should be considered.

Material Diversion

Establishing thorough procedures to describe the handling of nonconformances, including out-of-specification or out-of-trend results that requires product stream diversion during manufacturing, is critical. Procedures describing when the product stream is to be diverted and when collection is to be reinitiated should be established prior to the occurrence of nonconformance. If a nonconforming material is detected, it should be diverted at the next appropriate point. If the in-process monitoring detects that a certain amount of material needs to be diverted, this diverted material should be investigated before determining the disposition of the batch.

Production Floor Product Monitoring

PAT tools are more likely to be implemented in continuous manufacturing processes on the production floor (in-line, at-line, or on-line). As with batch manufacture, if a discrepancy is identified on the production floor, it should be investigated prior to material disposition. For example, if in-line testing results are trending toward failure, end-product testing is not sufficient to release associated material; an associated investigation should be performed.

Breakdown in the monitoring equipment should also be investigated. A procedure can often be established for the use of alternative testing or monitoring approaches in cases of equipment failure. The alternative approach could involve the use of end-product testing or other options while maintaining an acceptable level of quality.

Raw Material Variability

For continuous manufacturing processes, it is important to consider raw material variability as a potential root cause when performing an investigation. In a batch process, multiple raw material batches are typically mixed at the start of manufacturing. This may not be true for continuous manufacturing, where different lots of raw material can be used during the production campaign. Multiple raw material lots used in a single product batch, though they might meet specification, could introduce variability into the finished product.

15.8.9 Cleaning Validation

Cleaning and cleaning validation considerations for continuous manufacturing equipment and systems are primarily the same as those for noncontinuous manufacturing equipment and systems. Since continuous unit operations may be smaller than those in batch processing, the surface area to volume rations will differ. For continuous manufacturing, either dedicated or nondedicated equipment could be utilized. Principles for determining the acceptance criteria for cleaning agents, bioburden, endotoxin, and degradation products for cleaning validation of dedicated equipment are essentially the same as for nondedicated equipment.

If dedicated equipment is utilized for continuous manufacturing, cross-contamination of the active ingredient from the previous product to the next product is not an issue. Therefore, cleaning validation related to the active ingredient itself is generally not considered a requirement for dedicated equipment. However, cleaning validation should be considered for dedicated equipment if carryover of the cleaning agent or the contribution of bioburden or degradation by-products to the next manufactured batch is a concern. Manufacturers should conduct risk assessments for all cleaning scenarios in order to ensure that manufactured product using that equipment complies with quality expectations (e.g., residue removal, lot integrity, and regulatory expectations). It is considered to be best practice to document effectiveness of a cleaning process for dedicated equipment even if "visually clean" is the only criteria.

The cleaning process and frequency of cleaning should be defined and the effectiveness verified periodically.

Considerations for the design and verification of the cleaning process include the following:

1. Material holdup and buildup on equipment, piping, instruments (e.g., on-line analyzers, sensors), filters
2. Degradation of the material within the process
3. Microbiological growth
4. Formation of chemical films
5. Cleaning agent removal, if applicable
6. Product changeover, if applicable
7. Equipment size and complexity (e.g., equipment used for continuous manufacturing may be smaller in size and may have more intricate parts and components that may be more difficult to clean)

The cleaning frequency for continuous manufacturing equipment and systems could be defined in terms of the following:

1. Elapsed operating time
2. Quantity of material processed
3. History of process conditions or deviations
4. Product changeover, if applicable

Cleaning strategies employed for continuous manufacturing can include the following:

1. Stopping production or diverting or tagging material as nonreleasable material (e.g., if reliant on a single analyzer that requires attention)
2. Providing a second, duplicate piece of equipment or instrument (e.g., analyzer)
3. In-process cleaning of instruments such as sensors (e.g., air washes)

15.8.10 Equipment Failure

Continuous processing may pose challenges due to degrading performance of the equipment, which could occur gradually over a long period and hence might not be easily observed during batch processing or short tests runs. The control system should consider how equipment performance is to be detected and addressed.

Risk analysis, practical tests, and modeling techniques can be beneficial to determine and evaluate potential challenges in maintaining stable process conditions during the operation of a continuous process over the full length of the required production run.

Control systems can be designed to address where one unit operation within a process line is determined to be disproportionally vulnerable (for example, due to degradation or lack of robustness) or prone to equipment failure. Such considerations could include rapid change over or redundancy/parallelization/duplication of critical equipment elements.

15.9 Quality Considerations for Bridging Existing Batch Manufacturing to Continuous Manufacturing

Situations could arise where a continuous manufacturing process is proposed after using a different process, such as a batch process, to make the clinical, bioequivalence, or registration stability batches. A company could also wish to introduce a continuous process as a postapproval manufacturing change. In such situations, a risk-based approach can be useful to determine the type of bridging information to support continued product quality.

A change from batch to continuous processing is likely to result in changes to equipment, process parameters, control strategy, and facility or manufacturing area. A comparison of the two processes and the input materials (or formulation) can be a starting point for assessing the risk of the process change. In addition to differences within individual unit operations and equipment, a holistic assessment of the overall processes can be beneficial to evaluate risks due to differences in system dynamics.

Factors such as dosage form, strength, drug load, potency, release profile, and route of administration can also impact risk to product quality. For example, the risk of a high drug load immediate-release tablet is likely to be less than that of a low drug load extended-release tablet. A discussion of the proposed change and the bridging strategy with the respective regulatory agency is advisable to gain agreement prior to conducting the studies.

Some aspects to consider when bridging batch and continuous processes are discussed below.

15.9.1 Physicochemical Equivalence Considerations

A change from batch to continuous manufacturing typically necessitates establishing physicochemical equivalency. Evaluation of a change from batch to continuous operation could include a comparison of individual unit operations, process parameters, equipment, CQAs, and control strategies. Comparative batch data, particularly with respect to physical properties, impurity profiles, and drug release profiles, and bridging stability data can be helpful to support chemical equivalency.

15.9.2 Bioequivalence Considerations

When the risk of change to product quality attributes (e.g., polymorphicity, dissolution, impurities, stability, etc.) is low, demonstration of chemical equivalence could be sufficient to support the change from batch to a continuous process. In many instances, this could be when the continuous process is based on the same unit operations and formulation as used for the batch process. However, there could be situations, albeit rare, when significant changes or novel approaches are used in moving from batch to a continuous process. For example, the continuous process could incorporate a novel crystallization method that changes crystal form or a significant formulation change. Also, the drug product characteristics (e.g., a dosage form with a complex release profile or a very low drug load) may need to be considered in evaluating risk. Significant changes or high-risk products are more likely to require bridging by bioequivalence studies.

15.9.3 Batch Definition

An important aspect of continuous manufacturing is the definition of a batch. There are specific references to “batch” and “lot” in the US Code of Federal Regulations, which are applicable and need to be considered.

21CFR 210.3

A *batch* is defined as a specific quantity of a drug or other material that is intended to have uniform character and quality, within specified limits, and is produced according to a single manufacturing order due the same cycle of manufacturer. Therefore, batch refers to the quantity of material and does not specify the mode of manufacture.

Lot – a batch, or a specific identified portion of a batch, having uniform character and quality within specified limits; or, in the case of a drug product *produced by continuous process*, it is a specific identified amount produced in a *unit of time or quantity* in a manner that assures its having uniform character and quality within specified limits.

The 21 CFR definitions for both "batch" and "lot" are applicable to continuous manufacturing.

21 CFR 211

Documentation of Manufacturing 21CFR 211.188

- Batch product and control records shall be prepared for each batch of drug product produced and shall include complete information relating to the production and control of each batch.

While the regulations in the CFR provide flexibility in the area of documentation, the definition of a batch/lot at collection is not specifically described.

ICH Q7

A Batch or Lot is defined as:

> A specific quantity of material produced in a process or series of processes so that it is expected to be homogeneous within specified limits. In the case of *Continuous production*, a batch may correspond to a defined fraction of the production. The batch size can be defined either by a fixed quantity or by the amount produced in a fixed time interval

The batch definition is a fundamental element of continuous processes. It is linked to in-process testing, specifications, batch disposition, and many critical aspects of cGMP compliance.

ICH Q10

State of Control: A condition in which the set of controls consistently provides assurance of continued process performance and product quality (ICH Q10).

Bibliography

Allison G, et al. Regulatory and quality considerations for continuous manufacturing May 20–21, 2014 continuous manufacturing symposium. J Pharm Sci. 2015;104:803–12.

ICH Quality documents Q3, Q7, Q8 Q8(R2), Q9, Q10, Q11

FDA Guidance for Industry PAT A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance

cGMP Guidance

Woodcock J. FDA, AAPS annual meeting. Oct 2011.
http://www.fda.gov/ScienceResearch/AboutScienceResearchatFDA/ucm342928.htm
PhRMA White Paper Implementation and Application of Quality by Design Feb 2013.
http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/02/WC500162136.pdf
EU Guidelines for Process validation.
EU Guideline on NIR.
EU Guideline on Real Time Release testing.
ASTM Standards E2537, E2968, and E2474.
21CFR210.3.
21CFR211.118.
2011 FDA Guidance for Industry: Process Validation: General Principles and Practices.
2019 FDA Quality Considerations for Continuous Manufacturing Guidance for Industry.

Index

A
Adiabatic saturation temperature, 307
Advance angle, 279
Advanced modelling, 319
Alexanderwerk, 233, 234, 239, 241, 249, 252, 253, 260
Analytical, 229, 235, 249
Angle orientation, 279
Antifouling control (AFC), 165
Antiretroviral, 485
API, 228, 242, 258, 263
Artemisinin, 485
Auger, 228, 231, 232, 234, 239, 241, 246–248, 253
Avicel, 244
Axial dispersion, 52
Axial mixing, 276

B
Back pressure regulators, 97–101
Batch crystallization, 141
Blend, 228, 242–245, 247, 254, 255, 259, 260
Bohle, 249, 252, 253
Boundary conditions, 230, 232, 233, 235, 240, 256
Breakage, 279, 291
Brittle, 243–245, 254, 256
Bubble flow reactor, 30
Bubble formation, 317
Bubble phase, 312, 314
Bubble rise velocity, 318
Business case, 478, 479, 494, 497, 504, 509

C
C-curve, 53
Channel fill, 275
Cheekplates, 231, 233, 234, 252, 253
Circumferential, 231
Classical modelling, 309
Comb mixing elements, 282
Compacted, 228, 229, 231, 242, 244, 252, 254–256
Compartmental population balance modelling, 290
Compressibility, 237, 241, 242, 244, 245, 261
Compressible, 237, 254
Compression, 235, 243–247, 252, 255–257, 259
Computational, 229, 249
Concentration control (C-control), 412
Consolidation, 292
Constitutive, 230, 235
Content uniformity, 194
Continuous crystallization, 142
Continuous fluidized bed drying, 303–354
Continuous improvement, 408
Continuous processing, 75, 92, 478, 480, 482–487, 490, 494, 504, 506
Continuous stirred tank crystallizers, 88–94
Continuous stirred tank reactors (CSTRs), 35
Control strategy, 513
Conveying elements (CE), 271
Critical moisture content, 307, 336
Critical process parameters (CPP), 371
Critical quality attributes (CQAs), 258, 259, 371

Z. K. Nagy et al. (eds.), *Continuous Pharmaceutical Processing*,
AAPS Advances in the Pharmaceutical Sciences Series 42,
https://doi.org/10.1007/978-3-030-41524-2

Crystalline, 243, 245
Crystallization, 132, 410
Current state, 483, 485, 496, 497, 506, 507
Cytotoxic API, 37

D
Data reconciliation (DR), 204, 408
Deaerator, 249, 250
Deformation, 232, 244
Demands, 229
DEM/FEM, 235
DEM-PBM, 295
Density, 229–235, 237, 240–242, 244, 247, 251, 252, 254–260, 262, 263
Density distributions, 231–235, 248, 252, 262
Design of experiment (DoE), 258–261, 263, 371
Detection, 437–438
Diagnosis, 437–438
Dibasic calcium phosphate (DCP), 243–245
Die filling, 228
Digital, 478, 482, 489
Digital twins, 509
Dimensional, 229, 240, 260
Dimensional analysis, 286
Dimensionless, 260
Direct compression, 415
Discrete element method (DEM), 234, 278
Disintegrants, 228, 243
Dispersive mixing, 276
Disposable equipment, 37
Distributed, 486, 498, 505, 506
Distributive missing elements (DME), 271
Downstream, 230, 231, 235, 241, 258
Drucker-Prager/Cap (DPC), 232–235
Drug substance, 243
Drying curve, 306, 331
Drying kinetics, 315, 333
Drying of mixture of two solids, 351–353
Drying rate, 307
Ductile, 243, 255
Dwell time, 255, 256, 263

E
Effective friction, 236
Effective friction angle, 241
Elastic, 235
Elastic recovery, 256
Entry angle, 239
Environmental assessment, 495, 496
Excess gas velocity, 318
Excipients, 228, 243–245

F
Factorial, 229, 258, 259
Fatty acids, 245
F-curve, 53
Feed auger, 235, 241, 242, 247–250, 253–256, 259
Feed design, 247
Feeder, 233, 234, 247
Feeding, 232, 246, 247, 251, 255
Feeding-blending system, 195
Feeding system, 242
Feed system, 247
FEM, 230, 232–235, 240, 246
Fillers, 228, 243–245, 255
Filtration, 177
Fine, 228, 231, 244, 245
Fine fraction, 244–246, 252, 254, 257
Finite element, 230–232, 240, 241, 256, 262
Fitzpatrick, 249
Float, 234
Floating, 250, 253, 254
Flow, 230–232, 235, 237, 242, 246, 247
Flowability, 228, 243, 245, 246, 257
Flowsheet of continuous fluidized bed, 324
Flowsheet simulation, 494, 496, 506
Fluid dynamics of fluidized bed, 317
Fluidize, 248
Fluidized bed drying, 304
Force, 228, 230, 232, 233, 235, 242, 244, 246–250, 252–254, 257, 259, 263
Formulations, 242–246, 260
Friction, 228, 233, 236, 241, 246, 252
Friction angle, 230, 231, 233, 234, 240, 242, 250
Future state, 483, 485, 496, 497, 506, 507

G
Galerkin, 230, 232
Gap, 228, 233, 240, 253–255
Gap control, 246, 253, 255, 259
Gap width, 246, 247, 252, 253
Gas cylinders, 449
Gerteis, 234, 239, 248–250, 252, 253, 257
Granular, 235
Granulation, 228, 255, 258, 262, 263
Granulators, 271
Granule growth, 277
Granules, 228, 229, 231, 244–246, 254, 255, 257, 259, 260, 262, 263, 282
Granule size distribution (GSD), 273
Grignard, 41
Growth, 135

H
Heat and mass transfer, 306
Heterogeneous, 235, 248
High-pressure hydrogenation, 100
High-shear blender, 199
High shear wet granulation, 272
Homogeneous, 246
Hopper, 228, 247–249
Hydraulic, 250, 253
Hydroformylation, 27

I
ICH, 258
Imaging camera, 283
Imidazole cyclization, 24–27
Inlet, 230–235, 238–240, 248–250, 253, 254, 257, 262
Inlet moisture distribution, 350
In-line probe, 283
Interlocks, 470
Intermittent flow, 39, 88
Interstitial, 248, 249
Intragranular, 255
Inventory, 478, 484, 486–488, 506, 507, 509

J
Jenike-Shield, 236, 237, 239
Johanson, 229–231, 235–242, 251, 256, 257, 261, 262

K
Kinematic, 231, 232
Kinetics, 315
Kinetics of drying, 308
Kneading elements (KE), 271

L
Laboratory fume hoods, 450
Lactose, 243–245
Laminar, 63
Layering, 279
Leakage, 246, 248, 249, 252
Liquid distribution, 272, 274
Load cell, 232, 241, 242
Location of Information in Regulatory Submissions, 523
Loss-in-weight (LIW) feeder, 197
Lovastatin, 485, 491, 496, 497
Low-shear blender, 199
Lubricant, 228, 243–246
Lubrication, 244–246

M
Magnetic suspension balance, 326
Major principal stress, 236
Mannitol, 243, 252, 254, 255
Manufacturing, 486, 498, 505, 506
Mass and heat transfer, 316
Mass correction factor, 241
Material holdup, 275
Materials of construction (MOC), 466
Mechanical strength, 243, 245, 246
Mechanistic, 229, 242, 246, 261
Mesoscale plug flow crystallizer, 159
Metallic salts, 245
Metastable zone width (MSZW), 134
Metformin, 485
Microcrystalline cellulose (MCC), 243–245, 252, 254, 255
Micro-factory, 498, 505–509
Mill, 228, 229, 231, 242, 254, 256–259, 263
Mini-Pactor, 234, 239, 248, 253, 257
Mitigation, 437–438
Mixed-suspension-mixed-product removal (MSMPR), 89, 143–157
Model application, 321
Modelling, 309
Model predictive control (MPC), 213
Model validation, 351
Mohr's circle, 236
Moisture content, 284
Multivariate, 260

N
Near-infrared (NIR), 262, 283
Net Reynolds number (Re_n), 161
Network configuration, 483, 504–507
Network design, 488, 504, 506
Nip, 240, 241, 243, 246, 247
Nip angle, 230–232, 235, 237, 238, 240, 241, 246, 250, 251, 254, 261
Nipped, 233, 241
Nipping, 232, 233
Nonuniform stress distribution, 235
Normalized drying curve, 331, 337
Normal stress, 230, 237–240, 257
Normal stress distribution, 230, 231, 235, 238, 240, 241
Normal stress-volume, 237

No-slip, 237–239, 241
Nucleation, 134, 290
Nucleophilic aromatic substitution reaction (S_NAr), 43

O
Offset angle, 275
One-dimensional, 229–231, 237, 262
Oscillatory baffle reactor (OBR), 153
Oscillatory flow baffled crystallizer (OFBC), 160
Oscillatory Reynolds number (Re_o), 154
Over-lubrication, 244–246

P
Paracetamol, 491, 494–500, 506
Partial least squares (PLS) regression, 364
Particle image velocimetry (PIV), 278
Particle size, 243, 244
Particle size analysis, 283
Peclet number, 65, 276
Pharmaceutical Quality Systems (PQS), 524
Pharmaceutical supply chain, 480, 484, 487, 507
Phenomenological, 235
Plastic, 232, 243, 244, 255
Plug flow crystallizers (PFCs), 158
Plug flow reactors, 24
Plugging/Fouling, 446
Population balance, 305, 314
Population balance model (PBM), 138, 156, 288
Porosimeter, 232
Porosity, 282
Pre-compression, 249
Pre-consolidation, 260
Predensification, 250, 254
Pressure, 243–246, 250, 254–256
Pressure-driven transfer, 147
Pressure relief devices, 444
Pressure testing, 444
Principal component analysis (PCA), 364
Process analytical technology (PAT), 12, 167, 202, 262
Process condition model, 15
Process control, 210
Process control systems, 13
Process intensification, 178
Process validation, 513
Proportional integral derivative (PID), 242, 247, 250, 253, 255, 261

Q
Quality by Control (QbC), 398
Quality by Design (QbD), 258, 396
Quality by Testing (QbT), 395
Quasi-static, 235, 256

R
Raman spectroscopy, 283, 368
Real-time optimization (RTO), 18, 429–431
Real-time process management (RTPM), 429
Real-Time Release Testing, 18
Recycle, 103
Reductive amination, 30
Regulatory aspects, 513
Representative stability batches, 522
Residence time, 4, 55, 272
Residence time distribution (RTD), 5, 54, 285
Reworkability, 229, 231
Ribbon, 228–235, 240–242, 244–248, 251, 252, 254–260, 262, 263
Ribbon density, 234
Ribbon width, 241
Rimmed roll, 234, 252, 253
Rimmed roll compactor, 252
Roll compaction, 228–263
Roll diameter, 230, 231, 236, 260, 261
Roll dimensions, 241, 253
Roll force, 233, 234, 239–242, 244, 246, 250–256, 259, 261, 262
Roll gap, 228, 230, 233–237, 246, 250, 253–256, 258–261
Roll gap control, 250
Roll gap width, 242, 249, 250
Roll hydraulic pressure, 259
Rolling angle, 239, 241, 242
Rolling theory, 235, 239, 242, 261
Roll orientation, 249
Roll pressure, 239, 250, 253, 254, 258–260
Roll speed, 246, 250, 254–256, 258–260, 262
Roll surface, 230, 235, 236, 241, 246, 250, 251, 253, 261
Roll width, 239, 241, 259, 261
Rotational speed, 241, 250

S
Safety checklist, 444–451
Scale-out, 286
Scale-up, 229, 259–262, 286
Schotten-Baumann reaction, 35
Screw configuration, 276
Screw elements, 272
Seals, 234, 252

Seeding, 138
Semi-continuous filtration, 94
Shape, 282
Shear-elongation, 279
Shear stress, 382
Side seals, 246, 252, 253, 255
Simulation, 347
Single particle installation, 326, 327
Single particle measurements, 336
Single particle moisture measurement, 326
Slab, 229–231
Slip, 230, 236, 238, 239, 243, 244, 246
Solid dosage, 251
Solid fraction, 241, 244, 245
Solid phase balance, 314
SolidSim, 306, 321
Solvent exchange, 95
Sorption isotherms, 331
Spanwise, 231, 233, 234
Springback, 235
State-of-control, 3, 149
Static electricity, 446
Static mixer, 159
Stationary model, 306
Statistical, 242, 259
Stick, 245, 251
Sticking, 246
Stock vessels, 74
Streamwise, 230, 231, 233, 234, 239, 240, 257, 262
Streamwise normal stress, 239
Stress, 230, 232–242, 244, 246, 260, 262
Stress distribution, 232, 235, 242, 261
Stress-density, 237
Stress-strain, 230, 232
Stress-strain constitutive, 231, 232, 240
Superficial fluidizing gas velocity, 318
Supply network configuration, 481, 484, 487, 491, 494, 501
Supply network design, 478, 488, 491, 503
Surge vessels, 75
Suspension phase, 312, 313
Suzuki–Miyaura cross-coupling, 39
Systems integration, 17

T
Tablet, 228, 231, 243, 244, 246, 252, 254–256, 258, 259
Tablet compression, 243
Technology selection, 500
Tensile strength, 243–245, 252, 255, 256, 258
Thermal deprotection, 32
Thermal drying, 304
Thermal expansion, 70
Three-dimensional, 230–234
Torque, 232, 240
Traceability, 521
Tracer, 285
Twin screw granulator (TSG), 271
Two-dimensional, 231, 232, 240, 241
Two-phase model, 312

U
Uniform, 232, 234, 247
Uniform inlet, 232
Uniform stress, 232–234
Uniform velocity, 232
Upper-bound theorem, 231
Upstream, 233, 235, 236, 249

V
Vacuum, 249, 250, 253
Vacuum deaeration, 249
Variographic analysis, 384
Velocity, 230–233, 237, 240, 247, 256
Volume, 237
Volume-variety matrix, 483, 501, 502

W
Water activity, 304
Wet granulation, 271
Width, 231, 233, 234, 260, 263
Work hardening, 244, 245